ISSN 0138 - 3248

Rostocker Mathematisches Kolloquium

Heft 40

Wilhelm-Pieck-Universität Rostock

Information

Rostocker Informatik-Berichte ISSN 0233-0784

Diese Schriftenreihe der Sektion Informatik der Wilhelm-Pieck-Universität Rostock erscheint seit 1985.

Bisher liegen folgende Hefte vor:

Heft 1 (1985)	20 Jahre Rechenzentrum/Sektion Informationsverarbeitung
Heft 2 (1985)	DIGRA'84 (Internationale Tagung, November 1984, Ahrenshoop)
Heft 3 (1986)	Beiträge zur Digitalgraphik und ihren Anwendungen aus Institutionen und Kombinaten der DDR
Heft 4 (1986)	Arbeiten aus der Sektion Informatik
Heft 5 (1987)	Beiträge des Problemseminars "Graphisch-Interaktive Systeme"
Heft 6 (1988)	Problemseminar "Graphische Stanardisierung" und Probleme der Modellierung und Simulation
Heft 7 (1989)	Computergraphik und Anwendung DIGRA'88
Heft 8 (1989)	Computergraphik und Anwendung DIGRA'88
Heft 9 (1989)	Computergraphik und Anwendung DIGRA'88

Bezugsmöglichkeiten

- Bestellungen aus der DDR über die Wilhelm-Pieck-Universität Rostock, Abt. Wissenschaftspublizistik, Vogelsang 13/14, Rostock, DDR-2500.
- Bestellungen aus dem Ausland über die Firma Buchexport, Volkseigener Außenhandelsbetrieb der DDR, Leninstr. 16, Leipzig, DDR-7010.

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

ROSTOCKER MATHEMATISCHES KOLLOQUIUM

Heft 40

Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
1990

Herausgeber: Der Rektor der Wilhelm-Pieck-Universität Rostock

Wissenschaftliche Leitung: Prof. Dr. Gustav Burosch (Sektionsdirektor)
Prof. Dr. Gerhard Maeß

Redaktionelle Bearbeitung: Dr. Werner Plischke

Herstellung der Druckvorlage: Dipl.-Lehrer Andreas Straßburg

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

Redaktionsschluß: 30. 01. 1990

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

Außerdem bestehen Bezugsmöglichkeiten für

- Bestellungen aus der DDR über die Wilhelm-Pieck-Universität Rostock, Abteilung Wissenschaftspublizistik, Vogelsang 13/14, Rostock, DDR-2500.
- Bestellungen aus dem Ausland über die Firma Buchexport, Volkseigener Außenhandelsbetrieb der DDR, Leninstraße 16, Leipzig, DDR-7010.

Zitat-Kurztitel: Rostock. Math. Kolloq. (1990) 40

Wilhelm-Pieck-Universität Rostock,
Abteilung Wissenschaftspublizistik,
Vogelsang 13/14, Telefon 369 577, Rostock, DDR-2500

Druck: Ostsee-Druck Rostock BT Ribnitz II-15-14·0,50
1000

https://doi.org/10.18453/rosdok_id00003865

Inhalt Seite

THIELCKE, Beate : Zur Untergruppeneigenschaft von Klassenvereinigungen in endlichen Gruppen 4

WITHALM, Claudio : Supplierende pseudoholomorphe Funktionen 19

BEYER, Klaus : Zur Approxmation mittels Laméscher Multipolpotentiale 29

SCHEMPP, Walter : Elementary holograms, artificial neural networks, and theta - null values 35

SCHOTT, Dieter; KAPITANOWA, Maria; KOSTADINOW, Georgi; KOSTADINOW, Stepan : Über einen abstrakten Volterraoperator 51

MISHRA, S.N.; SINGH, S.L.: Some results on coincidences and fixed points 58

FISCHER, Hartmut : Stichprobenumfangsbestimmung für Konfidenzintervalle höherer Ordnung 71

DUCHRAU, Petra; FRISCHMUTH, Kurt : A numerical search algorithm for experimental design 83

Autorreferat

STADE, Carsten : Gaußsche Markovprozesse höherer Ordnung als Zeitreihenmodelle und ihre numerische Behandlung - mit Anwendungen auf Daten aus einem Wachstumsprozeß mit Tieren 85

Rostock. Math. Kolloq. 40, 4 - 18 (1990) 20D99

Beate Thielcke

Zur Untergruppeneigenschaft von Klassenvereinigungen in endlichen Gruppen

Herrn Prof. Dr. L. Berg zum 60. Geburtstag gewidmet

1. Einleitung

In der vorliegenden Arbeit werden Vereinigungen K von Konjugiertheitsklassen einer endlichen Gruppe G betrachtet, welche die {1}-Klasse enthalten und deren Mächtigkeiten die Ordnung von G teilen. Offensichtlich bilden derartige Komplexe K, kurz CNS-Komplexe genannt, in G im allgemeinen keine Untergruppe, verhalten sich jedoch invariant gegenüber inneren Automorphismen von G. Diese Tatsache gibt Anlaß zur folgenden

Definition: Wir nennen eine endliche Gruppe G eine Converse Normal Subgroup-Gruppe (kurz CNS-Gruppe), falls jeder in G enthaltene CNS-Komplex eine Untergruppe bildet.

Natürlich sind CNS-Komplexe in CNS-Gruppen stets sogar Normalteiler. Es erweist sich, daß sowohl hinreichend komplizierte Beispiele als auch Gegenbeispiele für CNS-Gruppen existieren, deren Untersuchung teils theoretisch, teils unter Einsatz von Computern erfolgt.
Bezeichnet Z_n die zylische Gruppe der Ordnung n, so ist leicht einzusehen, daß Z_1 und Z_p für jede Primzahl p sowie die Vierergruppe $V \cong Z_2 \times Z_2$ zu den CNS-Gruppen gehören, Z_n für eine zusammengesetzte Zahl n hingegen nicht.
Für das Zentrum Z(G) und für die Faktorgruppe G/G' nach der Kommutatorgruppe G' von CNS-Gruppen ergeben sich spezielle Gruppentypen, welche zur Auffindung der abelschen und der p-Gruppen unter den CNS-Gruppen führen.

2. Verwendete Zeichen

S_n symmetrische Gruppe des Grades n
A_n alternierende Gruppe des Grades n

Z_n	zyklische Gruppe der Ordnung n
V	Kleinsche Vierergruppe
Q	Quaternionengruppe
D_8	Diedergruppe der Ordnung acht
$\bar{x} = \{g^{-1}xg : g \in G\}$	bezeichnet die Konjugiertheitsklasse des Elements x in G
$o(\bar{p}(\bmod q))$	Ordnung der primen Restklasse $\bar{p}$ mod q

Unter der G-Klassengleichung

$$n = \sum_{i=1}^{t} n_i d_i \qquad (1)$$

eines unter G invarianten Komplexes N der Mächtigkeit n verstehen wir die Summe der Mächtigkeiten der in N gelegenen Konjugiertheitsklassen von G. Hierbei gibt n_i die Anzahl der d_i-elementigen Konjugiertheitsklassen von G in N an.
Weitere im Text verwendete Zeichen werden z.B. in [2] erklärt.

3. Uber das Zentrum und die Faktorgruppe nach der Kommutatorgruppe von CNS-Gruppen

Satz 1: Es sei G eine CNS-Gruppe. Dann ist Z(G) isomorph zu einer der Gruppen Z_1, Z_p mit einer Primzahl p oder V. Die Faktorgruppe G/G' besitzt die CNS-Eigenschaft.

Beweis: Stets ist Z(G) die Vereinigung aller einelementigen Konjugiertheitsklassen von G. Mit Ausnahme der in der Behauptung aufgelisteten Fälle für Z(G) läßt sich immer ein CNS-Komplex von G in Z(G) konstruieren, welcher in G keine Untergruppe bildet. Dieses ist leicht nachzuweisen und wird daher nicht weiter ausgeführt.
Um die CNS-Eigenschaft von G/G' zu zeigen, betrachten wir die Nebenklassenzerlegung

$$G = \bigcup_{x_i \in S} x_i G'$$

von G nach G'. Bekanntlich bestehen diese Nebenklassen aus vollständigen Konjugiertheitsklassen von G. Es sei

$$T := \{ xG' : x \in S_1 \subseteq S \}$$

ein in G/G' gelegener invarianter Komplex, welcher G' enthalte, mit einer die Ordnung von G/G' teilenden Mächtigkeit $|S_1|$. Dann ist

$$U := \bigcup_{x \in S_1} xG'$$

in G ein invarianter Komplex, der die $\{1\}$-Klasse enthält und dessen Mächtigkeit die Ordnung von G teilt. Weil G die CNS-Eigenschaft besitzt, ist U eine Untergruppe in G. Somit bildet auch U/G' eine Untergruppe in G/G'. Hieraus folgt die behauptete CNS-Eigenschaft von G/G'.
Andererseits ist G/G' eine abelsche Gruppe, so daß sie notwendigerweise isomorph zu Z_1, Z_q mit einer Primzahl q oder V ist.

Folgerung: Eine nichtabelsche p-Gruppe mit $p \neq 2$ besitzt nicht die CNS-Eigenschaft.

Beweis: Andernfalls folgt aus Satz 1, daß G/G' zyklisch ist. Demnach muß G selbst zyklisch sein im Widerspruch zur Voraussetzung.

4. Die CNS-Gruppen der Ordnung 2^n

Zunächst werden Hamiltonsche Gruppen auf CNS-Eigenschaft untersucht.

Satz 2: Eine Hamiltonsche Gruppe ist genau dann CNS-Gruppe, wenn sie Quaternionengruppe ist.

Beweis: Es sei G eine Hamiltonsche Gruppe mit CNS-Eigenschaft. Nach [1], Satz 10.2.5. gilt

$$G \cong Q \times A \times B,$$

wobei Q die Quaternionengruppe bezeichnet, A eine abelsche Gruppe, in der alle Elemente ungerade Ordnung besitzen, und B eine abelsche Gruppe vom Exponenten ≤ 2 ist. Bekanntlich gilt

$$G' = Q' \times A' \times B' = Q'.$$

Demnach ist $Q/Q' = Q/G'$ eine zur Vierergruppe isomorphe Untergruppe von G/G'. Nach Satz 1 muß $G/G' \cong V$ erfüllt sein,

woraus sich $G/G' = Q/G'$, also $G = Q$ ergibt.
Wir werden im folgenden die 2-Gruppen auf CNS-Eigenschaft prüfen. Aus Satz 1 folgt, daß $|G/G'| = 4$ ist. Der Satz von O. Taussky (vgl. [2], Kap. 3, Satz 11.9.) liefert, daß G isomorph ist zur Quasidiedergruppe, zur Diedergruppe oder zur verallgemeinerten Quaternionengruppe. Unter den genannten Gruppen besitzen nur Q und D_8 die CNS-Eigenschaft, wovon man sich durch Nachrechnen leicht überzeugt. Zusammengefaßt haben wir damit folgendes bewiesen:

Satz 3: Eine nichtabelsche 2-Gruppe G besitzt die CNS-Eigenschaft genau dann, wenn G isomorph zu einer der Gruppen Q und D_8 ist.

Die Klärung der CNS-Eigenschaft von 2-Gruppen läßt sich nun leicht auf nilpotente Gruppen erweitern.

Folgerung: Es sei G eine nichtabelsche nilpotente CNS-Gruppe. Dann ist G isomorph zu Q oder D_8.

Beweis: Nach Voraussetzung muß G/G' isomorph zu V sein. Lemma 1.v. aus [3] liefert, daß G eine 2-Gruppe ist. Folglich gilt $G \cong Q$ oder D_8.

5. Über metabelsche CNS-Gruppen

In diesem Abschnitt werden die metabelschen Gruppen (Gruppen mit abelscher Kommutatorgruppe) auf CNS-Eigenschaft untersucht. Unter ihnen betrachten wir gesondert die einstufig nichtkommutativen Gruppen, kurz Rédei-Gruppen genannt; das sind diejenigen nichtabelschen Gruppen, welche nur abelsche Untergruppen haben. Unseren bisherigen Erkenntnissen zufolge, interessieren nur die Gruppen, welche Nichtprimzahlpotenzordnung haben. Wie man in [4], Kap. xii, Satz 445 nachliest, besitzen sie die Presentation: Man gebe zwei verschiedene Primzahlen p und q sowie eine natürliche Zahl n vor, bezeichne mit m die Ordnung der primen Restklassen $\tilde{p}$ mod q, also

$$m := o(\,p(\bmod q)), \qquad (1)$$

nehme den Körper

$$K = GF(p^m) \qquad (2)$$

und aus der multiplikativen Gruppe K^* von K ein beliebiges, festes Element ω mit

$$o(\omega) = q \tag{3}$$

und betrachte eine zyklische Gruppe $\langle a\rangle$ der Ordnung q^n. Dann bilden alle Paare

$$(\nu, a^i) \quad \text{mit} \quad \nu \in K, \; i = 0,1,\ldots,q^n-1, \tag{4}$$

mit der Multiplikationsregel

$$(\nu, a^i)(\mu, a^j) = (\nu+\mu\omega^i, a^{i+j}), \; \nu,\mu \in K, \; i,j = 0,1,\ldots,q^n-1, \tag{5}$$

eine Gruppe der Ordnung $p^m q^n$. Diese Gruppe ist bis auf Isomorphie die einzige Rédei-Gruppe der Ordnung $p^n q^m$. Üblicherweise identifiziert man die Paare (ν, a^0) mit ν und $(0, a^i)$ mit a^i. Aus (5) folgt, daß G' aus den sämtlichen ν, $\nu \in K$, besteht. Somit ist $|G'| = p^m$. Weiter sieht man, daß G/G' zyklisch ist, woraus wir Satz 1 zufolge $n = 1$ erhalten. Mithin ist $Z(G) = \langle 1\rangle$, und G' besitzt die G-Klassengleichung

$$p^m = 1 + \left(\frac{p^m - 1}{q}\right) q, \tag{6}$$

weil der Index von G' in G gleich q ist.
Wir betrachten die außerhalb von G' gelegenen Konjugiertheitsklassen von G. Es sei

$$D = N_G(a) \cap G' \tag{7}$$

Offenbar sind $\langle a\rangle$ und D Normalteiler in $N_G(a)$. Aus Gründen der Elementordnungen ist $\langle a\rangle \cap D = \langle 1\rangle$, so daß $\langle a\rangle$ und D sogar elementweise vertauschbar sind. Andererseits gilt für alle $g \in G'\setminus\{1\}$ stets $N_G(g) = G'$, woraus $D = \langle 1\rangle$ folgt. Somit besitzt jede von 1 verschiedene a-Potenz genau p^m Konjugierte in G, und folglich lautet die Klassengleichung von G

$$p^m = 1 + \left(\frac{p^m - 1}{q}\right)q + (q - 1)p^m. \tag{8}$$

Aus (8) ergibt sich, daß der einzige eigentliche CNS-Komplex von G die Kommutatorgruppe G' ist und G folglich die CNS-

Eigenschaft besitzt.
Wir fassen unsere Erkenntnisse in dem folgenden Satz zusammen:

Satz 4: Unter den einstufig nichtkommutativen Gruppen besitzen genau die zentrumslosen Gruppen die CNS-Eigenschaft.

Im weiteren sei G eine beliebige metabelsche CNS-Gruppe. Wir beweisen zunächst das folgende

Lemma 1: Es sei G eine metabelsche Gruppe mit CNS-Eigenschaft. Dann ist G' eine p-Gruppe.

Beweis: Gemäß Satz 1 haben wir hinsichtlich G/G' zwei Fälle zu unterscheiden:

1. Fall: $G/G' \cong Z_q$, q Primzahl.

Dann gilt $Z(G) < G'$, und G' ist die Vereinigung von ein- oder q-elementigen Konjugiertheitsklassen von G. Wir schreiben G' als das direkte Produkt seiner Sylowgruppen:

$$G' = P_1 \times P_2 \times \ldots \times P_r, \ |P_i| = p_i^{t_i}, \ i = 1,2, \ldots, r. \tag{9}$$

Offensichtlich sind die Sylowgruppen von G' Normalteiler in G. Falls $Z(G) \cong Z_1$ ist, so bildet für $r \geq 2$ in (9)

$$K := (P_2 \backslash \bar{x}) \cup \bar{y}, \ x \in P_2 \backslash \{1\}, \ y \in P_1 \backslash \{1\}, \tag{10}$$

einen CNS-Komplex in G, welcher keine Gruppe ist.
Wenn $Z(G) \cong Z_p$ mit einer Primzahl p ist, so sei o.B.d.A. $Z(G) < P_1$ und $|P_1| = p^t$. Für $r \geq 2$ und $t > 1$ läßt sich K wie in (10) konstruieren, wobei $y \in P_1 \backslash Z(G)$ gewählt wird. Ist $t = 1$, so bildet

$$p_2^{t_2} = 1 + k \cdot q, \ k \geq 1, \tag{11}$$

die G-Klassengleichung von P_2 und

$$p \cdot p_2^{t_2} = p + (k \cdot p) \cdot q \tag{12}$$

die G-Klassengleichung von $P_1 \times P_2$. Aus (11) und (12) folgt, daß sich K wie in (10) konstruieren läßt, wobei y aus $(P_1 \times P_2) \backslash (P_1 \cup P_2)$ gewählt wird.
Schließlich kann noch $Z(G) \cong V$ auftreten. Dann schließen wir

wie im Fall $Z(G) \cong Z_p$ auf $r = 1$ in (9).

2. Fall: $G/G' \cong V$.

Wir schreiben G' als das direkte Produkt seiner Sylowgruppen P_i, $i = 1,2,\dots,r$, wie in (9) und nehmen $r > 1$ an. In P_1 und P_2 wählen wir jeweils einen minimalen Normalteiler N_1 und N_2 von G. Dann gilt $\Phi(N_i) = \langle 1\rangle$, d.h., die N_i sind elementarabelsch für i = 1,2. Wir betrachten N_1 und N_2 als Darstellungsmoduln für V über $GF(p_1)$ und $GF(p_2)$. Als minimale Normalteiler von G liefern N_1 und N_2 jeweils irreduzible Darstellungen von V, welche wegen $\{1,-1\}\subseteq GF(p_i)$, i = 1,2, sogar absolut irreduzibel sind. Folglich besitzen sie den Grad eins, woraus sich $N_i \cong Z_{p_i}$ (i = 1,2) ergibt. Wir wenden Satz 4.5. aus [2], Kap. 1, an, wonach $N_G(N_i)/C_G(N_i) = G/C_G(N_i)$ isomorph ist zu einer Untergruppe U_i der Automorphismengruppe Aut N_i von N_i für i = 1,2. Weil $|N_i| = p_i$ gilt, sind Aut N_i und damit U_i (i = 1,2) zyklisch.

Im Hinblick auf $C_G(N_i) \geq G'$ gilt $G/C_G(N_i) \cong Z_2$ oder Z_1. Folglich bestehen N_1 und N_2 aus ein- oder 2-elementigen Konjugiertheitsklassen von G. Wir schreiben die G-Klassengleichungen von N_1 und N_2 in der Form

$$p_1 = z\cdot 1 + h\cdot 2,\ z\in\{1,p_1\}, \tag{13}$$

$$p_2 = \quad 1 + k\cdot 2. \tag{14}$$

Wenn $h \geq 1$ ist, so gilt $z = 1$, und wir können o.B.d.A. $p_2 > 3$ annehmen. Wir bilden mit $x\in N_2\backslash\{1\}$, $y\in N_1\backslash\{1\}$ den CNS-Komplex

$$K := (N_2\backslash\bar{x}) \cup \bar{y}. \tag{15}$$

Weil N_2 von Primzahlordung ist, besitzt K nicht die Gruppeneigenschaft.
Andernfalls ist $h = 0$, und folglich gilt $z = p_1$. Die G-Klassengleichung von $N_1 \times N_2$ hat die Form

$$p_1\cdot p_2 = p_1\cdot 1 + (p_1\cdot k)\cdot 2. \tag{16}$$

Wir konstruieren K wie in (15), wobei $\bar{x} = \{x, x^g\} \subset N_2$ mit $g\in G$ sei und $\bar{y} := \overline{xz} = \{xz, x^g z\}$ mit $z\in N_1\backslash\{1\}$ gewählt wird.

Da $\bar{y}$ in $(N_1 \times N_2) \setminus (N_1 \cup N_2)$ liegt, liefert K für $p_2 \neq 3$ einen Widerspruch zur CNS-Eigenschaft von G. Wenn $p_2 = 3$ ist, so gehen (13) und (14) über in

$$p_1 = p_1 \cdot 1 \tag{13'}$$

$$3 = 1 + 2. \tag{14'}$$

Wir erhalten mit

$$K := \{1\} \cup \{z\} \cup \bar{x},\ z \in N_1 \setminus \{1\},\ x \in N_2 \setminus \{1\}, \tag{17}$$

einen in $N_1 \times N_2$ gelegenen CNS-Komplex von G. Dieser bildet wegen $|K| \nmid |N_1 \times N_2|$ keine Gruppe. Somit ist $r = 1$ in (9) und das Lemma bewiesen.

Satz 5: Es sei G eine metabelsche CNS-Gruppe, die keine Primzahlpotenzordnung habe. Dann ist G entweder eine zentrumslose Rédei-Gruppe oder von der Form

$$G \cong Z_q \ltimes (M \times N)$$

mit einer Primzahl q, wobei $Z_q \ltimes M$ und $Z_q \ltimes N$ zentrumslose Rédei-Gruppen sind und zudem

$$|M| = |N| = 1 + q \tag{18}$$

gilt.

Bemerkung: Aus (18) folgt, daß entweder $q = 2$ und $|M| = |N| = 3$ gilt oder $|M| = |N| = 2^s$ und $q = 2^s-1$ eine Mersennesche Primzahl ist.

Beweis: Nach Lemma 1 können wir $|G'| = p^t$, $t \geqslant 1$, mit einer Primzahl p setzen.
Zunächst sei G/G' isomorph zu Z_q mit einer von p verschiedenen Primzahl q. Dann gilt $Z(G) \leqslant G'$.
Wenn $Z(G) \cong Z_p$ ist, so besitzt G' die G-Klassengleichung

$$p^t = p + r \cdot q, \tag{19}$$

wobei $r > 0$ ist. Andernfalls liefert $Z(G) = G'$ die Nilpotenz von G, was nicht gelten sollte. Aus (19) folgt

$$p^{t-1} \equiv 1 \pmod q.$$

Wir betrachten einen CNS-Komplex K mit der G-Klassengleichung

$$p^{t-1} = 1 + s\cdot q,\ 0 < s < r. \tag{20}$$

Dann ist auch

$$L := (K\setminus\bar{x}) \cup \bar{y},\ x\in K\setminus\{1\},\ y\in G'\setminus(K\cup Z(G)) \tag{21}$$

ein CNS-Komplex in G, dessen Durchschnitt D mit K genau $(p^{t-1}-q)$ Elemente enthält. Weil $D \leq G'$ ist, folgt $|D| \mid |G'|$ oder m.a.W.

$$p^{t-1} = 1 + q. \tag{22}$$

Andererseits ist $K\cdot L$ ein in G' gelegener Normalteiler von G mit der Ordnung $p^{t-1}\cdot p^{t-1}$, welcher $|G'| = p^t$ teilt. Das ist nur für $t = 2$ möglich. Unter Beachtung von (22) erhalten wir $p = 3$ und $q = 2$. Anstelle von (19) steht

$$3^2 = 3\cdot 1 + 3\cdot 2. \tag{19'}$$

Wir bilden den CNS-Komplex

$$M := \{1\} \cup \{z\} \cup \bar{x} \cup \bar{y},\quad z\in Z(G),\ x,y\in G'\setminus Z(G).$$

Weil $(6=)|M|\nmid|G'|$, erhalten wir im Widerspruch zur CNS-Eigenschaft von G, daß M keine Gruppe ist.
Falls $Z(G) \cong V$ ist, geht (19) über in

$$2^t = 4 + r\cdot q,\ r > 0. \tag{19''}$$

Es sei K ein CNS-Komplex mit der G-Klassengleichung

$$2^{t-1} = 2 + s\cdot q,\ 0 < s < r,$$

und

$$L := (K\setminus\{z\})\cup\{z'\},\ z\in(K\cap Z(G))\setminus\{1\},\ z'\in Z(G)\setminus(K\cap Z(G)).$$

Wir erhalten $|K \cap L| = (2^{t-1}-1)|2^t$. Hieraus folgt $2^{t-1} = 2$,

also $t = 2$ im Widerspruch zu $r > 0$ in (19"). Somit gilt $Z(G) = \langle 1 \rangle$.
Angenommen, in G' existiert ein Normalteiler N von G der Ordnung p^s mit $s \geq 1$. Wir bilden

$$M := (N\backslash\bar{x}) \cup \bar{y},\ x \in N\backslash\{1\},\ y \in G'\backslash N, \tag{23}$$

und erhalten $|M \cap N| = (p^s - q) | p^t$. Es folgt

$$p^s = 1 + q. \tag{24}$$

Andererseits ist $M \cdot N$ ein Normalteiler in G. Wenn $t > 2 \cdot s$ ist, so liefert der CNS-Komplex

$$L := (M \cdot N)\backslash\bar{x} \cup \bar{w},\ x \in MN\backslash\{1\},\ w \in G'\backslash MN, \tag{25}$$

wegen $|L \cap MN| = (p^{2s} - q) \nmid p^t$ einen Widerspruch zur CNS-Eigenschaft von G. Folglich muß $t = 2 \cdot s$ gelten. Die außerhalb von G' gelegenen Elemente besitzen alle p^t-elementige Konjugiertheitsklassen in G, so daß

$$p^t \cdot q = 1 + (q + 2) \cdot q + (q - 1) \cdot p^t \tag{26}$$

die Klassengleichung von G ist. Weiterhin stellt $G' = M \times N$ als direktes Produkt zweier minimaler Normalteiler von G eine elementare abelsche Gruppe dar. Aus (26) folgt, daß neben den in G' gelegenen minimalen Normalteilern von G und G' selbst keine nichttrivialen CNS-Komplexe in G bildbar sind und G somit die CNS-Eigenschaft besitzt.
Wenn in G' keine minimalen Normalteiler von G auftreten, so gehört G zu den zentrumslosen Rédei-Gruppen, welche nach Satz 4 CNS-Gruppen sind.

Wir wenden uns nun $G/G' \cong V$ zu und setzen gemäß Lemma 1 $|G'| = p^t$. Aus dem Satz von Schur/Zassenhaus (vgl. [1], Kap. 6, Satz 6.2.2.) folgt die Existenz einer Untergruppe V in G, welche isomorph zu G/G' ist. Wir schreiben

$$G = V \ltimes G',\ V = \{1, a, b, c\}, \tag{27}$$

und erhalten aus der Kommutativität von G' und V, daß $G' = [V, G']$ gilt. Nach Voraussetzung ist G nichtabelsch, so daß mindestens ein $v \in V$ existiert mit $v \notin C_G(G')$. Wir setzen o.B.d.A. $v = a$ und bilden $a^{-1}g^{-1}ag = h \in G'$ mit $g \in G'$. Hieraus folgt $a^{-1}ga = gh^{-1}$. Wegen $a^2 = 1$ ergibt sich $g = gh^{-1}(h^{-1})^a$. Somit gilt $h^a = h^{-1}$. Wir definieren

$$M_a := \{ h: h \in G': h^a = h^{-1} \}.$$

M_a bildet wegen der Kommutativität von G' einen in G' gelegenen Normalteiler von G. Analog lassen sich die Normalteiler M_b und M_c konstruieren, welche gegebenenfalls auch trivial sein können. Wir wählen N minimal mit $N \leq M_a$, $N \trianglelefteq G$. Dann gilt $\Phi(N) = \langle 1 \rangle$, also ist N elementarabelsch. Wir betrachten N als Darstellungsmodul für V über $GF(p)$. Als minimaler Normalteiler von G liefert N eine irreduzible Darstellung von V. Diese ist wegen $\{1,-1\} \subset GF(p)$ absolut irreduzibel und besitzt folglich den Grad eins, woraus sich $|N| = p$ ergibt. Wir wenden Satz 4.5. aus [2], Kap. 1, an, wonach $G/C_G(N)$ isomorph ist zu einer Untergruppe U von Aut N. Weil $|N| = p$ gilt, sind Aut N und damit U zyklisch. Somit gilt $G/C_G(N) \cong Z_2$. Mithin besteht N aus ein- und 2-elementigen Konjugiertheitsklassen von G. Weil $|N| = p$ eine ungerade Primzahl ist und nach Konstruktion für $n \in N$ stets $n^a = n^{-1}$ gilt, besitzt N genau eine einelementige Konjugiertheitsklasse ($\bar{1} = \{1\}$) von G. Wir unterscheiden hinsichtlich $Z(G)$ folgende beiden Fälle:

1. Fall: $|Z(G)|>1$.

Es sei $n \in N$ mit $\bar{n} = \{n, n^{-1}\}$. Damit gilt $\overline{nz} = \{ nz, n^{-1}z\}$ für ein beliebiges $z \in Z(G) \setminus \{1\}$ und $\overline{nz} \notin N$. Wir bilden

$$K := (N \setminus \bar{n}) \cup \overline{nz}$$

und erhalten einen Widerspruch zur CNS-Eigenschaft von G, da $nz(n^{-1}z)^{-1} = n^2 \notin K$ ist.

2. Fall: $Z(G) = \langle 1 \rangle$.

Es seien N_a, N_b und N_c zu M_a, M_b und M_c gehörige minimale Normalteiler von G. Weil $Z(G) = \langle 1 \rangle$ vorausgesetzt wird, kann keine der Gruppen M_a, M_b und M_c trivial sein. Es gelte $N_a = N_b = N_c$. Dann existiert $h \in G' \setminus \{1\}$ mit $h^a = h^b = h^c = h^{-1}$. Andererseits ist nach (27) $h^c = h^{ab} = h$, womit wir einen Widerspruch erhalten haben. Folglich können wir o.B.d.A. $N_a \neq N_b$ setzen.

Wie wir oben erkannten, liegen in N_a und N_b mit Ausnahme von $\bar{1} = \{1\}$ nur zweielementige Konjugiertheitsklassen von G. Für $p \neq 3$ liefert wegen $N_a \cong Z_p$ der CNS-Komplex

$$K := (N_a \backslash \bar{n}_a) \cup \bar{n}_b, \quad n_a \in N_a \backslash \{1\}, \; n_b \in N_b \backslash \{1\},$$

einen Widerspruch zur CNS-Eigenschaft von G.
Wenn $p = 3$ ist, so bildet N_aN_b einen Normalteiler der Ordnung neun in G. Für $|G'| > 9$ liefert

$$L := ((N_aN_b) \backslash \bar{n}) \cup \bar{g}, \; g \in G' \backslash N_aN_b, \; n \in N_a \backslash \{1\}, \; |\bar{g}| = 2,$$

bzw.

$$M := ((N_aN_b) \backslash (\bar{n} \cup \bar{m})) \cup \bar{g}, \; g \in G' \backslash N_aN_b, \; |\bar{g}| = 4, \; n \in N_a \backslash \{1\}, \; m \in N_b \backslash \{1\},$$

wegen $| L \cap N_aN_b | \nmid 3^m$, $| M \cap N_aN_b | \nmid 3^m$ einen Widerspruch zur CNS-Eigenschaft von G. Für $|G'| = 9$ betrachten wir $v \in V$ mit $| \bar{v} | = 3$. Der CNS-Komplex

$$K := \{1\} \cup \bar{v}$$

bildet keine Gruppe in G. Andernfalls wäre $G = K \times G'$ das direkte Produkt seiner Sylowgruppen K und G' und folglich eine 2-Gruppe. Das sollte nicht gelten, womit der Beweis komplett ist.

6. Die symmetrischen und alternierenden CNS-Gruppen

Satz 5: Die symmetrische Gruppe S_n ist genau dann CNS-Gruppe, wenn $n \leq 4$ ist.

Beweis: Die symmetrischen Gruppen besitzen bis zum Grad vier die CNS-Eigenschaft, wovon man sich durch Nachrechnen leicht überzeugt.
Im folgenden sei $n > 4$ beliebig, aber fest gewählt. Die alternierende Gruppe A_n ist nach [1], Satz 7.4.2., einziger eigentlicher Normalteiler der S_n. Es genügt also, einen in S_n gelegenen, nichttrivialen und von A_n verschiedenen CNS-Komplex K zu konstruieren.
Wenn n gerade ist, dann liegt die Menge

$$M = \{ \overline{(12)(34\ldots n)} \}$$

in der A_n und besteht aus $\binom{n}{n-2}\cdot(n-3)!$ Elementen. Andererseits liegt

$$M' = \{\overline{(1)(2)(34\ldots n)}\}$$

nicht in A_n. M und M' sind gleichmächtig. Wir bilden

$$K := (A_n \backslash M) \cup M'.$$

Der so konstruierte CNS-Komplex K ist von A_n verschieden und liefert folglich für gerades n die Behauptung.
Falls n ungerade ist, so ist $M' \subseteq A_n$ und $M \nsubseteq A_n$. Wir bilden

$$K := (A_n \backslash M') \cup M$$

und erhalten wie oben die Behauptung für ungerades n. Hiermit ist der Beweis vollständig.

Wir bezeichnen von nun an der Kürze halber mit

$$\overline{(i_1)(i_2)\ldots(i_r)},\ i_1, i_2, \ldots, i_r \in \{1, 2, \ldots, n\}, \tag{1}$$

die Menge aller Permutationen π aus S_n, welche in der Zyklenschreibweise aus einem Zyklus der Länge i_1, einem Zyklus der Länge i_2, ..., und einem Zyklus der Länge i_r bestehen. Dabei werden Zyklen der Länge eins bei allen $\pi \neq id$ weggelassen, und in $\pi = id$ steht $\overline{(1)}$ anstelle von (1). Bekanntlich bilden alle derartigen Permutationen (welche Permutationen gleichen Typs ausmachen) eine vollständige Konjugiertheitsklasse in der S_n.
Die alternierenden Gruppen wurden mittels Einsatz eines Computers untersucht. Sie gehören bis zum Grade acht zu den CNS-Gruppen. Die A_9 bis A_{13} besitzen nicht die CNS-Eigenschaft, weil z.B. folgende CNS-Komplexe gebildet werden können:

in A_9

$$K_9 = \{\overline{(1)}, \overline{(3)(3)(3)}, \overline{(5)}, \overline{(7)}, \overline{(4)(4)}, \overline{(2)(2)(2)(2)}, \overline{(4)(2)},$$

$\overline{(5)(2)(2)}, \overline{(2)(2)}\}$ mit $| K_9 | = | A_9 | / 3$;

in A_{10}

$K_{10} = \{\overline{(1)}, \overline{(3)(3)(3)}, \overline{(5)}, \overline{(7)}, \overline{(3)(3)}, \overline{(2)(2)(2)(2)}, \overline{(4)(2)},$

$\overline{(5)(2)(2)}, \overline{(2)(2)}, \overline{(3)}, \overline{(5)(3)}, \overline{(5)(5)}, \overline{(7)(3)}\}$

mit $| K_{10} | = | A_{10} | / 3$;

in A_{11}

$K_{11} = \{\overline{(1)}, \overline{(3)(3)(3)}, \overline{(5)}, \overline{(7)}, \overline{(4)(4)}, \overline{(2)(2)(2)(2)}, \overline{(4)(2)},$

$\overline{(5)(2)(2)}, \overline{(2)(2)}, \overline{(8)(2)}, \overline{(5)(4)(2)}, \overline{(5)(5)}, \overline{(7)(2)(2)},$

$\overline{(11)}, \overline{(4)(2)(2)(2)}\}$

mit $| K_{11} | = | A_{11} | / 2$;

in A_{12}

$K_{12} = \{\overline{(1)}, \overline{(9)(3)}, \overline{(5)}, \overline{(6)(6)}, \overline{(7)(5)}, \overline{(2)(2)(2)(2)}, \overline{(7)(2)(2)},$

$\overline{(3)}, \overline{(5)(5)}, \overline{(5)(3)}, \overline{(7)(3)}, \overline{(11)}, \overline{(5)(4)(2)}\}$

mit $| K_{12} | = | A_{12} | / 2$;

in A_{13}

$K_{13} = \{\overline{(1)}, \overline{(3)(3)(3)}, \overline{(9)}, \overline{(13)}, \overline{(4)(4)}, \overline{(2)(2)(2)(2)}, \overline{(10)(2)},$

$\overline{(5)(3)(3)}, \overline{(2)(2)}, \overline{(3)(3)}, \overline{(5)(5)(3)}, \overline{(9)(3)}, \overline{(4)(4)(3)},$

$\overline{(11)}, \overline{(3)(3)(3)(3)}, \overline{(3)}, \overline{(2)(2)(2)(2)(2)(2)}, \overline{(6)(5)(2)},$

$\overline{(3)(3)(3)(2)(2)}\}$ mit $| K_{13} | = | A_{13} | / 2$.

Weil die betrachteten alternierenden Gruppen einfach sind, bilden die aufgelisteten CNS-Komplexe in der jeweiligen A_n keine Untergruppen.
Die sämtlichen für A_9 bis A_{13} gefundenen CNS-Komplexe - von denen oben fünf angegeben wurden - lassen keinen Zusammenhang erkennen, weder hinsichtlich der generell auftretenden Mächtigkeiten noch hinsichtlich der auftretenden Konjugiertheitsklassen, so daß eine Vermutung über die CNS-Eigenschaft der A_n für beliebiges n nicht getroffen werden kann.

7. Eine allgemeine lineare Gruppe mit CNS-Eigenschaft

In diesem Abschnitt wird GL(3,2) auf CNS-Eigenschaft untersucht. Da es sich bei den Elementen von GL(3,2) um 3×3 Matrizen über GF(2) handelt, bietet sich die Lösung mittels eines Rechnerprogramms an. Im Algorithmus wurden die Mächtigkeiten sämtlicher Konjugiertheitsklassen berechnet. Sie liefern die folgende Klassengleichung

$$| GL(3,2) | = 168 = 1 + 21 + 42 + 56 + 24 + 24.$$

Einzig bildbarer CNS-Komplex ist $K = \langle 1 \rangle$. Wir erhalten hiermit den folgenden

Satz 6: GL(3,2) ist eine einfache Gruppe mit CNS-Eigenschaft.

Abschließend möchte ich meinem Betreuer, Herrn Prof. Dr.G. Pazderski, und Herrn Dr. W. Bannuscher für die freundliche Beratung danken.

Literatur

1. Kochendörffer, R.: Lehrbuch der Gruppentheorie unter besonderer Berücksichtigung der endlichen Gruppen. Leipzig 1966

2. Huppert, B.: Endliche Gruppen. Berlin 1967

3. Schenkman, E.: Group Theory. Princeton New Jersey 1968

4. Rédei, L.: Algebra, 1. Leipzig 1959

eingegangen: 17. 10. 1989

Verfasser :

Dipl.-Math. Beate Thielcke
Wilhelm-Pieck-Universität Rostock
Sektion Mahematik
Universitätsplatz 1
Rostock
DDR-2500

Rostock. Math. Kolloq. 40, 19 - 28 (1990)   30A93
35J45

Claudio Withalm

Supplierende pseudoholomorphe Funktionen

Herrn Prof. Dr. L. Berg anläßlich seines 60. Geburtstages herzlichst zugeeignet

Abstract

Representing solutions of the homogeneous Vekua differential equation as a pseudoholomorphic function in the sense of L. Bers on certain conditions, in the correlative vector space one can specify supplementary elements of which the real and the imaginary part of the corresponding pseudoholomorphic functions of the second kind prove to be conjugated - harmonic with respect to those of the original solution.

1. Vorbemerkungen

1.1. Wir betrachten die Lösungen der homogenen Vekua-Gleichung ([6, 3])

$$w_{z} - aw - b\bar{w} = 0 \tag{1}$$

als $\langle F,G\rangle$-pseudoholomorphe Funktionen $w = F\varphi + F\psi$ im Bers'schen Sinne ([1, 2]), worin also $E := \langle F,G\rangle$ das Erzeugendensystem mit $E \in H^1_{D_0} \times H^1_{D_0}$ 1), $\operatorname{Im}(\bar{F}G) > 0$ und φ, ψ reell in $D \subset D_{\bullet} \subset \mathbb{C}$ seien, und die korrelative E-Ableitung vermöge der komplex-linearen Approximation modulo E von w an $z_0 \in D$

$$\begin{aligned} &w(z) = F(z)\varphi(z_0) + G(z)\psi(z_0) + \lambda(z-z_0) + \varrho|z-z_0| \\ &\lim_{|z-z_0|\to 0} \varrho(z,z_0) = 0 \end{aligned} \tag{2}$$

erklärt ist. (2) ist äquivalent mit

$$\lambda =: w(z_0) =: \frac{d_E w}{dz}(z_0) = \lim_{z\to z_0} \frac{w(z) - F(z)\varphi(z_0) - G(z)\psi(z_0)}{z - z_0}. \tag{3}$$

3*

Existiert (3) überall in D eigentlich, so heißt w dort eine pseudoholomorphe Funktion modulo E erster Art (E-pseudoholomorphe Funktion erster Art) und $\varphi + i\psi$ die korrespondierende pseudoholomorphe Funktion modulo E zweiter Art (E-pseudoholomorphe Funktion zweiter Art); diese eineindeutige Zuordnung werde vermöge $p_E w = \varphi + i\psi$ $(p_E^{-1}(\varphi + i\psi) = F\varphi + G\psi = w)$ ausgedrückt. $P_D(E)$ bezeichne den additiven Vektorraum über $\mathbb{R}$ aller über D erklärten pseudoholomorphen Funktionen modulo E.

1.2. Es sei $U = U(z_o) \subset D$ eine Umgebung von z_o; ist dann $w \in P_u(E)$, so existieren φ_z, ψ_z, $\varphi_{\mathbf{z}}$, $\psi_{\mathbf{z}}$ an z_o, und dort gilt

$$F\varphi_{\mathbf{z}} + G\psi_{\mathbf{z}} = 0, \qquad F\varphi_z + G\psi_z = \dot{w}. \tag{4}$$

Sind umgekehrt $\varphi, \psi \in C_u^1$ [1] und gilt $F\varphi_{\mathbf{z}} + G\psi_{\mathbf{z}} = 0$ daselbst, so ist $w \in P_u(E)$.

1.3. F und G sind selbst E-pseudoholomorphe Funktionen, die charakteristischen Koeffizienten a_E, b_E, A_E, B_E von $E = \langle F,G\rangle$ bestimmen sich durch

$$F_{\mathbf{z}} = a_E F + b_E \bar{F}, \qquad F_z = A_E F + B_E \bar{F},$$
$$G_{\mathbf{z}} = a_E G + b_E \bar{G}, \qquad G_z = A_E G + B_E \bar{G}.$$

Sind $a, b \in H_{D_o}$, beschränkt, so gibt es stets ein Erzeugendensystem $E = \langle F,G\rangle$ in D_o, so daß $a_E = a$ und $b_E = b$ [2]. (4) und seine Umkehrung erfahren somit die äquivalente Form: Ist $w \in P_u(E)$, so existieren w_z und $w_{\mathbf{z}}$ an z_o, und es gilt dort (vgl. (1))

$$w_{\mathbf{z}} - aw - b\bar{w} = 0, \qquad w_z - Aw - B\bar{w} = \dot{w}. \tag{5}$$

Ist umgekehrt $w \in C_u^1$ und gilt dort $w_{\mathbf{z}} - aw - b\bar{w} = 0$, so ist $w \in P_u(E)$. Ist $\langle F,G\rangle \equiv \langle 1,i\rangle =: I$, so gehen alle Aussagen ersichtlich in die entsprechenden klassischen Sätze über; $P_D(I)$ bezeichnet demnach einfach die Algebra (über $\mathbb{K}$) der über D definierten holomorphen Funktionen.

1.4. Ist $w \in P_D(E)$, so ist $p_E w \in H_D^2$, $\dot{w} \in H_D^1$, und es gilt $(\dot{w})_{\mathbf{z}} = a\dot{w} - B\bar{\dot{w}}$. Insbesondere die letzte Aussage zeigt, daß $\dot{w}$ einer Vekua-Gleichung (1) genügt, woraus auf die Existenz eines Erzeugendensystems $E^1 := \langle F_1, G_1\rangle$ geschlossen werden kann, so daß $w \in P_D(E^1)$ ist. Man hat also Anlaß, E^1 als Nachfolger von E (E Vorgänger von E^1) zu erklären, wenn $a_{E^1} = a_E$ und $b_{E^1} = -B_E$ gilt. Iterativ erkennt man, daß sich $E =: E^o$ in eine Erzeugendenfolge $(E^\nu)_{\nu \in \mathbb{Z}}$ einbetten läßt und daß

$$w^{(n+1)} = \frac{d^n_E w}{dz}, \quad w^{(o)} = w, \quad w^{(1)} = \dot{w} = \frac{d_{E^o} w}{dz}, \qquad n = 0, 1, \ldots$$

gilt. Gilt für festes $\mu \geq 1$ stets $E^{\nu+\mu} = E^\nu$, so heißt μ die Periode der Erzeugendenfolge. (Minimalperiodenproblem siehe [1, 4]) Insbesondere ist $\mu = 1$ genau dann, wenn $\partial/\partial x(G/F) = 0$, was einfach aus $b_E + B_E = 0$ folgt.

1.5. Zur Untersuchung einiger wichtiger Klassen pseudoholomorpher Funktionen spielt die Abhängigkeit der Komponenten F und G eines Erzeugendensystems E eine wesentliche Rolle (siehe z.B. [1, 7]). Aber auch die Relationen zwischen zwei Erzeugendensystemen geben Anlaß für weitreichende theoretische Überlegungen (vergleiche z.B. [1, 2]). So heißen etwa E und $\tilde{E}$ zueinander ähnlich, wenn es eine Funktion H gibt, so daß $\tilde{E} = HE$ ist; dabei ist $w \in P_D(E)$ genau dann, wenn $Hw \in P_D(HE)$ ist. Trifft dies zu, so gilt $d_{HE}w/dz = Hd_E w/dz$.
E und $\hat{E}$ heißen äquivalent bezüglich w, wenn mit $w \in P_D(E)$ auch $w \in P_D(\hat{E})$ gilt et vice versa ([1, 2, 7, 8]).

2. Ein Hauptterzeugendensystem

2.1. Das Ähnlichkeitsprinzip ([1, 2], Vekua: Reziprozitätsformel, [6, 5]) lehrt, daß die Lösungen der Vekua-Gleichung (1) in der Form $w = e^s f$ mit $s = T(w_{\bar{z}}/w)$ [3] und $f \in P_D(I)$ darstellbar sind. Ersichtlich lassen sich viele klassische Sätze mit Hilfe des Ähnlichkeitsprinzips auf die Theorie der pseudoholomorphen Funktionen verallgemeinert übertragen ([1, 2, 3, 5, 6]).

2.2. Satz: Es seien $w = F\varphi + G\psi \in P_D(E)$ bzw. $w = e^s f$ Darstellungen der Lösung von (1) in $D \subset\subset D_o$, letztere gemäß dem Ähnlichkeitsprinzip; dann sind E und $S := \langle e^s, ie^s \rangle$ äquivalente Erzeugendensysteme bezüglich w.

Beweis: Es ist e^s gleichmäßig stetig auf $\bar{D}$, und s_z und $s_{\bar{z}}$ existieren dort Hölder-stetig, also ist $S \in H_D^1 \times H_D^1$; da e^s auf $\bar{D}$ nicht verschwindet, ist $\mathrm{Im}(\overline{e^s} ie^s) = \mathrm{Im}(|e^s|^2) > 0$, also ist S ein Erzeugendensystem auf $\bar{D}$. Zudem gilt für $f = u + iv$ einfach $w = e^s u + ie^s v$ mit $e^s u_{\bar{z}} + ie^s v_{\bar{z}} = e^s f_{\bar{z}} = 0$ wegen der Holomorphie von f. Also ist $w = e^s f \in P_D(S)$. Die Umkehrung folgt aus der Reversibilität des Ähnlichkeitsprinzips.

Erzeugendensysteme dieser Struktur heißen auch Haupterzeugendensysteme ([7, 8]) und haben für viele Untersuchungen über funktionentheoretische Eigenschaften pseudoholomorpher Funktionen maßgebende Konsequenzen. Als einfache Folgerung hat man für die Ableitung modulo S mit $d_S w/dz = e^s u_z + ie^s v_z = e^s f'$ (vgl. 1.1) wieder ein Element aus $P_D(S)$, insbesondere ist also die Minimalperiode $\mu = 1$. (Natürlich ist auch e^s/ie^s als nur von y abhängig deutbar, siehe 1.4.)

2.3. Ursprünglich ([1, 6]) wurde $w = e^s f$ mit $f \equiv \mathrm{const}$ als verallgemeinerte Konstante bezeichnet; der vorstehende Satz gibt aber Anlaß, allgemeiner $w = F\varphi + G\psi \in P_D(E)$ mit (reell) konstanten φ und ψ als verallgemeinerte Konstante modulo E zu definieren. Der korellative Unterraum von $P_D(E)$ werde mit $P_D^o(E)$ bezeichnet. Insbesondere gilt für ein $w \in P_D(E)$, daß $\dot{w}(z) \equiv 0$ genau dann, wenn $w \in P_D^o(E)$!

3. Supplierende pseudoholomorphe Funktionen

Genau die Struktur pseudoholomorpher Funktionen mit holomorphem Faktor (Ähnlichkeitsprinzip) und Haupterzeugendendarstellung gibt Anlaß zu folgenden Überlegungen:

3.1. Es sei nun $w = F\varphi + G\psi \in P_D(E)$, und w genügt somit den (Pseudo-) Differentialgleichungen (4) bzw. (5). Frage: Unter welchen Voraussetzungen kann φ (und somit, wie sich herausstellt, simultan ψ) als Realteil (o.E.d.A.) einer holomorphen Funktion Φ (simultan Ψ, Satz 3.4., Satz 3.5.), also als konjugiert-harmonische Funktion interpretiert werden?
Vor der Kompatibilitätsprüfung (siehe 4.) seien einige Konsequenzen untersucht: Dazu sei $\tilde{w} := F\tilde{\varphi} + G\tilde{\psi}$ mit $\varphi + i\tilde{\varphi} =: \Phi \in P_D(I)$ und $\psi + i\tilde{\psi} =: \Psi \in P_D(I)$ angenommen [4].
Für das pseudoholomorphe Verhalten sind also im wesentlichen

(vgl. (4) und (3)) $F\tilde{\varphi}_{\bar{z}} + G\tilde{\psi}_{\bar{z}}$ und $F\tilde{\varphi}_{z} + G\tilde{\psi}_{z}$ zu untersuchen; wegen der Holomorphie von Φ und Ψ und wegen $w \in P_D(E)$ gilt aber $F\tilde{\varphi}_{\bar{z}} + G\tilde{\psi}_{\bar{z}} = iF\varphi_{\bar{z}} + iG\psi_{\bar{z}} = 0$ und $F\tilde{\varphi}_{z} + G\tilde{\psi}_{z} =$ $-iF\varphi_z - iG\psi_z = -i\dot{w} = -id_E w/dz$, d.h., unter den zusätzlichen Prämissen $\Phi, \Psi \in P_D(I)$ ist mit w auch $\tilde{w}$ (und somit auch $w + \tilde{w}$) ein Element von $P_D(E)$, und es gilt $\hat{w} = -i\dot{w}$ oder allgemeiner (vgl. die Definition ähnlicher Erzeugendensysteme, 1.5.)

$$\frac{d_E w}{dz} = i\,\frac{d_E \tilde{w}}{dz} = \frac{d_{iE}\, i\tilde{w}}{dz}. \tag{6}$$

3.2. <u>Definition:</u> Unter obiger Hypothese heißen w und $\tilde{w}$ supplierende pseudoholomorphe Funktionen modulo E erster Art, $p_E w$ und $p_E \tilde{w}$ supplierende pseudoholomorphe Funktionen modulo E zweiter Art.

3.3. <u>Definition:</u> Unter obiger Hypothese sei $w + i\tilde{w} =: W$, und $\dot{w} + i\hat{w} =: \dot{W}$. $W^* := w - i\tilde{w}$ heißt die suppletorische Konjugierte zu W.

Zunächst hat man Anlaß, die Funktion $W = w + i\tilde{w}$ zu untersuchen. Wegen (6) erhält man unmittelbar die (überbestimmten [5]) Formeln (vgl. 3.1.)

$$(F\varphi_{\bar{z}} + G\psi_{\bar{z}}) + (iF\tilde{\varphi}_{\bar{z}} + G\tilde{\psi}_{\bar{z}}) = F\Phi_{\bar{z}} + G\Psi_{\bar{z}} = 0$$

$$(F\varphi_{z} + G\psi_{z}) + (iF\tilde{\varphi}_{z} + iG\tilde{\psi}_{z}) = F\Phi' + G\Psi' \tag{7}$$

$$= \frac{d_E w}{dz} + \frac{d_{iE}\, i\tilde{w}}{dz} = \frac{d_E w}{dz} + i\,\frac{d_E \tilde{w}}{dz} = 2\,\frac{d_E w}{dz} = \dot{W}$$

bzw. (vgl. 1.2.)

$$F\varphi_{\bar{z}} + G\psi_{\bar{z}} = F\overline{\tfrac{1}{2}\Phi'} + G\overline{\tfrac{1}{2}\Psi'} = 0 \qquad F\tilde{\varphi}_{\bar{z}} + G\tilde{\psi}_{\bar{z}} = F\overline{\tfrac{i}{2}\Phi'} + G\overline{\tfrac{i}{2}\Psi'} = 0$$

$$F\varphi_{z} + G\psi_{z} = F\tfrac{1}{2}\Phi' + G\tfrac{1}{2}\Psi' = \dot{w} \qquad F\tilde{\varphi}_{z} + G\tilde{\psi}_{z} = F\tfrac{-i}{2}\Phi' + G\tfrac{-i}{2}\Psi' = -i\dot{w}$$

usw.

An dieser Stelle sei hervorgehoben, daß, unter den vereinbarten Annahmen, $W = w + i\tilde{w}$ in der Darstellung $W = F\Phi + G\Psi$ in zwei Summanden pseudoholomorpher Funktionen modulo der Haupterzeugendensysteme $\langle F, iF\rangle$ bzw. $\langle G, iG\rangle$ entsprechend 2.1. und 2.2. (Φ, Ψ holomorph) zerfällt erscheint.
Ist gemäß 1.4. E^1 Nachfolger von E, so daß $w \in P_D(E^1)$ ist, so gilt noch wegen (7)

$$\dot{W} = \frac{d_E w}{dz} + \frac{d_{iE} i\tilde{w}}{dz} = 2\,\frac{d_E w}{dz} = 2w \in P_D(E^1). \tag{8}$$

(7) und (8) rechtfertigen auch Definition 3.3. Vorbehaltlich des Kompatabilitätsnachweises haben wir also gezeigt

3.4. Satz:

$$[\ w = F\varphi + G\psi \in P_D(E) \wedge \Phi = \varphi + i\tilde{\varphi} \in P_D(I) \wedge \Psi = \psi + i\tilde{\psi} \in P_D(I)\ ] \Longrightarrow [\ \tilde{w} = F\tilde{\varphi} + G\tilde{\psi} \in P_D(E)\ ]$$ [6].

Umgekehrt gilt

3.5. Satz:

$$[\ w = F\varphi + G\psi \in P_D(E) \wedge \tilde{w} = F\tilde{\varphi} + G\tilde{\psi} \in P_D(E)^{6)} \wedge \Phi = \varphi + i\tilde{\varphi} \in P_D(I)\ ] \Longrightarrow [\ \Psi = \psi + i\tilde{\psi} \in P_D(I)\ ].$$

Beweis:

$$[\ 0 = (F\varphi_{\bar{z}} + G\psi_{\bar{z}}) + (iF\tilde{\varphi}_{\bar{z}} + iG\tilde{\psi}_{\bar{z}}) = F\Phi_{\bar{z}} + G\Psi_{\bar{z}} = G\Psi_{\bar{z}}\ ]$$

$$\Longrightarrow [\ \Psi_{\bar{z}} = 0\ ]$$ [7]. ∎

(Gemäß Definition 3.2. sind dann also w und $\tilde{w}$ bzw. $p_E w$ und $p_E\tilde{w}$ supplierende pseudoholomorphe Funktionen modulo E erster bzw. zweiter Art.)
Es seien also $w, \tilde{w} \in P_D(E)$ supplierende pseudoholomorphe Funktionen modulo E (3.2.); dann rechnet man zunächst leicht nach (1.3.), daß für die charakteristischen Koeffizienten

$$a_{iE} = a_E(=a), \qquad A_{iE} = A_E(=A),$$

$$b_{iE} = -b_E(=-b), \qquad B_{iE} = -B_E(=-B)$$

gilt, d.h., die Vekua-(Pseudo-) Differentialgleichungssysteme für w und $\tilde{w}$ (5) liefern durch Addition für $W = w + i\tilde{w}$ und ihre suppletorische Konjugierte (Definition 3.3.) $W^* = w - i\tilde{w}$ das System

$$W_{\bar{z}} - aW - b\overline{W^*} = 0, \qquad W_z - AW - B\overline{W^*} = \dot{W},$$

welches formal (5) entspricht.

3.6. Analog zur Theorie der pseudoholomorphen Funktionen erlaubt W eine Darstellung gemäß dem Ähnlichkeitsprinzip; denn setzt man $w = T(W_{\bar{z}}/W)$, so ist $\chi = e^{-w}W \in P_D(I)$, gilt doch $\chi_{\bar{z}} = e^{-w}(-w_{\bar{z}})W + e^{-w}W_{\bar{z}} = e^{-w}[(-a - b\overline{W^*}/W)W + aW + b\overline{W^*}\,] = 0$, χ beschränkt in einer Umgebung der Nullstellen von W.

4. Kompatibilität

4.1. Ist $w \in P_D^o(E)$, so sind alle Überlegungen von Abschnitt 3, unabhängig von F und G, per se erfüllt.

4.2. Es seien also $w = F\varphi + G\psi \in P_D(E)$; gemäß 1.4. ist $p_E w \in H_D^2$, also $p_E w \in C_D^2$. Zur Abkürzung sei $-iG/F =: \sigma + i\tau$, $-iF/G =: \hat{\sigma} + i\hat{\tau}$ (σ, τ, $\hat{\sigma}$, $\hat{\tau}$ reell); dann ist $F\varphi_{\bar{z}} + G\psi_{\bar{z}} = 0$ (1.2.) äquivalent mit dem elliptischen System (n.b. Im(FG > 0)

$$\varphi_x = \tau\psi_x + \sigma\psi_y\,, \qquad \psi_x = \hat{\tau}\varphi_x + \hat{\sigma}\varphi_y\,,$$

bzw.

$$\varphi_y = -\sigma\psi_x + \tau\psi_y\,, \qquad \psi_y = -\hat{\sigma}\varphi_x + \hat{\tau}\varphi_y,$$

woraus durch nochmalige Differentation (legitim), das System

$$\Delta\varphi + \hat{\alpha}\varphi_x + \hat{\beta}\varphi_y = 0\,, \qquad \alpha = \frac{\sigma_x+\tau_y}{\sigma} \quad \beta = \frac{\sigma_y-\tau_x}{\sigma}\,,$$

mit (9)

$$\Delta\psi + \alpha\psi_x + \beta\psi_y = 0\,, \qquad \hat{\alpha} = \frac{\hat{\sigma}_x+\hat{\tau}_y}{\hat{\sigma}} \quad \hat{\beta} = \frac{\hat{\sigma}_y-\hat{\tau}_x}{\hat{\sigma}}$$

hergeleitet werden kann. Interpretiert man φ bzw. ψ als Realteile holomorpher Funktionen $\Phi = \varphi + i\tilde{\varphi}$ bzw. $\Psi = \psi + i\tilde{\psi}$

(φ, ψ hinreichend glatt) und die entsprechenden Cauchy- Riemannschen Differentialgleichungen $\Phi_{\bar z} = (\varphi + i\tilde{\varphi})_{\bar z} = 0$ bzw. $\Psi_{\bar z} = (\psi + i\tilde{\psi})_{\bar z} = 0$ als Gleichungssysteme für die unbekannten Funktionen $\tilde{\varphi}$ bzw. $\tilde{\psi}$, so lauten die Integrabilitätsbedingungen für diese Systeme $\Delta\varphi = 0$ bzw. $\Delta\psi = 0$; sie führen mit (9) also auf die Bedingungen $\hat{\alpha}\varphi_x + \hat{\beta}\varphi_y = 0$ bzw. $\alpha\psi_x + \beta\psi_y = 0$. Man hat also zu ermitteln, wann dieses partielle Differentialgleichungssystem nicht-konstante Lösungen besitzt.

4.3. Eine Klasse von Lösungen bietet sich - unabhängig von φ und ψ - sofort aufgrund der Struktur von α, β, $\hat{\alpha}$, $\hat{\beta}$ ((9)) an: Ist nämlich G/F (und somit auch F/G) eine anti-holomorphe Funktion, so erfüllen $\sigma + i\tau$ bzw. $\hat{\sigma} + i\hat{\tau}$ die Systeme $(\sigma + i\tau)_z = 0$ bzw. $(\hat{\sigma} + i\hat{\tau})_z = 0$, oder

$$\sigma_x = -\tau_y\,, \qquad \hat{\sigma}_x = -\hat{\tau}_y\,,$$

$$\sigma_y = \tau_x\,, \qquad \hat{\sigma}_y = \hat{\tau}_x,$$

in (9) ist also $\alpha = \beta = \hat{\alpha} = \hat{\beta} = 0$, die Integrabilitätsbedingungen $\Delta\varphi = 0$ bzw. $\Delta\psi = 0$ sind somit erfüllt, und man findet bekanntlich $\tilde{\varphi}$ bzw. $\tilde{\psi}$, die dann bis auf eine reelle additive Konstante eindeutig bestimmt sind, in Form der in einem reellen Gebiet ($U_\varepsilon(z_o)$) wegunabhängigen reellen Kurvenintegrale

$$\begin{aligned}\tilde{\varphi}(z) &= \int_{z_o}^{z} \tilde{\varphi}_x dx + \tilde{\varphi}_y dy = \int_{z_o}^{z} -\varphi_y dx + \varphi_x dy,\\ \tilde{\psi}(z) &= \int_{z_o}^{z} \tilde{\psi}_x dx + \tilde{\psi}_y dy = \int_{z_o}^{z} -\psi_y dx + \psi_x dy,\end{aligned} \tag{10}$$

denn wegen $\Delta\varphi = 0$ bzw. $\Delta\psi = 0$ sind $-\varphi_y dx + \varphi_x dy$ bzw. $-\psi_y dx + \psi_x dy$ exakte Differentiale in $U_\varepsilon(z_o)$, die Funktionen $\varphi + i\tilde{\varphi} = \Phi$ bzw. $\psi + i\tilde{\psi} = \Psi$ sind also holomorphe Funktionen und bis auf jeweils rein imaginäre Konstanten eindeutig bestimmt.

4.4. Für den allgemeinen Fall überlegt man nochmals, daß die Integrabilitätsbedingung $\Delta\varphi = 0$ bzw. $\Delta\psi = 0$ genau dann erfüllt ist, wenn $\hat{\alpha}\varphi_x + \hat{\beta}\varphi_y = 0$ bzw. $\alpha\psi_x + \beta\psi_y = 0$ gelten; für vorgegebenes $E = \langle F,G\rangle$ werden die allgemeinen Integrale dieser homogenen Differentialgleichungen (α, β, $\hat{\alpha}$, $\hat{\beta}$ reell, glatt) nach der Charakteristikenmethode durch $\varphi = \eta(g(x,y))$ bzw. $\psi = \Theta(h(x,y))$ dargestellt, wenn $\varphi = c_\varphi$, $g(x,y) = c_g$ bzw.

$\psi = c_\psi$, $h(x,y) = c_h$ [8] die allgemeinen Lösungen der korrespondierenden charakteristischen Differentialgleichungssysteme $d\varphi = 0$, $dy/dx = \hat{\beta}/\hat{\alpha}$ bzw. $d\psi = 0$, $dy/dx = \beta/\alpha$ sind [8]. Die allgemeinen Lösungen $\varphi = \eta(g(x,y))$ bzw. $\psi = \Theta(h(x,y))$ sind dann die allen Bedingungen $\Delta\varphi = 0$ bzw. $\Delta\psi = 0$ und $F\varphi_{\bar z} + G\psi_{\bar z} = 0$ genügenden Funktionen, und man findet die konjugiertharmonischen Funktionen $\tilde\varphi$ bzw. $\tilde\psi$ und somit das Supplement $\tilde w = F\tilde\varphi + G\tilde\psi \in P_D(E)$ wieder nach (10).

5. Konklusion

Zusammenfassend läßt sich sagen, daß für $w \in P_D^o(E)$ per se bei antiholomorpher Abhängigkeit von G und F (vgl. 4.3.) jedes Element $w = F\varphi + G\psi$ aus $P_D(E)$ dortselbst eine Supplierende $\tilde w = F\tilde\varphi + \tilde G\tilde\psi$ besitzt, bei beliebig vorgegebenen $E = \langle F,G \rangle$ (vgl. 4.4.) je nach Lösungsangebot nach der Charakteristikenmethode ($\hat\beta/\hat\alpha$, β/α!). Es liegt nahe, insbesondere bei Berücksichtigung von Bemerkung 3.6., eine Funktionentheorie für Funktionen W supplierender Summanden zu entwickeln.

---

1 Hölder-stetig differenzierbar bzw. stetig differenzierbar.

2 Wo Mißverständnisse nicht zu befürchten sind, darf also die E-Indizierung unterdrückt werden.

3 Der T-Operator ist wie üblich definiert.

4 Ist $\Phi = \varphi + i\tilde\varphi \in P_D(I)$, so gelten bekanntlich die einfachen Rechenregeln: $\varphi_z = \frac{1}{2}\Phi' = i\tilde\varphi_z$, $\varphi_{\bar z} = \frac{1}{2}\overline{\Phi'} = -i\tilde\varphi_{\bar z}$; $\Phi_z = \overline{\Phi_{\bar z}}$, $\bar\Phi_{\bar z} = \overline{\Phi_z}$, $\overline{\varphi_z} = \varphi_{\bar z}$. Analog für $\Psi = \psi + i\psi \in P_D(I)$.

5 Man erfährt sogleich (Satz 3.4., Satz 3.5.), daß sich Φ und Ψ als simultan holomorph erweisen.

6 bzw. $i\tilde w =: F\tilde\varphi + iG\tilde\psi \in P_D(iE)$

7 $G \neq 0$ in D_o, Ψ hinreichend glatt.

8 Bekanntlich sind c_φ, c_g bzw. c_ψ, c_h reelle Konstanten, während η bzw. Θ gemäß Charakteristikenmethode willkürliche Funktionen bedeuten.

References:

1. Bers, L.: Theory of Pseudoanalytic Functions. New York 1953

2. Bers, L.: Local theory of pseudoanalytic functions. In: Kaplan, W. (ed.): Lectures on Functions of a Complex Variable. pp. 213-244. Ann Arbor 1955

3. Lanckau, E., and Tutschke, W.: Complex Analysis - Methods, Trends and Applications. Berlin 1983

4. Protter, M.H.: The periodicity problem for pseudoanalytic functions. Ann. of Math. (2) 64, 154-174 (1956)

5. Tutschke, W.: Partielle komplexe Differentialgleichungen in einer und mehreren komplexen Variablen. Berlin 1977

6. Vekua, I.N.: Verallgemeinerte Analytische Funktionen. Berlin 1963

7. Withalm, Cl.: Main generating systems of pseudoholomorphic functions. Erscheint in Math. Balkanica 11 (1981)

8. Withalm, Cl.: Über einen Integraloperator für pseudoholomorphe Funktionen modulo eines Haupterzeugendensystems. Complex Variables Theory Appl. 4, 2, 155-161 (1985)

eingegangen: 25. 4. 1989

Verfasser:

Univ.-Prof. Dr. Claudio Withalm
Institut für Mathematik
Karl-Franzens-Universität Graz
Elisabethstraße 16
Graz
A-8010

Rostock. Math. Kolloq. 40, 29 - 34 (1990) 35A40
35J55

Klaus Beyer

Zur Approximation mittels Laméscher Multipolpotentiale

Herrn Prof. Dr. L. Berg zu seinem 60. Geburtstag gewidmet

Bezeichne E_j die Spaltenvektoren der zu den Laméschen Gleichungen $Au = -\Delta u - \alpha\nabla \operatorname{div} u = 0$ $(\alpha > 0)$ gebildeten Grundlösungsmatrix

$$E_{ij}(x-y) = \frac{1}{8\pi}\left(\frac{\alpha}{1+\alpha}\frac{\partial^2 |x-y|}{\partial x_i \partial x_j} - 2\frac{\delta_{ij}}{|x-y|}\right). \tag{1}$$

Es ist wohlbekannt, daß ein auf dem Rand Γ eines beschränkten Gebietes $\Omega \subset \mathbb{R}^3$ definiertes Verschiebungsfeld $u(x) = (u_1(x), u_2(x), u_3(x))$ in der Norm des Raumes $L^2(\Gamma, \mathbb{R}^3)$ beliebig genau durch Linearkombinationen der Vektoren E_j approximiert werden kann, sofern nur die Pole y auf einer Nachbarfläche dicht liegen (siehe [9], wo dieser Umstand auch zur numerischen Konstruktion herangezogen wird). Wir beweisen hier die folgenden Verschärfungen:

Satz 1: Sei $\Gamma \subset \mathbb{R}^3$ eine geschlossene Fläche der Regularitätsklasse C^∞, und Γ berande ein beschränktes Gebiet Ω mit zusammenhängendem Komplement. $\tilde{\Gamma}$ sei eine Nachbarfläche mit den gleichen Eigenschaften und $\operatorname{dist}(\Gamma, \tilde{\Gamma}) > 0$. Liegt nun die Punktfolge $(x_j)_{j\geq 1} \subset \tilde{\Gamma}$ auf $\tilde{\Gamma}$ dicht, dann bilden die Vektoren $E_i(x-x_j)$, $i = 1,2,3$, $j \geq 1$, eine in jedem der Sobolewräume $H^s(\Gamma, \mathbb{R}^3)$, $s \geq 0$, totale Menge; ihre endlichen Linearkombinationen liegen also in diesen Räumen dicht.

Nach Satz 2 (siehe unten) genügt es, die Pole so zu verteilen, daß aus dem Verschwinden einer analytischen Funktion f auf der Trägermenge schon $f = 0$ gefolgert werden kann, - wie in [7] heiße die Trägermenge dann fundamental. Z.B. (siehe [10]) gilt der folgende

Hilfssatz: Sei $f(z) = f(z_1,\dots,z_n) = \sum a_\alpha z^\alpha$ konvergent für $|z_j| < r_j$. Seien ferner $(z_{jk})_{k\geq 1}$, $0 < |z_{jk}| < r_j$, n Folgen derart, daß

$$\lim_{k\to\infty} z_{jk} = 0 \quad \text{für} \quad j = 1,\dots,n.$$

Verschwindet f dann in jedem Punkt $(z_{1k_1},\dots,z_{nk_n})$, wobei $k_1,\dots,k_n$ unabhängig voneinander die Werte 1,2,... durchlaufen, so verschwindet jeder Koeffizient der Reihe.

Satz 2: Ist $(x_j)_{j\geq 1}$ eine außerhalb Γ gelegene fundamentale Folge mit $\mathrm{dist}((x_j), \Gamma) > 0$, so ist das System $E_i(x-x_j)$ ebenfalls total in den Räumen $H^s(\Gamma, \mathbb{R}^3)$.

Ersetzt man das obige System durch die zu einem fixierten Pol $x_o \notin \Gamma$ gebildeten Multipolpotentiale

$$D^\alpha E_i(x-x_o), \quad |\alpha| \geq 0, \quad i = 1,2,3, \tag{2}$$

dann gilt ähnlich

Satz 3: Unter den gleichen Voraussetzungen über Γ bildet (2) eine in jedem der Räume $H^s(\Gamma, \mathbb{R}^3)$ totale Menge.

Wir weisen auf den Zusammenhang zu den Approximationssätzen von Beckert [2] (vgl. auch [4,5,12]) hin. Für skalare elliptische Gleichungen mit konstanten Koeffizienten wurden ähnliche Dichtheitssätze von Hamann [6] bewiesen. Unsere Überlegungen folgen der Arbeit [3].

Wir bemerken noch, daß die Sätze 1 - 3 sinngemäß auch bei der Approximation mittels Spannungsverteilungen

$$(Tu)_i = \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) n_j + n_i(\alpha-1)\,\mathrm{div}\,u \tag{3}$$

gültig bleiben. $(Tu)_i$ bezeichnet hier die i-te Komponente des Bildvektors Tu und $n = (n_1,n_2,n_3)$, $|n| = 1$, die Außennormale von Γ (bezüglich Ω); über doppelte Indizes ist zu summieren. Natürlich sind jetzt die bekannten Lösbarkeitsbedingungen für die zweite Randwertaufgabe der linearen Elastizitätstheorie zu berücksichtigen.

Im Folgenden wird Satz 3 bewiesen, die Beweise der beiden ersten

Sätze verlaufen ähnlich. Wir notieren zunächst die für den Laméschen Operator A und seine Iterierten

$$A^m u = (-1)^m(\Delta^m u + ((\alpha+1)^m - 1)\nabla\Delta^{m-1}\mathrm{div}u)$$

gültigen Greenschen Formeln:

$$a(u,v) := \int_\Omega ((\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i})\frac{\partial v_i}{\partial x_j} + (\alpha-1)\frac{\partial u_i}{\partial x_i}\frac{\partial v_j}{\partial x_j})dx$$

$$= \int_\Omega v\ Au\ dx + \int_\Gamma v\ Tu\ d\Gamma$$

und

$$a(A^m u, A^m v) = \int_\Omega v\ A^{2m+1}u\ dx$$

$$+ \sum_{i=o}^{m} \int_\Gamma A^i v\ TA^{2m-i}u\ d\Gamma - \sum_{i=o}^{m-1} \int_\Gamma TA^i v\ A^{2m-i}u\ d\Gamma, \qquad (4)$$

wobei T den in (3) eingeführten Spannungsoperator bezeichnet. Wir weisen noch darauf hin, daß die Operatoren A^i, TA^j ($O\leq i\leq m$, $O\leq j\leq m-1$) längs Γ ein Dirichletsystem (siehe [11]) der Ordnung $2m+1$ bilden.

Die folgenden Überlegungen beziehen sich auf Felder mit dem Definitionsbereich $\mathbb{R}^3$. Bezeichne V^m den Faktorraum

$$V^m = \{\ u = (u_1,u_2,u_3):\ u_i \in \mathcal{D}'(\mathbb{R}^3),\ D^\alpha u_i \in L^2(\mathbb{R}^3) \text{ für } i = 1,2,3 \text{ und } |\alpha| = 1,m\ \}/\mathbb{R}^3.$$

V^m ist Hilbertraum mit der Norm

$$\|u\|_m^2 = \sum_{i=1}^{3} \sum_{|\alpha|=m} \frac{m!}{\alpha!} \int_{\mathbb{R}^3} |D^\alpha u_i|^2 dx + \sum_{i=1}^{3} \int_{\mathbb{R}^3} |\nabla u_i|^2 dx.$$

Felder der Regularitätsklasse C_o^∞ liegen in V^m dicht. Dafür gilt zudem

$$a(A^m u, A^m v) = \int ((\frac{\partial}{\partial x_j}\Delta^m u_i + \frac{\partial}{\partial x_i}\Delta^m u_j)$$

$$+ 2((\alpha+1)^m-1)\frac{\partial^2}{\partial x_i \partial x_j}\Delta^{m-1}\mathrm{div}u(\frac{\partial}{\partial x_j}\Delta^m u_i$$

$$+ ((\alpha+1)^m-1)\frac{\partial^2}{\partial x_i \partial x_j}\Delta^{m-1}\mathrm{div}u))dx + (\alpha-1)\int |\mathrm{div}\ A^m u|^2 dx$$

$$= \int \frac{\partial}{\partial x_j}\Delta^m u_i \frac{\partial}{\partial x_j}\Delta^m u_i \, dx + ((\alpha+1)^{2m+1}-1) \int |\Delta^m \operatorname{div} u|^2 dx$$

$$\geq \int \frac{\partial}{\partial x_j}\Delta^m u_i \frac{\partial}{\partial x_j}\Delta^m u_i \, dx$$

$$= \sum_{i=1}^{3} \sum_{|\alpha|=2m+1} \frac{|\alpha|!}{\alpha!} \int |D^\alpha u_i|^2 dx$$

(Integration über $\mathbb{R}^3$), also

$$a(A^m u, A^m u) + a(u,u) \geq \|u\|^2_{2m+1} \tag{5}$$

für $u \in C_0^\infty(\mathbb{R}^3,\mathbb{R}^3)$.

Wir nehmen nun an, daß im Widerspruch zur Behauptung von Satz 3 ein von Null verschiedenes lineares Funktional

$$f \in H^{-2m-(1/2)}(\Gamma,\mathbb{R}^3)$$

existiert mit der Eigenschaft

$$\langle f, D^\alpha E_i \rangle = 0 \quad \text{für} \quad i = 1,2,3$$

und alle Multiindizes α. $D^\alpha E_i$ ist hier (nach Einschränkung auf Γ) als Element von $H^{2m+(1/2)}(\Gamma,\mathbb{R}^3)$ anzusehen, und die eckigen Klammern kennzeichnen den Wert des linearen Funktionals.

Wegen (5) und der stetigen Einbettung $V^m \subset L^6(\mathbb{R}^3,\mathbb{R}^3)$ für $m \geq 1$ besitzen die Variationsgleichungen

$$a(A^m u, A^m v) + a(u,v) = \langle f,v \rangle \quad \text{für alle} \quad v \in V^{2m+1} \tag{6}$$

eine eindeutig bestimmte Lösung $u \in V^{2m+1}$ Der Vergleich mit (4) zeigt, daß u den Gleichungen

$$A^{2m+1}u + Au = 0 \quad \text{in} \quad \Omega \quad \text{und in} \quad \complement\Omega \tag{7}$$

sowie den längs Γ geltenden Sprungrelationen

$$[A^j u] = [TA^{j-1}u] = 0 \quad \text{für} \quad m+1 \leq j \leq 2m,$$
$$[TA^{2m}u + Tu] = f \tag{8}$$

genügt; die geraden Klammern bezeichnen die Differenz des inneren und äußeren Grenzwerts (der in bekannter Weise im distributionellen Sinn aufzufassen ist, vgl. [1]).

Im nächsten Schritt konstruiere man die Regularisierten

$$E_{ij}^{\varepsilon} = \psi^{\varepsilon} * E_{ij}(\cdot - x_o),\ \varepsilon > 0,$$

der Kerne E, wobei $0 \leq \psi \in C_o^{\infty}(\mathbb{R}^3)$, $\int \psi\, dx = 1$ und $\psi^{\varepsilon}(x) = \varepsilon^{-3}\psi(\frac{x}{\varepsilon})$ zu wählen ist. Wegen $\frac{\partial}{\partial x_k} D^{\alpha}E_{ij}^{\varepsilon} = D^{\alpha}\psi^{\varepsilon} * \frac{\partial}{\partial x_k} E_{ij}$ und der Hardy-Littlewood-Sobolewschen Ungleichung (siehe [8]) gilt hierfür

$$\left\| \frac{\partial}{\partial x_k} D^{\alpha}E_{ij}^{\varepsilon};\ L^2(\mathbb{R}^3) \right\| \leq C \| D^{\alpha}\psi^{\varepsilon};\ L^{6/5}(\mathbb{R}^3) \|.$$

Die Ableitungen $D^{\alpha}E_{ij}^{\varepsilon}$ gehören demzufolge zu den Räumen V^m. Einsetzen von $D^{\alpha}E_i^{\varepsilon} = (D^{\alpha}E_{1i}^{\varepsilon}, D^{\alpha}E_{2i}^{\varepsilon}, D^{\alpha}E_{3i}^{\varepsilon})^T$ in (6) liefert

$$a(A^m u, A^m D^{\alpha}E_i^{\varepsilon}) + a(u, D^{\alpha}E_i^{\varepsilon}) = \langle f, D^{\alpha}E_i^{\varepsilon} \rangle. \tag{9}$$

Wegen $x_o \notin \Gamma$ konvergiert die rechte Seite dieser Beziehung gegen $\langle f, D^{\alpha}E_i \rangle = 0$, wenn ε gegen Null strebt. Partielle Integration der linken Seite führt zu

$$a(A^m u, A^m D^{\alpha}E_i^{\varepsilon}) + a(u, D^{\alpha}E_i^{\varepsilon})$$

$$= \int (A^{2m+1}D^{\alpha}E_i^{\varepsilon} + AD^{\alpha}E_i^{\varepsilon})u\, dx$$

$$= (-1)^{|\alpha|} \int \psi^{\varepsilon}(x-x_o)(D^{\alpha}(A^{2m} + 1)u)_i\, dx,$$

wobei $(AE_i^{\varepsilon})_j = (\psi^{\varepsilon} * AE_i)_j = \psi^{\varepsilon}(x-x_o)\delta_{ij}$ berücksichtigt wurde. Aus (9) folgt deshalb nach dem Grenzübergang $\varepsilon \to 0$

$$(D^{\alpha}(A^{2m} + 1)u)_i(x_o) = 0 \quad \text{für} \quad i = 1,2,3$$

und alle Ableitungen D^{α}. Hieraus folgt wegen (7): $A^{2m}u + u = 0$ in der durch x_o bestimmten zusammenhängenden Komponente Ω^+ von $\mathbb{R}^3 \backslash \Gamma$. Bezeichnet die Signatur "+" den von Ω^+ aus gebildeten Grenzwert, so gilt mithin $A^{2m}u^+ + u^+ = 0$ längs Γ. Ein Blick auf (8) lehrt: $A^{2m}u^- + u^- = 0$ für den entsprechenden Grenzwert von der komplementären Komponente Ω^- her. Da $A^{2m}u + u$ auch in Ω^- Lösung der Laméschen Gleichungen ist, folgt $A^{2m}u + u = 0$ in Ω^- wegen der eindeutigen Lösbarkeit des Dirichletproblems für diese Gleichungen. (8) impliziert dann $f = [T(A^{2m}u + u)] = 0$. Satz 3 ist damit bewiesen.

Literatur:

1. Aubin, J.-P.: Approximation of Elliptic Boundary-Value Problems. New York 1972

2. Beckert, H.: Eine bemerkenswerte Eigenschaft der Lösungen des Dirichletschen Problems bei linearen elliptischen Differentialgleichungen. Math. Ann. 139, 255-264 (1960)

3. Beyer, K.: Approximation durch Lösungen elliptischer Differentialgleichungen. Rostock. Math. Kolloq. 26, 27-34 (1984)

4. Göpfert, A.: Über L_2-Approximationssätze - eine Eigenschaft der Lösungen elliptischer Differentialgleichungen. Math. Nachr. 31, 1-24 (1966)

5. Hamann, U.: Approximation durch Lösungen elliptischer Randwertprobleme auf geschlossenen Hyperflächen. Math. Nachr. 136, 285-301 (1988)

6. Hamann, U.: Approximation mittels Linearkombinationen von Fundamentallösungen elliptischer Differentialoperatoren. Ersch. in Math. Zeitschr.

7. Hermann, P., und Kersten, H.: Die Lösung der prae-Maxwellschen Gleichungen mit Hilfe vollständiger Lösungssysteme. Math. Methods. Appl. Sci. 2, 410-418 (1980)

8. Hörmander, L.: The Analysis of Linear Partial Differential Operators I. Berlin 1983

9. Kupradse, V.D.: Dreidimensionale Aufgaben der Mathematischen Elastizitäts- und Thermoelastizitätstheorie (russ.). Moskau 1976

10. Osgood, W.F.: Lehrbuch der Funktionentheorie II. Leipzig 1929

11. Schechter, M.: General boundary value problems for elliptic partial differential equations. Comm. Pure Appl. Math. 12, 457-486 (1959)

12. Wildenhain, G.: Approximation in Sobolew-Räumen durch Lösungen allgemeiner elliptischer Randwertprobleme bei Gleichungen beliebiger Ordnung. Rostock. Math. Kolloq. 22, 43-56 (1983)

eingegangen: 24. 1. 1989

Verfasser:

Prof. Dr. Klaus Beyer
Karl-Marx-Universität Leipzig
Sektion Mathematik
Karl-Marx-Platz
Leipzig
DDR-7010

Rostock. Math. Kolloq. 40, 35 - 50 (1990) 43A32
94A11

Walter Schempp

Elementary holograms, artificial neural networks, and theta - null values

Herrn Prof. L. Berg zu seinem 60. Geburtstag gewidmet

> Our struggles with digital computers have taught us much about how neural computation is not done; unfortunately, they have taught us relatively little about how it is done.
>
> Carver A. Mead (1989)

Abstract: Based on a unified nilpotent harmonic analysis approach to artificial neural net work models implemented with coherent optical or analog elelctronic neurocomputer architectures, the papaer establishes a new identity for the matching polynomials of complete bichromatic graphs.

1. Introduction. Real-time image processing, computer vision, automatic target recognition in robotics, speech understanding, and other areas of artificial intelligence (AI) need to process extremely large amounts of data with very high velocity. The problem of large-volume and high-speed computations can be solved by

- o data compression techniques,
- o parallel data processing.

Since their very beginning, artificial neural networks have been considered as massively parallel computing paradigms. The fundamental characteristics of all neurocomputer architectures are the interconnections of arrays of simple processing elements to form a concurrent distributed processing network of extensive connectivity. Large scale (LS) collective systems like artificial neural networks exhibit many properties, including robustness, reliability, and fault tolerance, an ability to deal with ill-posed problems and noisy data, which conventional digital computer architectures do not. Neurobiology provides existence theorem on effectiveness of neural network parallel algorithms on appropriate problems.

For artificial neural networks to become ultimately useful, neuromorphic hardware must be developed. Development efforts in the field of sixth generation computers have concentrated on one of two goals: to build

- o efficient hardware that effectively executes software simulations,
- o actual hardware emulators for specific neural network models.

Examples of the first are the Hecht-Nielsen Neurocomputer (HNC) accelerator board for conventional serial personal computers, and the Delta board by Science Applications International Corporation (SAIC). An important application of the SAIC neural network software simulation is the detection of explosives in checked airline baggage: the luggage is bathed in low energy (thermal) neutrons and the gamma rays resulting from neutron absorption by atomic elements in the luggage are analyzed. The artificial neural network software then searches for specific combinations of atomic elements that characterize explosives including dynamites and water gels.

Examples of the second are arrays of coherent optical processors ([2], [3], [18], [19], [20], [32], [34]) for the implementation of neural network models by holographic interconnections, and neural network analog very large scale integrated (VLSI) chips. For instance, the silicon models of the orientation-selective retina for pattern recognition ([14], [15], [16], [1]), and the analog electronic cochlea for auditory localization ([13], [14], [15]) belong to this category. The retinal and the cochleal VLSI chip are made with a standard complementary MOS (CMOS) process.

Although the implementation of the various neural network models needs to overcome many difficult design problems, their performance is modest compared with the powerful organizing principles found in biological neural wetware. The visual system of a single human being does more image processing than do the entire world's supply of supercomputers, and the nervous system of even a very simple animal like the common house-fly (Musca domestica) contains computing paradigms that are orders of magnitude more effective than are those found in systems made by humans. Presently the most advanced neural network analog VLSI chips model, to a first approximation, the time-frequency domain processing of two highly spectacular biological neural systems: the active auditory localization system of the horseshoe bats (Rhinolophidae), and the passive auditory localization system of the barn owl (Tyto alba) which both produce complete maps of the auditory space from the time-frequency coding pathways. Continuing evolution, however, of technology and of neuromathematics, the highly promising new field of studying how computations can be carried out in extensive networks formed by arrays of heavily interconnected simple processing elements, will create neurocomputers within the next decade which will be able to solve problems intractable for even the largest digital computers.

This paper concentrates on a unified approach to massively parallel coherent optical and analog electronic neurocomputer architectures which is based on harmonic analysis of the three-dimensional Heisenberg nilpotent Lie group. As a result, the analysis provides total approximating families in $L^2(\mathbb{R} \oplus \mathbb{R})$ of correlating and decorrelating code primitives of artificial neural networks. Finally, a series of new identities for theta-null values shows that studies in computational mathematics combined with synthetic neurobiology may have a spin-off in pure mathematics.

2. The holographic transform. Let $\mathcal{S}(\mathbb{R})$ denote the Schwartz space of complex-valued $\mathcal{C}^\infty$ functions on the real line $\mathbb{R}$ rapidly decreasing at infinity. Consider $\mathcal{S}(\mathbb{R})$ as a dense vector subspace of the complex Hilbert space $L^2(\mathbb{R})$ under its natural isometric embedding. In optical holography, a square-law detector encodes in a massively parallel way the optical path length difference $x \in \mathbb{R}$ and the phase difference $y \in \mathbb{R}$ of two coherent signals having the same carrier and their amplitudes ψ in φ in the space $\mathcal{S}(\mathbb{R})$ by simultaneously recording the coordinates (x,y) of the interference pattern written by the two-wave mixing $\psi \otimes \varphi$ into the hologram plane $\mathbb{R} \oplus \mathbb{R}$. The sesquilinear extension to $\mathcal{S}(\mathbb{R}) \otimes \mathcal{S}(\mathbb{R})$ of the mapping

$$\psi \otimes \varphi \mapsto H(\psi,\varphi;x,y) = \int_{\mathbb{R}} \psi(t-x)\overline{\varphi}(t)e^{2\pi i y t}dt$$

describes by semi-linear superposition the holographic angle encoding: each object to be globally stored by the coherent object signal beam in the hologram is encoded prior to its recording by mixing (or heterodyning) an unfocused linearly polarized coherent non-object-bearing reference signal beam having a particular angle between its wave vector and the normal vector of the hologram plane $\mathbb{R} \oplus \mathbb{R}$. Therefore the sesquilinear mapping $\psi \otimes \varphi \mapsto H(\varphi,\psi;.,.)$ is called the holographic transform of the writing amplitudes ([27], [28], [29]). The method of holography applies to all waves: to electron waves, X rays, light waves, acoustic waves, and seismic waves, providing the waves are coherent enough to form the required interference patterns in the hologram plane ([29]). In radar analysis, H is called the cross-ambiguity function ([23], [7], [21]). In the following it will be convenient to define the auto-ambiguity function by $H(\psi;.,.) :\doteq H(\psi,\psi;.,.)$.

Remark 1. High-resolution radar imagery of the terrain and optical holographic imaging are closely related concepts. In fact, airborne and spaceborne synthetic aperture radar (SAR) remote sensing systems use microwave holograms for data storage and can therefore be regarded as optical neurocomputers which implement a Doppler filter bank by a relatively static reflection pattern of the architecture mirror ([29]). The massive parallelism inherent to the optical data processing approach is in large part responsible for the success of SAR imaging.

Remark 2. Since the advent of optical holography or coherent wavefront reconstruction, there has been a strong interest in replacing the lenses used in optical systems by holographic optical elements (HOEs). In particular, optical SAR data processing systems may be realized by optical heads which include HOEs. Many HOEs are fabricated by recording the interference pattern between two mixing laser beams. The use of digital computer-generated hologram (CGH) techniques, however, avoids the technological difficulties involved in the interferometric HOE fabrication. Moreover, one benefit that digital CGHs can offer that is not available with optical holography is the ability to deal with objects that exist only mathematically. High quality digital CGHs may be fabricated with the same technology used in the manufacture of VLSI circuits. A digital computer controlled output device such as an electron-beam high-resolution microlithographic system writes the desired geometric pattern on photoresist, which is subsequently processed to produce the finished transmissive or reflective holographic element. Alternately, digital CGHs may be realized by writing the appropriate pattern on a spatial light modulator (SLM). In any case, digital CGHs form a necessarily imperfect bridge between digital computer and optical neurocomputer architectures.

Remark 3. A vital element of optical neurocomputer architectures is the medium for optical hologram recording because it plays the rôle of a holographic associative memory. Electro-optical photorefractive crystals (PRCs) are known to form reusable holographic storage materials that can be infinitely recycled and do not require additional processing. The crystals of the sillenite family, bismuth silicon oxide $Bi_{12}SiO_{20}$ (BSO), bismuth titanium oxide $Bi_{12}TiO_{20}$ (BTO), and bismuth germanium oxide $Bi_{12}GeO_{20}$ (BGO) exhibit the highest sensitivity to light among presently known PRCs ([33]). Optical holograms are recorded inside PRCs directly by illuminating the crystal with laser light. The light induces a charge redistribution inside the crystal ([8]) and in a certain characteristic time interval a dynamic equilibrium between distributions of the recording light intensity and internal electric charge is established. The electric charge induces an internal electrostatic field that changes the refractive index of the crystal by the electro-optical effect and forms a volume phase hologram. As the interference pattern undergoes changes, a new charge distribution is formed, hence a new optical hologram is recorded. This charge distribution again comes to a dynamic equilibrium with the recording interference pattern. If the period during which the interference pattern changes is sufficiently long, the electro-optical crystal rerecords an optical hologram. Hence the electro-optical PRCs can adapt itself to varying external conditions, such as occasional temperature-induced changes of the phase difference between the writing object signal beam and reference signal beam, or mechanical instabilities. This is an extremely important feature because it allows more reliable storage of scattering objects by almost-real-time holography.

3. The Heisenberg nilpotent Lie group. Let G denote the multiplicative group of all unipotent real matrices

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \quad := \quad (x,y,z)$$

Then G is a two-step nilpotent Lie group with one-dimensional center $C = \{(0,0,z) \mid z \in \mathbb{R}\}$; G is a realization of the three-dimensional Heisenberg group ([23]) with Lie algebra $\mathfrak{g}$ formed by the upper triangular matrices $\{(x,y,z)-(0,0,0) \mid x,y,z \in \mathbb{R}\}$. For each real number $\nu \neq 0$ the central character χ_ν: $(0,0,z) \mapsto e^{2\pi i\nu z}$ determines up to an isomorphism a unique infinite-dimensional irreducible unitary linear representation U_ν of G in $L^2(\mathbb{R})$ which acts on the vector subspace $\mathcal{S}(\mathbb{R})$ according to the rule

$$U_\nu(x,y,z)\psi(t) \quad = \quad e^{2\pi i\nu(z+yt)}\psi(t-x) \qquad (t \in \mathbb{R}).$$

Let $\bar{U}_\nu$ denote the contragredient representation of U_ν, so that

$$\bar{U}_\nu(x,y,z) \; = \; {}^tU_\nu((x,y,z)^{-1})$$

holds for all elements $(x,y,z) \in G$. Obviously

$$U_\nu|C = \chi_\nu, \quad \bar{U}_\nu|C = \chi_{-\nu} \qquad (\nu \in \mathbb{R},\ \nu \neq 0).$$

The flatness of the affine Kirillov coadjoint orbits $\mathcal{O}_\nu$ and $\mathcal{O}_{-\nu}$ associated with U_ν and $\bar{U}_\nu$ in the dual $\mathfrak{g}^*$ of the Heisenberg Lie algebra $\mathfrak{g}$, respectively, is equivalent to the square integrability modulo C of U_ν and $\bar{U}_\nu$. From these facts the central projection G-slice theorem follows:

Theorem 1. *The holographic transform is the coefficient function of the linear Schrödinger representation* U_1 *of the polarized Heisenberg group* G *projected along the center* C *onto* G/C, *i.e., the identities*

$$\begin{cases} H(\psi',\varphi';x,y) = \langle U_1(x,y,0)\psi' \mid \varphi'\rangle \\ \bar{H}(\psi,\varphi;x,y) = \langle \bar{U}_1(x,y,0)\bar{\psi} \mid \bar{\varphi}\rangle \end{cases}$$

hold for all points $(x,y) \in \mathbb{R} \oplus \mathbb{R}$.

The importance of the preceding result lies in the fact that the hidden symmetries of the holographic transform H can be expressed by the group of automorphisms of the Heisenberg nilpotent Lie group G keeping the center C pointwise fixed. This group, the metaplectic group **Mp**(1,$\mathbb{R}$), forms a twofold cover of the symplectic group **Sp**(1,$\mathbb{R}$) acting on the hologram plane $\mathbb{R} \oplus \mathbb{R}$ ([23]).

4. Readout of optical holograms. In the following, the isomorphic G-manifolds $\mathcal{O}_1 \in \mathfrak{g}^*/\mathrm{CoAd}(G)$, $\mathcal{O}_{-1} \in \mathfrak{g}^*/\mathrm{CoAd}(G)$, and the central projection G-slice G/C will be identified with the hologram plane $\mathbb{R} \oplus \mathbb{R}$. An application of Schur's lemma provides the biorthogonality relations ([23], [22], [31])

$$\iint_{\mathbb{R}\oplus\mathbb{R}} H(\psi',\varphi';x,y)\overline{H}(\psi,\varphi;x,y)dxdy \;=\; \langle\psi'\otimes\varphi|\psi\otimes\varphi'\rangle$$

for $\psi',\varphi',\psi,\varphi$ in $\mathscr{S}(\mathbb{R})$. Therefore the dyads

$$\begin{cases} E(\psi',.;x,y):\; \varphi' \mapsto H(\psi',\varphi';x,y)\overline{U}_1(x,y,0)\overline{\psi}' \\ \overline{E}(\psi,.;x,y):\; \varphi \mapsto \overline{H}(\psi,\varphi;x,y)U_1(x,y,0)\psi \end{cases} \qquad ((x,y) \in \mathbb{R}\oplus\mathbb{R})$$

which embed $\psi' \in \mathscr{S}(\mathbb{R})$ and $\psi \in \mathscr{S}(\mathbb{R})$, respectively, into the Hilbert-Schmidt (HS) operators acting on $L^2(\mathbb{R})$, define a U_1-system $(E(.,.;x,y))_{(x,y)\in\mathbb{R}\oplus\mathbb{R}}$, and a $\overline{U}_1$-system $(\overline{E}(.,.;x,y))_{(x,y)\in\mathbb{R}\oplus\mathbb{R}}$ of coherent states based on the hologram plane $\mathbb{R}\oplus\mathbb{R}$ ([17]).

Theorem 2. *For all writing amplitudes $\psi',\varphi',\psi,\varphi$ in $\mathscr{S}(\mathbb{R})$ the gain equations*

$$\iint_{\mathbb{R}\oplus\mathbb{R}} E(\psi',\varphi';x,y)dxdy \;=\; \|\psi'\|_2\overline{\varphi}'$$

$$\iint_{\mathbb{R}\oplus\mathbb{R}} \overline{E}(\psi,\varphi;x,y)dxdy \;=\; \|\psi\|_2\varphi$$

hold.

Remark 4. Similar inversion formulas can be established for the affine coherent states defined by the wavelet transform and the square integrable irreducible unitary linear representations of the non-unimodular affine Lie group of the real line $\mathbb{R}$. Wavelets are particularly useful code primitives for voice decomposition ([9]).

Remark 5. Turning from optical holography to computer-aided tomography ([5]), the preceding identities give rise by an application of the theory of the reductive dual pair $(\tilde{Sp}(1,\mathbb{R}),\tilde{O}(n,\mathbb{R}))$, ([25], [11]), to the singular value decomposition of the Radon transform $\mathcal{R}$: $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}\times S_{n-1})$ acting on functions $f \in \mathcal{S}(\mathbb{R}^n)$ according to

$$\mathcal{R}f(r,\omega) = \int_{\mathbb{R}^n} f(x)\varepsilon_{(r-\langle\omega|x\rangle)}dx.$$

It follows that the inversion problem for the Radon transform $\mathcal{R}$ which underlies computer-aided tomography (CT) is ill-posed. Neurocomputers, however, seem to be more appropriate to solve ill-posed problems than conventional digital computers.

As a special case we obtain from Theorem 2 supra the following result which describes the readout procedure of optical holograms.

<u>Corollary.</u> *Let $\varphi \in \mathcal{S}(\mathbb{R})$ and assume that $\psi \in \mathcal{S}(\mathbb{R})$ satisfies the normalization condition $\|\psi\|_2 = 1$. If $\mathcal{F}$ denotes the Fourier transform acting on $\mathcal{S}(\mathbb{R})$ then the reproducing scattering integrals of degenerate four-wave mixing*

$$\left\{\begin{aligned} \iint_{\mathbb{R}\oplus\mathbb{R}} H(\psi,\varphi;x,y)e^{-2\pi iyt}\bar{\psi}(t-x)dxdy &= \bar{\varphi}(t) \\ \iint_{\mathbb{R}\oplus\mathbb{R}} H(\psi,\varphi;y,x)e^{2\pi iyt}\mathcal{F}\bar{\psi}(t-x)dydx &= \mathcal{F}\bar{\varphi}(t) \end{aligned}\right. \qquad (t \in \mathbb{R})$$

hold.

The preceding integral equations prove the holographic reciprocity principle which governs the angle decoding of optical holograms: The amplitude and the phase of the conjugate object signal can be read out simultaneously by illuminating the hologram with the unfocused conjugate reference signal beam. The pair of reproducing scattering integrals describing the holographic filter bank are also at the basis of optical wavefront conjugation by means of real-time holography ([8]) in electro-optical PRCs. Wavefront conjugate mirrors provide retroreflection and optical tracking novelty filters. Therefore, Theorem 2 is at the basis of neural network models implemented by holographic interconnections in optical neurocomputer architectures ([2], [3], [18], [19], [20], [32]). If the holographic associative memory has net gain comparable with the losses in the resonator cavity, the output will converge to a real image of the globally stored object: the expanded conjugate reference signal beam acts as an optical scanner for readout of the associate information. In case of a linear resonator memory, gain is supplied by the wavefront conjugate mirror which provides regenerative feedback, whereas in case of a loop resonator memory, gain is supplied by an externally pumped electro-optical PRC.

5. Radial isotropy. A writing amplitude $\psi \in \mathscr{S}(\mathbb{R})$ is called radially isotropic if $H(\psi;.,.)$ is a radial function on the hologram plane $\mathbb{R} \oplus \mathbb{R}$, i.e., if $H(\psi;.,.)$ is invariant under the natural action of the orthogonal group $O(2,\mathbb{R})$.

Theorem 3. *The amplitude $\psi \in \mathscr{S}(\mathbb{R})$ is radially isotropic if and only if it admits the form of Hermite-Gaussian eigenmodes*

$$\psi = \zeta_n H_n$$

where $\zeta_n \in \mathbb{C}$ is a constant and $H_n(t) = e^{-t^2/2} h_n(t)$ is the Hermite function of degree $n \geq 0$.

The proof follows by classifying the irreducible unitary linear representations of the diamond solvable Lie group $\mathbb{T} \times G$ having U_ν as their restrictions to G ([24]).

The elementary holograms $(H(H_m, H_n;.,.))_{m\geq 0, n\geq 0}$ form a total orthogonal approximating family in the complex Hilbert space $L^2(\mathbb{R} \oplus \mathbb{R})$, hence a decorrelating family of code primitives. The Hermite-Gaussian eigenmodes $(H_n)_{n\geq 0}$ are crucial for the phenomenon of daydreaming in optical neurocomputers ([3]).

6. Scanout of pixel arrays. The implementation of pixel arrays by holographic optical interconnections ([2], [3], [18], [19], [20], [32], [34]), and analog VLSI wavefront arrays ([14], [15], [16], [1], [13]) suggests to look at restrictions of the holographic transform to lattices inside the hologram plane ([4]). The quadratic lattice $\mathbb{Z} \oplus \mathbb{Z}$ embedded in the hologram plane $\mathbb{R} \oplus \mathbb{R}$ may be considered as the projection onto G/C of the 3-cubic lattice $L_o := \{(\mu,\nu,\zeta) \mid \mu \in \mathbb{Z},\ \nu \in \mathbb{Z},\ \zeta \in \mathbb{Z}\}$ and the normal subgroup $L := \mathbb{Z} \oplus \mathbb{Z} \oplus C$

inside the three-dimensional Heisenberg nilpotent Lie group G along its center C. Then the compact Heisenberg nilmanifold $L_0\backslash G$ associated to G allows to realize by an application of the Weil-Brezin isomorphism

$$w_1 : \psi \mapsto ((x,y,z) \mapsto e^{2\pi i z} \sum_{n\in\mathbb{Z}} e^{2\pi i n y}\psi(n-x)) \qquad (\psi \in \mathscr{S}(\mathbb{R}))$$

the linear Schrödinger representation U_1 of G as the linear lattice representation $\delta_1 = \mathrm{Ind}_L^G(\chi_1)$ of G ([23]). It follows

$$H(\psi,\varphi;x,y) = \langle\delta_1(x,y,0)w_1(\psi)|w_1(\varphi)\rangle$$

for all points (x,y) of the pixel $]-1/2,+1/2] \times]-1/2,+1/2]$ in the hologram plane $\mathbb{R} \oplus \mathbb{R}$. Therefore the Parseval-Plancherel type pixel identity

$$\sum_{(\mu,\nu)\in\mathbb{Z}\oplus\mathbb{Z}} H(\psi;\mu,\nu)\overline{H}(\varphi;\mu,\nu) = \sum_{(\mu,\nu)\in\mathbb{Z}\oplus\mathbb{Z}} |H(\psi,\varphi;\mu,\nu)|^2$$

holds for writing amplitudes ψ,φ in $\mathscr{S}(\mathbb{R})$. If the Hermite functions H_m and H_n ($m\geq n\geq 0$) are inserted for ψ and φ, respectively, the radial symmetry of the terms of the left-hand side implies that the associated lattices of pixel arrays in the hologram plane have the crystallographic groups D_k ($k \in \{1,2,3,4,6\}$) of order 2k as their groups of symmetry ([26], [27], [34]). An application of the Weil-Brezin isomorphism w_1 to the readout formulae of the Corollary of Theorem 2 supra shows that the scanout of the pixel arrays of the holographic lattices may be performed by a time-multiplexing procedure.

Remark 6. The holographic lattices are at the basis of the detour phase method ([28], [30]) of writing digital CGHs of sampled images by use of the fast Fourier transform (FFT) algorithm. The height and the displacement of a single aperture centered at the sampling points of the holographic lattice are used to encode the amplitude and the phase of the complex wavefront. Thus the actual encoding of detour phase CGHs is performed without the explicit use of a reference beam. The holographic lattice corresponding to the crystallographic group D_6 of twelvefold symmetry offers substantial computational efficiency and a significant reduction of required data storage compared with rectangular sampling: the hexagonal FFT is 25% more efficient than the most efficient rectangular FFT algorithm. The scanout of the wavefront is achieved when the CGH is illuminated with a plane wave and focused with a Fourier-transforming lens.

Remark 7. The compact disks (CDs) may be regarded as one-dimensional digital CGHs that may be scanned out by the holographic optical head of a CD digital audio player. Another point of view is to consider CDs as one-dimensional artificial neural networks.

7. Artificial neural networks. In order to identify explicitly the terms of the Parseval-Plancherel type pixel identity indicated above, we denote by $K_{m,n}$ the complete bichromatic graph of $m + n$ vertices. Define $c(K_{m,n}, 0) := 1$ and let $c(K_{m,n}, l)$ denote the number of choices of $l \geq 1$ disjoint edges in $K_{m,n}$ each linking two vertices of different colours. Then

$$\Phi_{m,n}(X) := \sum_{0 \leq l \leq [(m+n)/2]} (-1)^l c(K_{m,n}, l) X^{m+n-2l}$$

denotes the matching polynomial ([8]) of variable X associated to the bipartite graph $K_{m,n}$.

Theorem 4. *The coefficients of the matching polynomial* $\Phi_{m,n}(X)$ *are the elementary synaptic weights* $(-1)^l c(K_{m,n}, l)$, $0 \leq l \leq [(m+n)/2]$, *where* $c(K_{m,n}, l)$ *denotes the number of disjoint synaptic interconnections of the neural network* $K_{m,n}$ $(m \geq n \geq 0)$ *activated by* l *simultaneously firing neurons.*

The next theorem describes the relationship between the elementary holograms and the matching polynomials attenuated by the Gaussian $(H_0 \otimes H_0) \in L^2(\mathbb{R} \oplus \mathbb{R})$ with distance: the farther away an input is from a point in the neural network, the less weight it is given.

Theorem 5. *Let* $m \geq n \geq 0$. *Then the elementary holograms admit the form*

$$H(H_m, H_n; x, y) = \frac{(-1)^n}{\sqrt{m!\,n!}} \; e^{-\pi(x^2+y^2)/2} \; \Phi_{m,n}(\sqrt{\pi}(x+iy))$$

for all pairs $(x, y) \in \mathbb{R} \oplus \mathbb{R}$.

Remark 8. In biological vision, the center-surround receptive field profiles of the retinal neurons ([5]) and the cells of the lateral geniculate nucleus are far from forming an orthogonal family in $L^2(\mathbb{R} \oplus \mathbb{R})$. Therefore the resulting neural representation remains highly correlated. Theorem 2 supra suggests to implement a matching filter bank by an adaptive artificial neural network model which is based on the central projection G-slice orbits

$$G_{(y,y')}: (x,x') \mapsto U_1(x,y,0) \otimes \bar{U}_1(x',y',0)(H_0 \otimes H_0) \qquad ((y,y') \in \mathbb{R} \oplus \mathbb{R})$$

in $L^2(\mathbb{R}\oplus\mathbb{R})$. The approximating family of Gabor functions $\{G_{(y,y')} \mid (y,y') \in \mathbb{R}\oplus\mathbb{R}\}$ is total in the complex Hilbert space $L^2(\mathbb{R} \oplus \mathbb{R})$ due to the irreducibility of the linear Schrödinger representation U_1 of G, but non-orthogonal. Early stages of biological visual systems pay for keeping $m = n = 0$ by the non-orthogonality of the center-surround receptive field profiles ([4], [12]). The retina and the lateral geniculate nucleus, however, act as decorrelators of the incoming signals. At the level of the mammalian visual cortex, the introduction of orientation selectivity through localized wave modulation combined with quadrature phase relations among paired cells results in a decorrelated neural representation with optimal image compression performance by the total orthogonal approximating family in $L^2(\mathbb{R} \oplus \mathbb{R})$ of elementary holograms $(H(H_m, H_n; .,.))_{m\geq 0, n\geq 0}$. Signal preprocessing and processing in the auditory parts of the cortex follow similar lines.

Theorem 5 supra implies

Theorem 6. *For* $m \geq n \geq 0$ *the identity*

$$\sum_{(\mu,\nu)\in\mathbb{Z}\oplus\mathbb{Z}} (-1)^{m+n} e^{-\pi(\mu^2+\nu^2)} \Phi_{m,m}(\sqrt{\pi}(\mu+i\nu))\Phi_{n,n}(\sqrt{\pi}(\mu+i\nu)) =$$

$$\sum_{(\mu,\nu)\in\mathbb{Z}\oplus\mathbb{Z}} e^{-\pi(\mu^2+\nu^2)} |\Phi_{m,n}(\sqrt{\pi}(\mu+i\nu))|^2$$

holds.

8. Theta-null values. The preceding theorem gives rise to the following special identities for the odd powers of π in terms of theta-null values $\vartheta(0,1) = \sum e^{-\pi\mu^2}$ ([23], [24]) where $\sum := \sum_{\mu\in\mathbb{Z}}$:

$m = 1,\ n = 0$

$$\pi = \frac{\sum e^{-\pi\mu^2}}{4 \sum \mu^2\, e^{-\pi\mu^2}}$$

$m = 2,\ n = 1$

$$\pi^3 = \frac{15 \sum (8\pi^2\mu^4 - 1)\, e^{-\pi\mu^2}}{32 \sum \mu^6\, e^{-\pi\mu^2}}$$

$m = 3,\ n = 2$

$$\pi^5 = \frac{45 \sum (16\pi^4\mu^8 - 140\pi^2\mu^4 + 21)\, e^{-\pi\mu^2}}{64 \sum \mu^{10}\, e^{-\pi\mu^2}}$$

$m = 4,\ n = 3$

$$\pi^7 = \frac{91 \sum (256\pi^6\mu^{12} - 15840\pi^4\mu^8 + 166320\pi^2\mu^4 - 25245)\, e^{-\pi\mu^2}}{1024 \sum \mu^{14}\, e^{-\pi\mu^2}}.$$

Theorem 5 supra shows that the preceding identities for the theta-null values $\vartheta(0,1)$ are of a combinatorial character.

Remark 9. The univariate impulse response of the ideal lowpass filter admits the Euler factorization

$$\operatorname{sinc} x = \prod_{n\geq 1} \left(1 - \frac{x^2}{n^2}\right) \qquad (x \in \mathbb{R}).$$

Its logarithmic derivative combined with the generating function of the Bernoulli polynomials $B_n(X)$ of degree $n \geq 0$ yields the classical Euler formulae for the even powers of π:

$$\pi^{2n} = (-1)^{n+1} \frac{2(2n)!}{2^{2n} B_{2n}}\, \zeta(2n) \qquad (n \geq 1),$$

where ζ denotes the Riemann zeta-function and $B_{2n} = B_{2n}(0)$ are the Bernoulli numbers.

Acknowledgments. The author acknowledges the support of the Visiting International Scholar Award 1988/89 from the University of Missouri-Saint Louis. It is his pleasure to thank Professors Gail D.L. Ratcliff and Grant V. Welland (St. Louis) for interesting discussions of these topics.

REFERENCES

1. **T. Allen, C. Mead, F. Faggin, and G. Gribble,** Orientation-selective VLSI retina, Visual Communications and Image Processing '88, T. Russell Hsing, Editor, Proc. SPIE 1001, 1040-1046 (1988).

2. **D.Z. Anderson,** Coherent optical eigenstate memory, Optics Letters 11, 56-58 (1986).

3. **D.Z. Anderson, M.C. Erie,** Resonator memories and optical novelty filters, Optical Engineering 26, 434-444 (1987).

4. **J.G. Daugman,** Relaxation neural network for complete discrete 2-D Gabor transforms, Visual Communications and Image Processing '88, T. Russell Hsing, Editor, Proc. SPIE 1001, 1048-1061 (1988).

5. **J.E. Dowling,** The retina: an approachable part of the brain, The Belknap Press of Harvard University Press, Cambridge, Massachusetts, and London 1987.

6. **N.H. Farhat, C.L. Werner, and T.H. Chu,** Prospects for three-dimensional projective and tomographic imaging radar networks, Radio Science 19, 1347-1355 (1984).

7. **E. Feig, C.A. Micchelli,** L^2-synthesis by ambiguity functions, Multivariate Approximation Theory IV, C.K. Chui, W. Schempp, and K. Zeller, Editors, Birkhäuser Verlag, Basel, Boston, Berlin 1989.

8. **J. Feinberg,** Applications of real-time holography, Holography, Lloyd Huff, Editor, Proc. SPIE 532, 119-135 (1985).

9. **A. Grossmann, J. Morlet,** Decomposition of functions into wavelets of constant shape and related transforms, Mathematics and Physics, Lectures on Recent Results, Vol. 1, L. Streit, Editor, World Scientific, Singapore, Philadelphia 1985.

10. **H. Hosoya,** Matching and symmetry of graphs, Comp. and Maths. with Appls. 12B, 271-290 (1986).

11. **R. Howe,** Dual pairs in physics: Harmonic oscillators, photons, electrons, and singletons, Applications of Group Theory in Physics and Mathematical Physics, M. Flato, P. Sally, and G. Zuckerman, Editors, 179-207, American Mathematical Society, Providence, Rhode Island 1985.

12. **J. Jones, L. Palmer,** An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex, J. of Neurophysiology 58, 1233-1258 (1987).

13. **J. Lazzaro, C.A. Mead,** A silicon model of auditory localization, Neural Computation 1, 47-57 (1989).

14. **C. Mead,** Analog VLSI and neural systems, Addison-Wesley, Reading, Massachusetts 1989.

15. **C. Mead, M. Ismail,** Analog VLSI implementation of neural systems, Kluwer, Norwell, Massachusetts 1989.

16. **C.A. Mead, M.A. Mahowald,** A silicon model of early visual processing, Neural Networks 1, 91-97 (1988).

17. **H. Moscovici,** Coherent state representations of nilpotent Lie groups, Commun. math. Phys. 54, 63-68 (1977).

18. **J. Ohta, M. Takahashi, Y. Nitta, S. Tai, K. Mitsunaga, and K. Kjuma,** A new approach to a GaAs/AlGaAs optical neurochip with three layered structure, Proc. IJCNN International Joint Conference on Neural Networks, II-477-480 (1989).

19. **Y. Owechko, E. Marom, B.H. Soffer, and G. Dunning,** Associative memory in a phase conjugate resonator cavity utilizing a hologram, IOCC-1986 International Optical Computing Conference, J. Shamir, A.A. Friesem, and E. Marom, Editors, Proc. SPIE 700, 296-7300 (1986).

20. **D. Psaltis, D. Brady, X. Gu, and K. Hsu,** Optical implementation of neural computers, Optical Processing and Computing, H. Arsenault, T. Szoplik, and B. Macukow, Editors, Academic Press, Boston, Orlando, San Diego, New York, Austin, London, Sydney, Tokyo, Toronto 1989.

21. **G. Ries,** Rotationssymmetrische Radar-Ambiguity-Funktion, Manuskript, Lehrstuhl für Elektrotechnik VIII - Hochfrequenztechnik, Universität Siegen 1989.

22. **W. Schempp,** Radar ambiguity functions, the Heisenberg group, and holomorphic theta series, Proc. Amer. Math. Soc. 92, 103-110 (1984).

23. **W. Schempp,** Harmonic analysis on the Heisenberg nilpotent Lie group, with applications to signal theory, Pitman Research Notes in Math., Vol. 147, Longman Scientific and Technical, Harlow, Essex, and J. Wiley & Sons, New York 1986.

24. **W. Schempp,** Group theoretical methods in approximation theory, elementary number theory, and computational signal geometry, Approximation Theory V, C.K. Chui, L.L. Schumaker, and J.D. Ward, Editors, 129-171, Academic Press, Boston, Orlando, San Diego, New York, Austin, London, Sydney, Tokyo, Toronto 1986.

25. **W. Schempp**, The oscillator representation of the metaplectic group applied to quantum electronics and computerized tomography, Stochastic Processes in Physics and Engineering, S. Albeverio, P. Blanchard, M. Hazewinkel, and L. Streit, Editors, 305-344, D. Reidel, Dordrecht, Boston, Lancaster, Tokyo 1988.

26. **W. Schempp**, Elementary holograms and 3-orbifolds, C.R. Math. Rep. Acad. Sci. Canada 10, 155-160 (1988).

27. **W. Schempp**, Holographic grids, Visual Communications and Image Processing '88, T. Russell Hsing, Editor, Proc. SPIE 1001, 116-120 (1988).

28. **W. Schempp**, The holographic transform, Numerical Methods and Approximation Theory III, G.V. Milovanovič, Editor, 67-91, University of Niš, Niš 1988.

29. **W. Schempp**, Holographic image processing, coherent optical computing, and neural computer architecture for pattern recognition, Lie Methods in Optics II, K.B. Wolf, Editor, Lecture Notes in Physics, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo 1989.

30. **D. Schreier**, Synthetische Holografie, Fachbuchverlag Leipzig 1984.

31. **D.S. Shucker**, Square integrable representations of unimodular groups, Proc. Amer. Math. Soc. 89, 169-172 (1983).

32. **B.H. Soffer, G.J. Dunning, Y. Owechko, and E. Marom**, Associative holographic memory with feedback using phase-conjugate mirrors, Optics Letters 11, 118-120 (1986).

33. **H.J. Tiziani**, Real-time metrology with BSO crystals, Optica Acta 29, 463-470 (1982).

34. **T. Yatagai**, Cellular logic architectures for optical computers, Applied Optics 25, 1571-1577 (1986).

eingegangen: 25. 7. 1989

Verfasser:

Prof. Dr. Walter Schempp
Lehrstuhl für Mathematik I
Universität Siegen
Siegen
D-5900

Rostock. Math. Kolloq. 40, 51 - 57 (1990) 45G10
47H99

Dieter Schott; Maria Kapitanowa;
Georgi Kostadinow; Stepan Kostadinow

Über einen abstrakten Volterraoperator

Das Ziel der vorliegenden Arbeit besteht in der Untersuchung der Eigenschaften eines nichtlinearen Operators vom Volterratyp und der Eigenschaften der entsprechenden inhomogenen Gleichung zweiter Art. Wir stützen uns dabei auf Resultate aus [3] und [4] und verallgemeinern Aussagen aus [1] und [2].
In [1] wird ein nichtlinearer beschränkter Operator K vom Volterratyp vorausgesetzt, der über dem Raum $C(\Omega)$ aller komplexwertigen stetigen Funktionen f mit einem kompakten metrischen Definitionsbereich Ω erklärt ist. Man findet dort hinreichende Bedingungen für die Existenz und Eindeutigkeit der Lösung von $f + Kf = 0$. Außerdem wird ein Analogon der Bellman-Gronwall-Ungleichung bewiesen.
In [2] werden die Voraussetzungen in der Weise abgeschwächt, daß die Werte der Funktionen f in einem beliebigen Banachraum liegen können und ihr Definitionsbereich Ω nicht mehr kompakt zu sein braucht. Es werden Eigenschaften der homogenen Gleichung $f + Kf = 0$ und der inhomogenen Gleichung $f + Kf = p$ angegeben.

Daran anknüpfend setzen wir folgende allgemeinere Situation voraus:

Es sei Ω ein topologischer Raum mit Borelschem Maß μ. Weiterhin sei Φ: $x \to M_x$ eine Abbildung, die jedem Element x aus Ω eine kompakte Teilmenge M_x von Ω zuordnet. Wir werden sagen, daß die Bedingung (A) erfüllt ist, wenn folgende Voraussetzungen gelten:

A1: Es ist $\omega = \sup\limits_{x\in\Omega} \mu(M_x) < \infty$.

A2: Für jedes $\varepsilon > 0$ und jedes $x \in \Omega$ existiert eine Umgebung $V = V(x,\varepsilon)$ von x, so daß für jedes $y \in V$ die Ungleichung
$\mu(M_x \Delta M_y) = \mu(\{M_x \setminus M_y\} \cup \{M_y \setminus M_x\}) < \varepsilon$
zutrifft.

Es sei B ein reeller Banachraum mit der Norm $\|.\|_B$. Wir betrachten folgende linearen Räume:

$$C^{-1} = C^{-1}(\Omega, B) = \{ f \mid f: \Omega \to B, \sup_{x\in\Omega} \|f(x)\|_B < \infty \},$$
$$C = C(\Omega, B) = \{ f \mid f: \Omega \to B \text{ stetig}, \sup_{x\in\Omega} \|f(x)\|_B < \infty \}.$$

Der Raum C^{-1} ist ein Banachraum bezüglich der Norm

$$\|f\| = \sup \{ \|f(x)\|_B : x \in \Omega \}. \tag{1}$$

Wenn Ω kompakt ist, so ist der Raum C ebenfalls ein Banachraum bezüglich der Norm (1).
Wir werden sagen, daß der Operator

$$F: \Omega \times C^{-1} \to C^{-1}$$

die <u>Bedingung B</u> erfüllt, wenn folgende Voraussetzungen gelten:

<u>B1</u>: Für jedes $\varepsilon > 0$ und jede Funktion $f' \in C^{-1}$ existiert eine Zahl $\delta = \delta(\varepsilon, f') > 0$, so daß für jede Funktion $f \in C^{-1}$ mit $\|f - f'\| < \delta$ und jedes $x \in \Omega$ die Ungleichung

$$\|F(x, f) - F(x, f')\| < \varepsilon$$

zutrifft.

<u>B2</u>: Für jedes $\varepsilon > 0$ und jedes $x' \in \Omega$ existiert eine Umgebung $U = U(x', \varepsilon)$ von x', so daß für jedes $x \in U$ und jede Funktion $f \in C^{-1}$ die Ungleichung

$$\|F(x, f) - F(x', f)\| < \varepsilon$$

zutrifft.

<u>B3</u>: Für jedes feste Element $x \in \Omega$ und jede feste Funktion $f \in C^{-1}$ ist die Abbildung

$$\varphi_f(y) = F(x, f)(y)$$

schwach meßbar bezüglich M_x $(x \in \Omega)$ und fast separabelwertig, d.h., zu jeder Menge M_x $(x \in \Omega)$ existiert eine Menge $E_x \subset M_x$ mit $\mu(E_x) = 0$, so daß $\varphi_f(M_x \setminus E_x)$ eine separable Teilmenge des Raumes B ist.

<u>B4</u>: Für jedes $x \in \Omega$ und jede Funktion $f \in C^{-1}$ besteht die Ungleichung

$$\|F(x,\ f)\| \leq A(f)\|f\| \quad (A\colon C^{-1} \to [0,\ \infty)).$$

Wir betrachten nun die Gleichung

$$f + Kf = p \qquad (f,\ p \in C^{-1}), \tag{2}$$

wobei der Operator K folgendermaßen definiert ist:

$$(Kf)(x) = \int_{M_x} F(x,\ f)(y)d\mu_y \qquad (x \in \Omega). \tag{3}$$

Die Existenz des Integrals in (3) ist nach den Sätzen 3.7.4 und 3.5.3 aus [3] durch die Voraussetzungen B3 und B4 gesichert. Wir geben zunächst Abbildungseigenschaften des Operators K an.

<u>Lemma 1:</u> Unter den Voraussetzungen A1, B3 und B4 bildet der Operator K den Raum C^{-1} in sich ab.

<u>Beweis:</u> Für beliebiges $f \in C^{-1}$ folgt aus A1 und B4 die Abschätzung

$$\|(Kf)(x)\| \leq A(f)\|f\|\omega.$$

Damit liegt auch Kf in C^{-1}.

<u>Lemma 2:</u> Unter den Voraussetzungen A1, B1, B3 und B4 bildet der Operator K den Raum C^{-1} stetig in sich ab.
Sind außerdem noch die Voraussetzungen A2 und B2 erfüllt, so daß die Bedingungen (A) und (B) gelten, dann bildet der Operator K den Raum C^{-1} stetig in den Raum C ab.

<u>Beweis:</u> Es sei (f_n) eine beliebige Folge aus C^{-1} mit

$$\lim_{n\to\infty} f_n = f' \in C^{-1}.$$

Weiterhin sei $\varepsilon > 0$ beliebig vorgegeben. Aus B1 folgt dann die Existenz einer natürlichen Zahl $n_o = n_o(\varepsilon,\ f')$, so daß für $n > n_o$ die Ungleichung

$$\|F(x,\ f_n) - F(x,\ f')\| \leq \varepsilon/\omega \qquad (x \in \Omega) \tag{4}$$

gilt. Aus (4) und A1 erhält man unmittelbar

$$\|Kf_n - Kf'\| \leq \varepsilon.$$

Somit ist die erste Behauptung des Lemmas nachgewiesen.
Wir betrachten nun ein beliebiges f aus C^{-1} und zeigen, daß Kf stetig ist. Dazu seien $\varepsilon > 0$ und $x' \in \Omega$ beliebig vorgegeben. Aus A2 und B2 ergibt sich die Existenz einer Umgebung $W = W(\varepsilon, x') = U(x', \varepsilon) \cap V(x', \varepsilon)$ von x', so daß für jedes $x \in W$ die Ungleichungen

$$\|F(x, f) - F(x', f)\| < \frac{\varepsilon}{3\omega}, \quad \mu(M_x \Delta M_{x'}) < \frac{\varepsilon}{3A(f)\|f\|} \tag{5}$$

bestehen. Für beliebiges, aber festes $x \in W$ folgt dann aus den Beziehungen (5)

$$\begin{aligned}
&\|(Kf)(x) - (Kf)(x')\|_B \\
&= \|\int_{M_x} F(x,f)(y)d\mu_y - \int_{M_{x'}} F(x',f)(y)d\mu_y\|_B \\
&\leq \|\int_{M_x \cap M_{x'}} [F(x,f)(y) - F(x',f)(y)]d\mu_y\|_B \\
&+ \|\int_{M_x \setminus M_{x'}} F(x,f)(y)d\mu_y\|_B + \|\int_{M_{x'} \setminus M_x} F(x',f)(y)d\mu_y\|_B \\
&< \frac{\varepsilon\omega}{3\omega} + 2A\|f\| \frac{\varepsilon\omega}{3A\|f\|\omega} = \varepsilon .
\end{aligned}$$

Lemma 3: Es seien die Bedingungen (A) und (B) erfüllt, wobei das Funktional A in B4 außerdem auf jeder beschränkten Menge von C^{-1} als beschränkt vorausgesetzt werden soll. Weiterhin sei $B(O,R)$ eine beliebige abgeschlossene Kugel des Raumes C^{-1} mit dem Mittelpunkt O und dem Radius R.
Dann ist die Menge $K(B(O,R))$ gleichmäßig beschränkt und gleichgradig stetig.

Wir verzichten hier auf den Beweis, weil er ähnlich verläuft, wie der von Lemma 2.

Für die weiteren Betrachtungen wird der folgende Satz aus [4] benötigt.

Satz 1: Es sei D eine offene beschränkte Teilmenge des reellen Banachraumes E. Der Operator $K\colon \bar{D} \to E$ sei kompakt und stetig. Dann hat die Gleichung

$$f + Kf = p$$

für jedes $p \in D$ mit der Eigenschaft, daß $f + tKf \neq p$ für alle $f \in \partial D$ und $t \in [0,1]$ gilt, eine Lösung in D.

Außerdem brauchen wir noch zwei Begriffe (siehe auch [2]). Es sei $p \in C^{-1}$, $p \neq 0$, und $f \in C^{-1}$ genüge der Bedingung $\|p - f\| < \|p\|$.

Definition 1: Wir sagen, daß f in bezug auf den Operator K normal zur Kugel $B(p,R_f)$ mit $R_f \in (\|p - f\|, \|p\|)$ liegt, wenn für jedes $g \in \partial B(p,R_f)$ ein $\bar{x} = \bar{x}(f,g) \in \Omega$ und eine Zahl $\bar{\alpha} = \bar{\alpha}(f, g) \in (0, 1/\mu(M_{\bar{x}}))$ existieren, so daß mit den Abkürzungen

$$u(\bar{x},h) = \sup_{y \in M_{\bar{x}}} \|F(\bar{x},h)(y)\|_B,$$

$$u(\bar{x},f,g) = \sup_{y \in M_{\bar{x}}} \|F(\bar{x},f)(y) - F(\bar{x},g)(y)\|_B$$

die Ungleichung

$$\max \{ u(\bar{x},f),\ u(\bar{x},g),\ u(\bar{x},f,g)\} < \bar{\alpha}\|f(\bar{x}) - g(\bar{x})\|_B \tag{6}$$

besteht.

Definition 2: Die Funktion $p \in C^{-1}(\Omega, B)$, $p \neq 0$, heißt reguläres Element der Gleichung (2), wenn ein Radius $R' \in (0, \|p\|)$ und eine Zahl $\varepsilon \in (0, \|p\| - R')$ existieren, so daß jedes $f \in \partial B(p, R')$ in bezug auf den Operator K normal zu einer Kugel $B(p, R_f)$ mit $R' + \varepsilon < R_f < \|p\|$ liegt.

Bemerkung 1: Beispiele für reguläre Elemente sind in [2] zu finden.

Wir kommen jetzt zur Formulierung des Hauptsatzes.

Satz 2: Es seien die Bedingungen (A) und (B) erfüllt. Weiterhin sei $p \in C^{-1}$ ein reguläres Element der Gleichung (2), für das $K(B(p, \|p\|))$ eine relativ kompakte Teilmenge von C^{-1} ist. Dann besitzt (2) eine Lösung $f \neq 0$ in der Kugel $B(p, \|p\|)$.

Beweis: Aus B4 folgt zunächst, daß $f = 0$ keine Lösung von (2) ist. Nun sei p ein Element mit den im Satz angegebenen Eigenschaften und $\{ B(p, R) : R \in (0, \|p\|) \}$ eine Familie abgeschlossener Kugeln.

Auf $\bar{D} = B(p, \|p\|)$ ist K ein kompakter, stetiger Operator. Wir nehmen an, daß für jede Kugel $B(p, R)$ eine Zahl $t_R \in [0, 1]$ und ein Element $f_R \in \partial B(p, R)$ mit

$$f_R + t_R K f_R = p \tag{7}$$

existieren. Andernfalls wäre Satz 2 aufgrund von Satz 1 schon bewiesen. Aus (7) folgt, daß $t_R \neq 0$ für alle R aus $(0, \|p\|)$ gilt. Darüber hinaus können wir auch voraussetzen, daß $t_R \neq 1$ für alle R aus $(0, \|p\|)$ zutrifft, da sonst die Behauptung schon gezeigt wäre.

Es seien $R' \in (0, \|p\|)$ und $\varepsilon \in (0, \|p\| - R')$ die in Definition 2 genannten Zahlen. Die Funktion $f' \in \partial B(p, R')$ und die Zahl $t' \in (0,1)$ seien so gewählt, daß sie der Beziehung

$$f' + t' K f' = p \tag{7'}$$

genügen. Aus der Regularität von p erhält man die Existenz einer Kugel $B(p, R_{f'})$ mit dem Radius $R_{f'} \in (R'+\varepsilon, \|p\|)$, zu der f' in bezug auf den Operator K normal liegt.

Daher gibt es unter Beachtung der Annahmen und der Definition 1 ein $g^* \in \partial B(p, R_{f'})$ und ein $t^* \in (0,1)$, so daß einerseits

$$g^* + t^* K g^* = p \tag{7"}$$

gilt und andererseits mit einem $\bar{x} = \bar{x}(f', g^*) \in \Omega$ und einer Zahl $\bar{\alpha} = \bar{\alpha}(f', g^*) \in (0, 1/\mu(M_{\bar{x}}))$ die Abschätzung

$$\begin{aligned} &\sup_{y \in M_{\bar{x}}} \|t' F(\bar{x}, f')(y) - t^* F(\bar{x}, g^*)(y)\|_B \\ &\leq \max \{ u(\bar{x}, f'), u(\bar{x}, g^*), u(\bar{x}, f', g^*) \} \\ &\leq \bar{\alpha} \, \|f'(\bar{x}) - g^*(\bar{x})\|_B \end{aligned} \tag{8}$$

folgt. Aus (7'), (7") und (8) erhält man aber den Widerspruch

$$\begin{aligned} &\|f'(\bar{x}) - g^*(\bar{x})\|_B = \|t'(Kf')(\bar{x}) - t^*(Kg^*)(\bar{x})\|_B \\ &\leq \mu(M_{\bar{x}}) \sup_{y \in M_{\bar{x}}} \|t' F(\bar{x}, f')(y) - t^* F(\bar{x}, g^*)(y)\|_B \\ &< \|f'(\bar{x}) - g^*(\bar{x})\|_B . \end{aligned}$$

Damit ist alles gezeigt.

Bemerkung 2: Ein ähnlicher Satz ergibt sich auch für den Raum $C(\Omega, B)$.

Literatur:

1. Bainov, D.D., Myshkis, A.D., and Zahariev, A.I.: On an abstract analogue of the Bellman-Gronwall inequality. Publ. Res. Inst. Math. Sci. 20, 903-911 (1984)

2. Bainov, D.D., Kostadinov, S.I., and Zahariev, A.I.: Abstract Volterra type integral equations. Bull. Inst. Math. Acad. Sinica 16, 93-104 (1988)

3. Hille, E., and Phillips, R.: Functional Analysis and Semigroups. Providence 1957

4. Hutson, V.C.L., and Pym, J.S.: Applications of Functional Analysis and Operator Theory. New York 1980

eingegangen: 22. 11. 1988 revidierte Fassung: 3. 2. 1989

Verfasser:

Doz. Dr. Dieter Schott
Pädagogische Hochschule "Liselotte Herrmann"
Sektion Mathematik/Physik
Goldberger Straße 12
Güstrow
DDR-2600

Dr. Maria Kapitanowa
Dr. Georgi Kapitanow
Dr. Stepan Kostadinow
Universität "Paissii Hilendarski"
Wissenschaftsbereich
"Angewandte Mathematik und Wahrscheinlichkeitsrechnung"
Zar Asen 24
Plovdiv
VR Bulgarien-4000

Rostock. Math. Kolloq. 40, 58 - 70 (1990) 47H10
54H25

S.N. Mishra; S.L. Singh

Some results on coincidences and fixed points

0. Introduction

Let (X,u) be a uniform space. A family $\{d_t\}$, t running over an indexing set I, of pseudometrics on X is called an associated family for the uniformity u if the family $B' = \{ V(t,r): t \in I, r > 0 \}$, where $V(t,r) = \{ (x,y): x,y \in X, d_t(x,y) < r \}$ is a subbase for the uniformity u. We may assume B' itself to be a base by adjoining finite intersections of members of B'. The corresponding family P^* of pseudometrics is called an augmented associated family for u (cf. Thron [13]).

For any nonempty subset A of X we define $\delta(A) = \sup\{ d_t(x,y): x,y \in A, t \in I \}$, where $\{ d_t: t \in I \} = P^*$. Then $\delta(A)$ is called an augmented diameter of A, and if $\delta(A) < \infty$, then A is called P^*-bounded (cf. Mishra [8]). For any nonempty subsets A and B of X we define

$$D_t(x,A) = \inf\{ d_t(x,A): a \in A, t \in I \}, (x \in X),$$

$$\delta_t(A,B) = \sup\{ d_t(a,b): a \in A, b \in B, t \in I \},$$

$$H_t(A,B) = \max\{ \sup D_t(a,B): a \in A, \sup D_t(A,b): b \in B \}.$$

Further, let

$$BN(X) = \{ A: A \text{ is a nonempty } P^*\text{-bounded subset of } X \},$$

$$B(X) = \{ A: A \text{ is a nonempty subset of } X \},$$

$$CL(X) = \{ A: A \text{ is a nonempty closed subset of } X \},$$

$$2^X = BN(X) \cap CL(X).$$

It is well-known that H_t is a pseudometric on 2^X, called the Hausdorff pseudometric induced by d_t, $t \in I$.

Let $U \in u$ be arbitrary. For each subset A of X, define

$U(A) = \{ y: (x,y) \in U$ for some $x \in X \}$. The Hausdorff uniformity 2^u on 2^X is defined by the base $2^{B'} = \{ \tilde{U}: U \in u \}$, where $\tilde{U} = \{ (A,B): A,B \in 2^X$ and $A \subseteq U(B),\ B \subseteq U(A) \}$. The augmented associated family P^* also induces a uniformity u^* on 2^X defined by the base $B^* = \{ V^*(t,r): t \in I,\ r > 0 \}$ where $V^*(t,r) = \{ (A,B): A,B \in 2^X,\ H_t(A,B) < r \}$. It is also well-known that the uniformities u^* and 2^u on 2^X are uniformly isomorphic. Thus 2^X equipped with u^* or 2^u is a uniform space, called the hyperspace of X.
In the sequel, unless stated otherwise, the terms uniform space and P^*-bounded will refer to Hausdorff uniform space defined by P^* and the P^*-boundedness, respectively. Further, by completeness of the space X or of its subspace we shall mean the sequential completeness of the same.

Definition 1: Let X be an uniform space, and let $h: X \to X$, $F,G: X \to 2^X$ be singlevalued and multivalued mappings. F and h (resp. G and h) are said to commute if for each $x \in X$ we have $Fhx = hFx$ (resp. $Ghx = hGx$). Further, F and G are said to commute if for each $x \in X$, we have $FGx = GFx$. As a consequence, we have $FGA = GFA$, where $FA = \cup \{ Fa: a \in A \}$, A being a nonempty subset of X.

Definition 2: Let X be an uniform space, and let $h: X \to X$, $F,G: X \to 2^X$. Mappings F,G and h are said to have a coincidence point z (resp. a common fixed point z) if $hz \in Fz \cap Gz$ (resp. $hz = z \in Fz \cap Gz$).

1. Results

Theorem 1: Let (X,u) be a uniform space. Suppose $F,G: X \to 2^X$ and $h: X \to X$ are mappings such that $F(X) \cup G(X) \subseteq h(X)$ and for each $t \in I$, there exists number $q_t \in (0,1)$ such that

$$H_t(Fx,Gy) \leq q_t \max\{ d_t(hx,hy),\ D_t(hx,Fx),\ D_t(hy,Gy),\ \frac{1}{2}[D_t(hx,Gy) + D_t(hy,Fx)] \} \tag{1.1}$$

for all $x,y \in X$. If $h(X)$ is a complete subspace of X and h is commuting with F and G, then F, G and h have a coincidence point.

Proof: Let $U \in u$ be arbitrary. Since B' is a base for u, there exists $V(t,r) \in B'$ such that $V(t,r) \subseteq U$. Pick $x_o \in X$. Since $F(X) \cup G(X) \subseteq h(X)$, we may choose a point $hx_1 \in Fx_o$ and a point $hx_2 \in Gx_1$ such that

$$d_t(hx_1, hx_2) \leq q_t^{-a_t} H_t(Fx_o, Gx_1) \quad \text{for some} \quad a_t \in (0,1).$$

In general, we can choose a sequence $\{x_n\}$ in X such that

$$hx_{2n+1} \in Fx_{2n},\ hx_{2n+2} \in Gx_{2n+1}$$

and

$$d_t(hx_{2n+1}, hx_{2n+2}) \leq q_t^{-a_t} H_t(Fx_{2n}, Gx_{2n+1}),$$

$$d_t(hx_{2n+2}, hx_{2n+3}) \leq q_t^{-a_t} H_t(Gx_{2n+1}, Fx_{2n+2}).$$

Consider the inequality,

$$d_t(hx_{2n+1}, hx_{2n+2}) \leq q_t^{-a_t} H_t(Fx_{2n}, Gx_{2n+1}).$$

Using condition (1.1), it can be easily verified that

$$d_t(hx_{2n+1}, hx_{2n+2}) \leq q_t^{1-a_t} d_t(hx_{2n}, hx_{2n+1}).$$

Similarly,

$$d_t(hx_{2n+2}, hx_{2n+3}) \leq q_t^{1-a_t} d_t(hx_{2n+1}, hx_{2n+2}).$$

Therefore for all $n \geq N(t,r)$, we have

$$d_t(hx_n, hx_{n+1}) \leq q_t^{1-a_t} d_t(hx_{n-1}, hx_n).$$

Since $0 < q_t^{1-a_t} < 1$, it follows that $d_t(hx_n, hx_{n+1}) < r$ and $(hx_n, hx_{n+1}) \in U$ for all $n \geq N(t,r)$. Therefore $\{hx_n\}$ is a Cauchy sequence and $hx_n \to b$ for some $b \in h(X)$. Hence there exists a z in X such that $hz = b$.

We note that $hx_{2n+1} \in Fx_{2n}$. Therefore $hhx_{2n+1} \in hFx_{2n} \subseteq Fhx_{2n}$.

Let $V \in u$ be arbitrary. Suppose $V(i,s) \in B'$ such that $V(i,s) \subseteq V$. Consider

$$D_i(hz,Gz) \leq d_i(hz,hx_{2n+1}) + D_i(hx_{2n+1},Gz)$$

$$\leq d_i(hz,hx_{2n+1}) + H_i(hx_{2n+1},Gz)$$

$$\leq d_i(hz,hx_{2n+1}) + q_i \max\{ d_i(hx_{2n},hz),$$

$$d_i(hx_{2n},hx_{2n+1}),\ D_i(hz,Gz),$$

$$\tfrac{1}{2}\, d_i(hx_{2n},hz) + D_i(hz,Gz) + d_i(hz,hx_{2n+1}) \},$$

wich follows from condition (1.1) and simple calculation. Since $hx_n \to hz$, it follows that $D_i(hz, Gz) \leq q_i D_i(hz,Gz)$. Consequently, $D_i(hz,Gz) = 0 < s$ for all $n \geq N(i,s)$. Thus $(hz,Gz) \in V$ and since V is arbitrary and X is Hausdorff, we have $hz \in Gz$. Similarly $hz \in Fz$. Hence $hz \in Fz \cap Gz$. ■

Remark 1: It is interesting to note that mappings F,G, and h satisfying the condition

$$H_t(Fx,Gy) \leq a_t\, d_t(hx,hy) + b_t\, D_t(hx,Fx) + c_t\, D_t(hy,Gy) +$$

$$e_t[D_t(hx,Gy) + D_t(hy,Fx)]$$

for all $x,y \in X$, a_t, b_t, c_t, e_t being nonnegative real numbers with $0 < a_t + b_t + c_t + 2e_t < 1$, will also satisfy the condition (1.1). Therefore several fixed point theorems for multivalued mappings in metric spaces due to Iseki [5], Kasahara [6], Nadler [10] and Reich [11] may be obtained as special cases of Theorem 1 upon choosing h to be an identity mapping. Also, if $F = G$ and h is an identity mapping then the completeness in the above theorem may be replaced by F-orbital completeness (cf. Ciric [1]). Theorem 1 remains valid if $H_t(A,B)$ is defined for all $A,B \in CL(X)$. Hence Circic's result ([1], Theorem 2) is also obtained as a special case of the above theorem upon choosing $F = G$ and h an identity mapping on a metric space X. The following corollaries are also easy consequences of Theorem 1.

Corollary 1: ([9], Corollary 1). Let X be a complete uniform space and let (F,G) be a pair of multivalued mappings from X to 2^X satisfying condition (1.1) with h being an identity mapping on X. Then F and G have a common fixed point in X.

Corollary 2: Let (X,u) be an uniform space, and let $f, g, h: X \to X$ be mappings such that $f(X) \cup g(X) \subseteq h(X)$ and for each $t \in I$, there exists nonnegative real numbers a_t, b_t, c_t, e_t with $0 < a_t + b_t + c_t + 2e_t < 1$ such that

$$d_t(fx,gx) \leq a_t\, d_t(hx,hy) + b_t\, d_t(hx,fx) + c_t\, d_t(hy,gy)$$

$$+ e_t[d_t(hx,gy) + d_t(hy,gx)]$$

for all $x,y \in X$. If $h(X)$ is a complete subspace of X and h is commuting with f and g, then f, g and h have a unique common fixed point.

Remark 2: Corollary 1 generalizes a result of Kubiak [7] which itself includes and improves many results including those of Singh and Whitfield [12].

Theorem 2: Let (X,u) be a uniform space, and let $F, G: X \to BN(X)$ and $h: X \to X$ be mappings such that $F(X) \cup G(X) \subseteq h(X)$. Suppose for each $t \in I$, there exists a number $q_t \in (0,1)$ such that

$$\delta_t(Fx,Gx) \leq q_t \max\{ d_t(hx,hy),\ H_t(hx,Fx),\ H_t(Hy,Gy),\ \frac{1}{4}[H_t(hx,Gy) + H_t(hy,Fx)] \} \tag{2.1}$$

for all $x,y \in X$. If $h(X)$ is a complete subspace of X and h is commuting with F and G, then F, G and h have a unique common fixed point.

Proof: Let $U \in u$ be arbitrary. Since B' is a base for u, there exists $V(t,r) \in B'$ such that $V(t,r) \subseteq U$. Define singlevalued mappings $f, g: X \to X$ such that $fx \in Fx$ and $gx \in Gx$ for all $x \in X$ and $d_t(hx,fx) \geq q_t^{a_t} H_t(hx,Fx)$, $d_t(hx,gx) \geq q_t^{b_t} H_t(hx,Gx)$ for some a_t and b_t in $(0,1)$. Note that $q_t^{a_t}, q_t^{b_t} > q_t$.

We have

$$d_t(fx,gy) \leqslant \delta_t(Fx,Gy)$$

$$\leqslant q_t \max\{ d_t(hx,hy),\ H_t(hx,Fx),\ H_t(hy,Gy),$$

$$\tfrac{1}{4}\,[H_t(hx,Gy) + H_t(hy,Fx)] \}.$$

Therefore one of the following relations holds:

$$d_t(fx,gy) \leqslant q_t\, d_t(hx,hy) \tag{2.2}$$

$$d_t(fx,gy) \leqslant q_t H_t(hx,Fx) \leqslant q_t^{1-a_t}\, d_t(hx,fx) \tag{2.3}$$

$$d_t(fx,gy) \leqslant q_t H_t(hy,Gy) \leqslant q_t^{1-b_t}\, d_t(hy,gy) \tag{2.4}$$

$$\begin{aligned} d_t(fx,gy) &\leqslant \tfrac{1}{4}\, q_t[H_t(hx,Gy) + H_t(hy,Fx)] \\ &\leqslant \tfrac{1}{4}\, q_t[q_t^{-b_t}(d_t(hx,hy) + d_t(hy,gy)) \\ &\quad + q_t^{-a_t}(d_t(hy,hx) + d_t(hx,fx)] \\ &\leqslant \tfrac{1}{4}\,(q_t^{1-a_t} + q_t^{1-b_t})d_t(hx,hy) + \tfrac{1}{4}q_t^{1-a_t}\, d_t(hx,fx) \\ &\quad + \tfrac{1}{4}\, q_t^{1-b_t}\, d_t(hy,gy) \end{aligned} \tag{2.5}$$

Since $f(X) \cup g(X) \subseteq F(X) \cup G(X) \subseteq h(X)$ and

$$0 < q_t,\ q_t^{1-a_t},\ q_t^{1-b_t} < 1,$$

$$0 < \tfrac{1}{4}(q_t^{1-a_t} + q_t^{1-b_t}) + \tfrac{1}{4}(q_t^{1-a_t} + q_t^{1-b_t}) < 1,$$

it follows from Corollary 2 that in each of the cases (2.2) - (2.5) f, g, and h have a unique common fixed point, say z. Then $0 = d_t(hz,fz) \geqslant q_t^{a_t} H_t(hz,Fz)$. Hence $H_t(hz,Fz) < r$ and consequently, $(hz,Fz) \in V(t,r) \subseteq U$.

Since U is arbitrary and X is Hausdorff, it follows that $hz = z \in Fz$. Similarly, $z \in Gz$. Hence $hz = z \in Fz \cap Gz$. The uniqueness of z as a common fixed point of F, G and h can be easily verified.

Remark 3: Mappings F, G and h satisfying the condition

$$\delta_t(Fx,Gy) \leq a_t[H_t(hx,Fx) + H_t(hy,Gy)] + b_t[H_t(hx,Gy) + H_t(hy,Fx)] + c_t d_t(hx,hy),$$

where $x,y \in X$, $t \in I$ and a_t, b_t, c_t are nonnegative real numbers with $0 < 2a_t + 2b_t + 2c_t < 1$, also satisfy condition (2.1). Thus in case $hx = x$ for all $x \in X$ and $F = G$, Theorem 2 presents a stronger version of Iseki [5, Theorem 2].

Theorem 3: Let (X,u) be a complete uniform space. Suppose $F, G: X \to BN(X)$ and $h: X \to X$ are mappings such that $F(X) \cup G(X) \subseteq h(X)$ and for each $t \in I$, there exists a real number $q_t \in (0,1)$ satisfying

$$\delta_t(Fx,Gy) \leq q_t \max\{ d_t(hx,hy), \delta_t(hx,Fx), \delta_t(hy,Gy), \frac{1}{2}[D_t(hx,Gy) + D_t(hy,Fx)] \} \tag{3.1}$$

for all $x,y \in X$. If $h(X)$ is a closed subspace of X, then F, G, and h have a coincidence point.

Proof: Let $U \in u$ be arbitrary, and let $V(t,r) \in B'$ be such that $V(t,r) \subseteq U$. Now pick an $x_o \in X$. We may assume that $\delta_t(hx_o,Fx_o) > 0$. Since $q_t^{a_t}\delta_t(hx_o,Fx_o) < \delta_t(hx_o,Fx_o)$ for some $a_t \in (0,1)$, in view of the fact that $F(X) \cup G(X) \subseteq h(X)$, we may choose a point $hx_1 \in Fx_o$ such that $d_t(hx_o, hx_1) \geq q_t^{a_t}\delta_t(hx_o,Fx_o)$. Therefore $\delta_t(hx_o,Fx_o) \leq q_t^{-a_t}d_t(hx_o,hx_1)$. Similarly we can choose a point $hx_2 \in Gx_1$ such that $\delta_t(hx_1,Gx_1) \leq q_t^{-a_t}d_t(hx_1, hx_2)$. Continuing this process we can select a sequence $\{x_n\}$ such that $hx_{2n+1} \in Fx_{2n}$ and $hx_{2n+2} \in Gx_{2n+1}$ and

$$\delta_t(hx_{2n},Fx_{2n}) \leq q_t^{-a_t}d_t(hx_{2n},hx_{2n+1}),$$

$$\delta_t(hx_{2n+1},Gx_{2n;1}) \leq q_t^{-a_t}d_t(hx_{2n+1},hx_{2n+2}).$$

From (3.1),

$$d_t(hx_{2n+1},hx_{2n+2}) \leq \delta_t(Fx_{2n},Gx_{2n+1})$$

$$\leq q_t \max\{\, d_t(hx_{2n},hx_{2n+1}),$$

$$\delta_t(hx_{2n},Fx_{2n}),\ \delta_t(hx_{2n+1},Gx_{2n+1}),$$

$$\tfrac{1}{2}[D_t(hx_{2n},Gx_{2n+1})]\ \}.$$

Therefore one of the following relations holds

$$d_t(hx_{2n+1},hx_{2n+2}) \leq q_t\ d_t(hx_{2n},hx_{2n+1}),$$

$$d_t(hx_{2n+1},hx_{2n+2}) \leq q_t\ \delta_t(hx_{2n},Fx_{2n})$$

$$\leq q_t^{1-a_t} d_t(hx_{2n},hx_{2n+1}),$$

$$d_t(hx_{2n+1},hx_{2n+2}) \leq q_t\ \delta_t(hx_{2n+1},Gx_{2n+1})$$

$$\leq q_t^{1-a_t} d_t(hx_{2n+1},hx_{2n+2}),$$

$$d_t(hx_{2n+1},hx_{2n+2}) \leq \tfrac{1}{2}q_t\ D_t(hx_{2n},Gx_{2n+1})$$

$$\leq \tfrac{1}{2}q_t\ d_t(hx_{2n},hx_{2n+2})$$

$$\leq \tfrac{1}{2}q_t[d_t(hx_{2n},hx_{2n+1}) + d_t(hx_{2n+1},hx_{2n+2})].$$

Hence

$$d_t(hx_{2n+1},hx_{2n+2}) \leq p_t\ d_t(hx_{2n},hx_{2n+1})$$

where $p_t = \max\{\, q_t,\ q_t^{1-a_t},\ q_t/2-q_t\ \} < 1.$

Similarly $d_t(hx_{2n+2},\ hx_{2n+3}) \leq p_t\ d_t(hx_{2n+1},\ hx_{2n+2})$. Therefore $\{hx_n\}$ is a Cauchy sequence and converges to some point $b \in h(X)$. Hence there exists a point $z \in X$ such that $hz = b$. So it follows from

$$\delta_t(hz,Gz) \leq d_t(hz,hx_{2n+1}) + \delta_t(hx_{2n+1},Gz)$$

$$\leq d_t(hz,hx_{2n+1}) + \delta_t(Fx_{2n},Gz)$$

$$\leq d_t(hz,hx_{2n+1}) + q_t \max\{\, d_t(hx_{2n},hz),$$

$$\delta_t(hx_{2n}, Fx_{2n}),\ \delta_t(hz, Gz),$$
$$\tfrac{1}{2}[D_t(hx_{2n}, Gz) + D_t(hz, Fx_{2n})] \}$$
$$\leq d_t(hz, hx_{2n+1}) + q_t \max\{ d_t(hx_{2n}, hz),$$
$$q_t^{-a_t} d_t(hx_{2n}, hx_{2n+1}),\ \delta_t(hz, Gz),$$
$$\tfrac{1}{2}[d_t(hx_{2n}, hz) + \delta_t(hz, Gz) + d_t(hz, hx_{2n+1})] \}$$

that $\delta_t(hz, Gz) = 0 < r$. Consequently, $(hz, Gz) \in V(t, r) \subseteq U$. U being arbitrary and X being Hausdorff, it follows that $hz \in Gz$. Similarly $hz \in Fz$. Hence $hz \in Fz \cap Gz$.

Remark 4: If h is an identity mapping in Theorem 3, then the above theorem generalizes a result of Ciric [1, Theorem 3].

Theorem 4: Let (X, u) be a complete uniform space. Suppose X is bounded and $F, G : X \to B(X)$, $h : X \to X$ are mappings such that $F(X) \cup G(X) \subseteq h(X)$ and for each $t \in I$, there exists a number $q_t \in (0,1)$ such that for all $x, y \in X$

$$\delta_t(Fx, Gy) \leq q_t \max\{ d_t(hx, hy), \delta_t(hx, Fx), \delta_t(hy, Gy), \delta_t(hx, Gy), \delta_t(hy, Fx) \}. \tag{4.1}$$

If $h(X)$ is a closed subspace of X and h is commuting with F and G, then F, G, and h have a unique common fixed point.

Proof: Let $t \in I$ be arbitrary. It follows from (4.1) that if $x \in A$ and $y \in B$ for $A, B \in B(X)$, then

$$\delta_t(FA, FB) \leq q_t \max\{ \delta_t(hA, hB), \delta_t(hA, FA), \delta_t(hB, GB), \delta_t(hA, GB), \delta_t(hB, FA) \}. \tag{4.2}$$

Since X is bounded, we note that

$$k_t = \sup \{ \delta_t(A, B) : A, B \in B(X) \} < \infty.$$

Now we shall prove that

$$\delta_t(F^n G^n A, F^n G^n B) \leq q_t^n k_t \tag{4.3}$$

for $A, B \in B(X)$ and $n = 1, 2, 3, \ldots$. Since F, G, and h are commuting, we have

$$\begin{aligned}\delta_t(FGA, FGB) &= \delta_t(FGA, GFB) \\ &\leqslant q_t \max\{ \delta_t(hGA, hFB),\ \delta_t(hGA, FGA), \\ &\quad \delta_t(hFB, GFB),\ \delta_t(hGA, GFB),\ \delta_t(hFB, FGA) \} \\ &\leqslant q_t k_t\end{aligned}$$

Therefore (4.3) is true for $n = 1$. Suppose it is true for $n = m$. Then we have

$$\delta_t(F^{m+1}G^{m+1}A, F^{m+1}G^{m+1}B) = \delta_t(F(F^mG^{m+1}A), G(G^mF^{m+1}B)).$$

Now applying condition (4.2) and a simple calculation, we get

$$\begin{aligned}\delta_t(F^{m+1}G^{m+1}A, F^{m+1}G^{m+1}B) \leqslant q_t \max\{ &\delta_t(F^mG^m(GhA), F^mG^m(FhB)), \\ &\delta_t(F^mG^m(GhA), F^mG^m(FGA)), \\ &\delta_t(F^mG^m(FhB), F^mG^m(FGB)), \\ &\delta_t(F^mG^m(GhA), F^mG^m(FGB)), \\ &\delta_t(F^mG^m(FhB), F^mG^m(FGA)) \}.\end{aligned}$$

Therefore

$$\delta_t(F^{m+1}G^{m+1}A, F^{m+1}G^{m+1}B) \leqslant q_t q_t^m k_t = q_t^{m+1} k_t.$$

Hence by induction, the relation (4.3) is true for all $n = 1, 2, 3, \ldots$.

For $\varepsilon > 0$, choose a positive integer N such that $q_t^N k_t < \varepsilon$. Then from (4.3) we have

$$\delta_t(F^nG^nA, F^mG^mB) < \varepsilon \quad \text{for all } m, n \geqslant N \text{ and } A, B \in B(X). \tag{4.4}$$

With a fixed $A \in B(X)$, choose a point $hx_{2n} \in F^nG^nhA$ and a point $hx_{2n+1} \in F^nG^{n+1}hA$ for $n = 1, 2, \ldots$. Such a choice is permissible as $F(X) \cup G(X) \subseteq h(X)$. Then

$$d_t(hx_{2n}, hx_{2n+1}) \leqslant \delta_t(F^nG^n(hA), F^nG^n(GhA)) \leqslant q_t^n k_t$$

and

$$d_t(hx_{2n+1}, hx_{2n+2}) \leqslant \delta_t(F^nG^n(GhA), F^nG^n(FGhA)) \leqslant q_t^n k_t.$$

Therefore $\{hx_n\}_{n=2}^{\infty}$ is a Cauchy sequence in $h(X)$ and hence $hx_n \to b$ for some $b \in h(X)$. Hence there exists a $z \in X$ such that $hz = b$. Let $i \in I$ be arbitrary. Then

$$\begin{aligned}\delta_i(hz,Gz) &\leqslant d_i(hz,hx_{2n}) + \delta_i(hx_{2n},Gz)\\ &\leqslant d_i(hz,hx_{2n}) + \delta_i(F^nG^nhA,Gz).\end{aligned}$$

Now using condition (4.2) we get

$$\begin{aligned}\delta_i(hz,Gz) \leqslant d_i(hz,hx_{2n}) + q_i \max \{ &\delta_i(F^{n-1}G^nhhA,hz),\\ &\delta_i(F^{n-1}G^{n-1}(GhhA),F^nG^nhhA),\ \delta_i(hz,Gz),\\ &\delta_i(F^{n-1}G^nhhA,Gz),\ \delta_i(hz,F^nG^nhA) \}.\end{aligned} \tag{4.5}$$

Also,

$$\begin{aligned}\delta_i(hz,F^nG^nhA) &\leqslant d_i(hz,hx_{2m}) + \delta_i(hx_{2m},F^nG^nhA)\\ &\leqslant d_i(hz,hx_{2m}) + \delta_i(F^mG^mhA,F^nG^nhA)\\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon \quad \text{for all} \quad m,n \geqslant N\end{aligned}$$

follows from (4.4) and the fact that $hx_{2m} \to hz$. Hence for all $n \geqslant N$,

$$\delta_i(hz,F^nG^nhA) \leqslant \varepsilon. \tag{4.6}$$

Similarly it can be proved that

$$\begin{aligned}&\delta_i(hz,F^{n-1}G^nhhA) < \varepsilon,\\ &\delta_i(hz,F^nG^nhA) < \varepsilon\end{aligned} \tag{4.7}$$

for all $n \geqslant N$. Therefore from (4.5) - (4.7), it follows for all $n \geqslant N$ that

$$\delta_i(hz,Gz) \leqslant \varepsilon + q_i\max\{ \varepsilon,\ \varepsilon,\ \delta_i(hz,Gz),\ \delta_i(hz,Gz) + \varepsilon,\ \varepsilon \}.$$

Hence $\delta_i(hz,Gz) \leqslant (1 + q_i)(1 - q_i)^{-1}\varepsilon$. Since $\varepsilon > 0$ is arbitrary and X is Hausdorff, it follows that $\{hz\} = Gz$. Similarly $\{hz\} = Fz$.

Now we shall prove that $hz = hhz$. Let $j \in I$ be arbitrary. Then

$$d_j(hhz,hz) \leq d_j(hhz,hx_{2n}) + d_j(hx_{2n},hz)$$
$$\leq d_j(hhz,hx_{2n}) + d_j(F^nG^nhA,Ghz).$$

Using condition (4.2), it follows that for all $n \geq N$

$$d_j(hhz,hz) \leq \varepsilon + q_j \max\{ \varepsilon, \varepsilon, 0, \varepsilon, \varepsilon \}$$
$$= (1 + q_j)\varepsilon.$$

This proves that $hz = hhz$. Thus hz is a common fixed point of F, G and h. The uniqueness of hz as a common fixed point of F, G, and h can be easily verified.

Remark 5: The above theorem generalizes the results of Fisher [2, 3, 4]. In fact, the main theorem of Fisher [3] can be obtained as a special case of the above theorem by choosing X to be a metric space and h an identity mapping on X.

References:

1. Ciric, L.B.: Fixed points for generalized multivalued contractions. Mat. Vesnik 9, 265-272 (1972)

2. Fisher, B.: Results on common fixed points on bounded metric spaces. Math. Sem. Notes Kobe Univ. 7, 73-80 (1979)

3. Fisher, B.: Setvalued mappings on bounded metric spaces. Indian J. Pure Appl. Math. 11, 8-12 (1980)

4. Fisher, B.: Setvalued mappings on metric spaces. Fund. Math. 112, 141-145 (1981)

5. Iseki, K.: Multivalued contraction mappings in complete metric spaces. Rend. Sem. Mat. Univ. Padova 53, 15-19 (1975)

6. Kasahara, S.: Common fixed points of multivalued mappings in L-spaces. Math. Sem. Notes Kobe Univ. 4, 181-193 (1976)

7. Kubiak, T.: Fixed point theorems for contractive type multivalued mappings. Math. Japon. 30, 89-101 (1985)

8. Mishra, S.N.: A note on common fixed points of multivalued mappings in uniform spaces. Math. Sem. Notes Kobe Univ. 9, 341-347 (1981)

9. Mishra, S.N.: Fixed points of contractive type multivalued mappings. Indian J. Pure Appl. Math. 18, 283-289 (1987)

10. Nadler, S.B., Jr.: Multivalued contraction mappings. Pacific J. Math. 30, 475-488 (1969)

11. Reich, S.: Fixed points of contractive functions. Boll. Un. Mat. Ital. 5, 26-42 (1972)

12. Singh, K.L., and Whitfield, J.H.M.: Fixed points for contractive type multivalued mappings. Math. Japon. 27, 117-124 (1982)

13. Thron, W.J.: Topological Structures. New York 1966

received: 23. 3. 1989

Authors:

Dr. S.N. Mishra
Departement of Mathematics & Computer Sciences
National University of Lesotho
P.O. Roma 180
Lesotho

Dr. S.L. Singh
Departement of Mathematics
G.K. University
Haridwar 249404
India

Rostock. Math. Kolloq. 40, 71 - 82 (1989) 62F12 62F25

Hartmut Fischer

Stichprobenumfangsbestimmung für Konfidenzintervalle höherer Ordnung

1. Einleitung

Maximum-Likelihood-Schätzer gehören zu den wirksamsten Schätzern von Parametern einer Verteilung. Sie werden deshalb oft zur Konstruktion von Konfidenzintervallen benutzt. Die vorliegende Arbeit bezieht sich auf kontinuierliche Verteilungen P_θ mit einer Dichte $f(y;\theta)$, θ sei der p-dimensionale Parametervektor, $p \geq 1$. Dann ist bekannt, daß unter gewissen Regularitätsbedingungen die Verteilung der Zufallsgröße $\underline{W}_n = \sqrt{n}\left(\hat{\underline{\theta}}_n - \theta_0\right)$ schwach gegen eine Normalverteilung mit Erwartungswertvektor $(0,\ldots,0)^T \in \mathbb{R}^p$ und Kovarianzmatrix $I^{-1}(\theta_0)$ konvergiert, wobei $\hat{\underline{\theta}}_n$ eine Maximum-Likelihood-Schätzung von θ_0 ist und $I(\theta_0)$ die Fishersche Informationsmatrix bezeichnet. Diese asymptotische Normalverteilung wird oft zur Konstruktion von Konfidenzintervallen genutzt, ohne jedoch zu berücksichtigen, daß dieses Vorgehen nur asymptotisch zulässig ist, das heißt, daß das Konfidenzniveau nur für $n \to \infty$ eingehalten wird.

Eine Möglichkeit, die Asymptotik zu verbessern, bieten die Edgeworthentwicklungen. Dabei werden zur asymptotischen Normalverteilung noch Korrekturterme addiert, die mit Potenzen in $\frac{1}{\sqrt{n}}$ versehen sind.

Unter 2. wird auf die allgemeine Herleitung einer Edgeworthentwicklung eingegangen und das spezielle Vorgehen bei der Approximation der Verteilung eines Maximum-Likelihood-Schätzers erläutert.

Im 3. Abschnitt wird davon ausgegangen, daß eine eindimensionale Edgeworthentwicklung zweiter Ordnung zur Verfügung steht. Diese wird benutzt, um Quantile gleicher Ordnung zu ermitteln.

Solche Quantile höherer Ordnung finden dann in 4. Verwendung bei der Stichprobenumfangsbestimmung für Konfidenzintervalle des Parameters θ_0 $(p = 1)$ bzw. einer seiner Komponenten $(p > 1)$.

2. Allgemeine Herleitung der Edgeworthentwicklung

Es sei $\{\underline{Z}_n\}_{n\geq 1}$ eine Folge unabhängig und identisch verteilter, r-dimensionaler Zufallsvektoren mit

$$\mu = E\underline{Z}_1 , \qquad \underline{Z} = \tfrac{1}{n}\sum_{i=1}^{n}\underline{Z}_i . \tag{2.1}$$

Wir betrachten jetzt die Statistik

$$\underline{W}_n = \sqrt{n}\Big(H(\underline{Z})-H(\mu)\Big) , \tag{2.2}$$

wobei H eine reellwertige meßbare Funktion auf dem $\mathbb{R}^r$ ist. $\underline{Z}_1$ habe endliche Momente zweiter Ordnung, und H sei stetig differenzierbar in einer Umgebung von μ. Es ist bekannt, daß dann $\underline{W}_n$ asymptotisch normalverteilt ist mit Erwartungswert 0 und Varianz $\sigma^2 = J_H \Sigma J_H^T$, wobei J_H den Gradienten von H an der Stelle μ und Σ die Kovarianzmatrix von $\underline{Z}_1$ bezeichnen. Es wird vorausgesetzt, daß σ^2 positiv ist. H sei $(s-1)$-mal stetig differenzierbar in μ. Ersetzt man $H(\underline{Z})$ in der Statistik $\underline{W}_n$ durch deren nach dem $(s-1)$-ten Glied abgebrochene Taylorreihe an der Stelle μ, so erhält man die Näherung $\underline{W}'_n$ von $\underline{W}_n$. Die Kumulanten $\kappa_{j,n}'$ von $\underline{W}'_n$ haben die Struktur

$$\kappa_{j,n}' = \mathbb{1}_{\{2\}}(j)\sigma^2 + \sum_{i=1}^{s-2} n^{-i/2}\beta_{ji} , \qquad j=1,\ldots,s, \tag{2.3}$$

wobei $\mathbb{1}_M(x)$ die Indikatorfunktion bezeichnet. Die β_{ji} hängen nur von den gemischten Momenten des Vektors $\underline{Z}-\mu$ ab, die sich aus den Momenten von $\underline{Z}_1$ berechnen lassen, und von den Ableitungen von H an der Stelle μ höchstens bis zur Ordnung $s-1$. Es ist leicht zu sehen, daß für gerades j nur gerade und für ungerades j nur ungerade Potenzen von $\frac{1}{\sqrt{n}}$ in der Entwicklung (2.3) auftreten.

Wenn man im Ausdruck $\exp(\ln C_{\underline{W}'_n}(t))$ den Logarithmus der charakteristischen Funktion von $\underline{W}'_n$ in eine Reihe entwickelt und nach dem s-ten Glied abbricht, so erhält man

$$\exp\Big((it)\kappa_{1,n}' + \frac{(it)^2}{2!}\big(\kappa_{2,n}'-\sigma^2\big) + \sum_{j=3}^{s}\frac{(it)^j}{j!}\kappa_{j,n}'\Big)\ \exp(-\tfrac{1}{2}\sigma^2 t^2) \tag{2.4}$$

als eine Approximation der charakteristischen Funktion von $\underline{W}'_n$. Entwickelt man den ersten Exponentialfaktor nochmals in eine

Potenzreihe, setzt die $\varkappa_{j,n}'$ ein und ordnet nach Potenzen von $\frac{1}{\sqrt{n}}$, so erhält man den Ausdruck

$$\left(1 + \sum_{k=1}^{s-2} n^{-k/2} T_k(it)\right) \exp(-\tfrac{1}{2}\sigma^2 t^2) , \tag{2.5}$$

wobei die T_k Polynome sind, deren Koeffizienten sich aus den β_{ji} ergeben und damit nicht von n abhängen. Die formale Edgeworthentwicklung $\Psi_{s-2,n}$ der Verteilungsfunktion Ψ_n von $\underline{W}_n$ ist bestimmt durch

$$\left.\begin{aligned} \Psi_{s-2,n}(x) &= \int_{-\infty}^{x} \psi_{s-2,n}(u)du , \\ \psi_{s-2,n}(u) &= \left(1 + \sum_{k=1}^{s-2} n^{-k/2} T_k\left(-\frac{d}{du}\right)\right) \varphi(u;\sigma^2) , \end{aligned}\right\} \tag{2.6}$$

wobei $\varphi(u;\sigma^2)$ die Dichte einer Normalverteilung mit Erwartungswert 0 und Varianz σ^2 bezeichnet. Die $T_k\left(-\frac{d}{du}\right)$ sind Differentialoperatoren, bei deren Anwendung sich die Korrekturpolynome der Edgeworthentwicklung ergeben. Es ist leicht zu sehen, daß die Fouriertransformierte von $\Psi_{s-2,n}$ gerade (2.5) ist.

<u>Definition 2.1:</u> Eine Entwicklung der Form (2.6) mit $q = s-2$ heißt Edgeworthentwicklung der Ordnung q der Verteilungsfunktion Ψ_n von $\underline{W}_n$, falls

$$\Psi_n(x) = \Psi_{q,n}(x) + o\left(n^{-q/2}\right) \tag{2.7}$$

gilt für $n \to \infty$.

<u>Bemerkung 2.1:</u> Analoges Vorgehen führt bei einer vektorwertigen (p-dimensionalen) Funktion H zu einer p-dimensionalen Edgeworthentwicklung der Verteilung des Vektors $\underline{W}_n$. J_H ist dann eine Matrix von Ableitungen vom Format $p \times r$. Die Summanden in (2.4) sind Multilinearformen, wobei das $t \in \mathbb{R}^1$ durch $\tau \in \mathbb{R}^p$ und σ^2 durch die asymptotische Kovarianzmatrix von $\underline{W}_n$ vom Format $p \times p$ zu ersetzen sind. Die Kumulanten $\varkappa_{j,n}'$ sind Tensoren der Ordnung j mit $\varkappa_{j,n}' \in \mathbb{R}^{p^j}$, $j = 1,..,s$, und die Komponenten von $\varkappa_{j,n}'$ sind die gemischten Kumulanten j-ter Ordnung von $\underline{W}_n'$. Sie sollen entsprechend ihrer Indizierung in den Gitterpunkten eines j-dimensionalen Würfels mit Kantenlänge p angeordnet sein. Das Integral in (2.6) ist ein Mehrfachintegral, und die T_k sind Differentialoperatoren mit partiellen Ableitungen nach p Vari-

ablen.

Bemerkung 2.2: Wir wollen jetzt die Struktur der formalen Edgeworthentwicklung einer eindimensionalen Randverteilung von $\underline{W}_n$ bestimmen. Für $p > 1$ sei mit W_n die i-te Komponente von $\underline{W}_n$ bezeichnet, $i \in \{1,..,p\}$. In der obigen Anordnung der Komponenten von $\varkappa_{j,n}'$ steht jeweils an der i-ten Stelle der Raumdiagonalen die j-te Kumulante $\dot{\varkappa}_{j,n}'$ von $\dot{W}_n'$. Den Logarithmus der charakteristischen Funktion von $\dot{\underline{W}}_n'$ erhält man, indem $\tau = te_i$ gesetzt wird, $t \in \mathbb{R}^1$, e_i der i-te Einheitsvektor im $\mathbb{R}^p$. Berechnet man mit diesem τ die in Bemerkung 2.1 beschriebenen Multilinearformen, so reduzieren sich diese zu gewöhnlichen Produkten wie in (2.4), wobei dann in (2.4) die $\varkappa_{j,n}'$ durch $\dot{\varkappa}_{j,n}'$ und σ^2 durch die asymptotische Varianz $\dot{\sigma}^2$ von $\dot{\underline{W}}_n$ zu ersetzen sind. Der weitere Weg bis zur Bestimmung der Edgeworthentwicklung entspricht dem in (2.4) bis (2.6) vorgezeichneten. Damit sind wir in der Lage, Edgeworthentwicklungen eindimensionaler Randverteilungen zu bestimmen, ohne die p-dimensionale Entwicklung zu kennen, deren Existenz jedoch vorausgesetzt werden muß.

Edgeworthentwicklungen von Verteilungen von Zufallsgrößen der Form (2.2) lassen sich aber nicht nur für explizit gegebene Funktionen berechnen, sondern auch für implizit gegebene H, weil nur die Ableitungen von H an der Stelle μ, die sukzessive ermittelt werden können, in deren Berechnung eingehen. Z.B. sei $H(\underline{Z}) = \hat{\underline{\theta}}_n$, $\hat{\underline{\theta}}_n$ der Maximum-Likelihood-Schätzer des p-dimensionalen Parametervektors θ_0 einer Verteilung mit der Dichte $f(y;\theta_0)$. Sei jetzt

$$f_\nu(y;\theta_0) = D^\nu \ln f(y;\theta_0) , \tag{2.8}$$

wobei

$$D^\nu = \frac{\partial^{|\nu|}}{\partial\theta_1^{\nu_1}..\partial\theta_p^{\nu_p}} ,$$

$$\nu = (\nu_1,..,\nu_p) , \quad \nu_j \in \mathbb{N}_0 , \quad j=1,..,p,$$

$$|\nu| = \nu_1+..+\nu_p , \quad \nu! = \nu_1!..\nu_p! .$$

D^ν ist also ein Differenzialoperator bezüglich des Parametervektors θ. Desweiteren sei $\{\underline{Y}_n\}_{n\geq 1}$ eine Folge unabhängiger Zufallsgrößen mit der Dichte $f(y;\theta_0)$. Dann erhält man für die p Loglikelihoodgleichungen

$$0 = \tfrac{1}{n}\sum_{i=1}^{n} f_{e_j}(\underline{Y}_i;\theta) \ , \qquad j=1,..,p, \tag{2.9}$$

deren Lösungen die ML-Schätzungen $\underline{\hat{\theta}}_{n,j}$ sind. e_j bezeichnet dabei den j-ten Einheitsvektor im $\mathbb{R}^p$. Werden die Gleichungen in eine Taylorreihe an der Stelle θ_0 entwickelt, so ergibt sich

$$\begin{aligned} 0 &= \tfrac{1}{n}\sum_{i=1}^{n} f_{e_j}(\underline{Y}_i;\theta_0) \\ &\quad + \sum_{|\nu|=1}^{s-1} \left(\tfrac{1}{n}\sum_{i=1}^{n} f_{e_j+\nu}(\underline{Y}_i;\theta_0)\frac{(\theta-\theta_0)^\nu}{\nu!} \right) + R_{n,j}(\theta) \qquad (2.10) \\ &= \bar{z}_{e_j} + \sum_{|\nu|=1}^{s-1} \bar{z}_{e_j+\nu} \frac{(\theta-\theta_0)^\nu}{\nu!} + R_{n,j}(\theta) \ , \qquad j=1,..,p, \end{aligned}$$

wobei die $\bar{z}_\varkappa$, $1\leq|\varkappa|\leq s$, die Komponenten von $\underline{Z}$ bezeichnen. Die Dimension dieses Vektors beträgt

$$r = \sum_{k=1}^{s}\binom{p+k-1}{k} \ , \tag{2.11}$$

wobei der k-te Summand die Zahl der verschiedenen partiellen Ableitungen k-ter Ordnung bei p Variablen angibt. Die r Komponenten von $\mu = E\underline{Z} = E\underline{Z}_1$ sind also

$$\mu_\nu = Ef_\nu(\underline{Y}_1;\theta_0) \ , \qquad 1\leq|\nu|\leq s. \tag{2.12}$$

Vernachlässigt man das Restglied in (2.10), so erhält man

$$\Pi_j(X,\theta) = x_{e_j} + \sum_{|\nu|=1}^{s-1} x_{e_j+\nu} \frac{(\theta-\theta_0)^\nu}{\nu!} = 0 \ , \qquad j=1,..,p, \tag{2.13}$$

als Aproximationen der rechten Seiten der p Gleichungen in (2.9) in den $r+p$ Variablen $X = \underline{Z}$ und θ. Mit $\theta = H(X)$ ist durch (2.13) eine implizit gegebene Funktion H erklärt, deren Ableitungen im Punkt μ zu berechnen sind, wobei $H(\mu) = \theta_0$ gilt. Damit ist ein Vorgehen gemäß (2.2) bis (2.6) auch für Verteilungen von Maximum-Likelihood-Schätzern prinzipiell möglich. In [1] sind Bedingungen dafür angegeben, daß die auf diese Weise formal hergeleitete Entwicklung auch tatsächlich eine asymptotische Entwicklung ist, das heißt, daß die Differenz zwischen der Verteilungsfunktion von $\underline{W}_n = \sqrt{n}\left(\hat{\underline{\theta}}_n - \theta_0\right)$ und der zu ermittelnden Näherungsfunktion von der Ordnung $o\left(n^{-(s-2)/2}\right)$ ist.

3. Asymptotische Entwicklungen für Quantile

Die folgenden Überlegungen gelten nur für eindimensionale Verteilungen von Maximum-Likelihood-Schätzern bzw. eindimensionalen Randverteilungen. Deshalb stehen im weiteren $\underline{W}_n$ und σ^2 auch gleichzeitig für $\dot{\underline{W}}_n$ und $\dot{\sigma}^2$ (siehe Bemerkung 2.2).

Bemerkung 3.1: Wir bezeichnen mit $\underline{Y}_n$ die standardisierte Zufallsgröße

$$\underline{Y}_n = \frac{1}{\sigma}\underline{W}_n = \frac{\sqrt{n}}{\sigma}\left(\hat{\underline{\theta}}_n - \theta_0\right) , \tag{3.1}$$

ihre Verteilungsfunktion mit F_n und die zugehörige Edgeworthentwicklung q-ter Ordnung mit $F_{q,n}$.

Folgerung: Damit ergibt sich folgender Zusammenhang zwischen der Verteilungsfunktion Ψ_n von $\underline{W}_n$ und $F_{q,n}$:

$$\Psi_n(x) = F_n\left(\frac{x}{\sigma}\right) = F_{q,n}\left(\frac{x}{\sigma}\right) + o\left(n^{-q/2}\right) . \tag{3.2}$$

Bemerkung 3.2: Eine Edgeworthentwicklung q-ter Ordnung der Verteilung einer standardisierten Zufallsgröße $\underline{Y}_n$ der Form (2.6) läßt sich in eine Edgeworthentwicklung q-ter Ordnung der Form

$$F_{q,n}(y) = \Phi(y) - \varphi(y)\sum_{k=1}^{q} n^{-k/2}P_k(y) \tag{3.3}$$

überführen, wobei die P_k Polynome sind, deren Koeffizienten sich aus der Anwendung der Differentialoperatoren $T_k\left(-\frac{d}{du}\right)$ auf $\varphi(u;\sigma^2)$ ergeben. Man erhält sie also aus den Koeffizienten der T_k und denen der Hermite-Polynome. $\Phi(y)$ bezeichnet dabei die Verteilungsfunktion der Standardnormalverteilung.

Folgerung: Für $q = 2$ erhält man eine eindimensionale Edgeworthentwicklung zweiter Ordnung der Gestalt

$$F_{2,n}(y) = \Phi(y) - \varphi(y)\left(\frac{1}{\sqrt{n}} P_1(y) + \frac{1}{n}P_2(y)\right) , \qquad (3.4)$$

wobei

$$P_1(y) = d_2y^2 + d_o ,$$

$$P_2(y) = d_5y^5 + d_3y^3 + d_1y$$

gilt (siehe z.B. [2]).

Wir wollen jetzt die Korrekturgrößen der Entwicklungen zugehöriger Quantile gleicher Ordnung bestimmen.

<u>Definition 3.1:</u> Eine Entwicklung der Form

$$y_{q,n}(\gamma) = Q_o(\gamma) + \sum_{k=1}^{q} n^{-k/2}Q_k(\gamma) \qquad (3.5)$$

heißt Cornish-Fisher-Entwicklung der Ordnung q (oder Quantil q-ter Ordnung) für das γ-Quantil $y_n(\gamma)$ der Verteilungsfunktion F_n, falls gilt

$$F_n(y_{q,n}(\gamma)) = \gamma + o\left(n^{-q/2}\right) . \qquad (3.6)$$

<u>Satz 3.1:</u> Gegeben sei eine Edgeworthentwicklung der Gestalt (3.4) einer Verteilungsfunktion F_n. Dann ist

$$y_{2,n}(\gamma) = Q_o(\gamma) + \frac{1}{\sqrt{n}} Q_1(\gamma) + \frac{1}{n}Q_2(\gamma) \qquad (3.7)$$

eine Cornish-Fisher-Entwicklung zweiter Ordnung, wobei

$$Q_o(\gamma) = z_\gamma ,$$

$$Q_1(\gamma) = P_1(z_\gamma) ,$$

$$Q_2(\gamma) = e_5z_\gamma^5 + e_3z_\gamma^3 + e_1z_\gamma$$

mit

$$e_5 = d_5 - \frac{1}{2}d_2^2 , \qquad e_3 = d_3 + 2d_2^2 - d_2d_o ,$$

$$e_1 = d_1 + 2d_2d_o - \frac{1}{2}d_o^2 .$$

Dabei bezeichnet $z_\gamma = \Phi^{-1}(\gamma)$ das γ-Quantil der Standardnormalverteilung.

Beweis: Wir setzen $t = \frac{1}{\sqrt{n}}$ und machen den Ansatz

$$y(t) = Q_o(\gamma) + tQ_1(\gamma) + t^2Q_2(\gamma) . \tag{3.8}$$

Wir wollen jetzt $Q_i(\gamma)$, $i = 0,1,2$, so bestimmen, daß

$$F_n(y(t)) = \gamma + o\left(t^2\right) \tag{3.9}$$

gilt. Dazu betrachten wir die Funktion

$$\begin{aligned} L(t) &= F_{2,n}(y(t)) \\ &= \Phi(y(t)) - \varphi(y(t))\left(tP_1(y(t)) + t^2P_2(y(t))\right) . \end{aligned} \tag{3.10}$$

Es ist also $L(t) = \gamma + o\left(t^2\right)$.

Dann ist $L(0) = \Phi(Q_o(\gamma)) = \gamma$, also

$$Q_o(\gamma) = \Phi^{-1}(\gamma) = z_\gamma . \tag{3.11}$$

Im folgenden werden Ableitungen nach t gebildet, zum Teil unter Anwendung der Kettenregel, wobei die Argumente auf der rechten Seite weggelassen werden. Zur Unterscheidung bezeichnen wir die Ableitung nach t mit einem Punkt und die nach y mit einem Strich. Dann ist

$$\dot{L}(t) = \varphi\dot{y} - \varphi'\dot{y}\left(tP_1+t^2P_2\right) - \varphi\left(P_1+tP_1'\dot{y}+2tP_2+t^2P_2'\dot{y}\right) . \tag{3.12}$$

Nun erhalten wir durch Koeffizientenvergleich an der Stelle $t = 0$ die Gestalt von $Q_1(\gamma)$. Man beachte, daß $y(0) = Q_o(\gamma)$, $\dot{y}(0) = Q_1(\gamma)$ und $\ddot{y}(0) = 2Q_2(\gamma)$ gilt. Somit ist

$$\dot{L}(0) = \varphi(Q_o(\gamma))Q_1(\gamma) - \varphi(Q_o(\gamma))P_1(Q_o(\gamma)) = 0 ,$$

also

$$Q_1(\gamma) = P_1(z_\gamma) . \tag{3.13}$$

Auf analoge Weise erhält man $Q_2(\gamma)$ durch Bildung der zweiten Ableitung von $L(t)$. Es gilt

$$\begin{aligned} \ddot{L}(t) &= \varphi'\dot{y}^2 + \varphi\ddot{y} - \left(\varphi''\dot{y}^2+\varphi'\ddot{y}\right)\left(tP_1+t^2P_2\right) \\ &\quad - 2\varphi'\dot{y}\left(P_1+tP_1'\dot{y}+2tP_2+t^2P_2'\dot{y}\right) \end{aligned} \tag{3.14}$$

$$- \varphi\left(P_1^{\cdot} y + P_1^{\cdot} \dot{y} + tP_1^{\cdot\cdot} \dot{y}^2 + tP_1^{\cdot} \ddot{y} + 2P_2 + 2tP_2^{\cdot} \dot{y} + 2tP_2^{\cdot} \dot{y} + t^2 P_2^{\cdot\cdot} \dot{y}^2 + t^2 P_2^{\cdot} \ddot{y}\right) .$$

Somit ist

$$\begin{aligned} \ddot{L}(0) &= \varphi^{\cdot} Q_1^2 + 2\varphi Q_2 - 2\varphi^{\cdot} Q_1 P_1 - \varphi\left(2P_1^{\cdot} Q_1 + 2P_2\right) \\ &= \varphi^{\cdot}\left(-P_1^2\right) + 2\varphi\left(Q_2 - P_1^{\cdot} P_1 - P_2\right) \\ &= \varphi\left(Q_0 P_1^2 + 2\left(Q_2 - P_1^{\cdot} P_1 - P_2\right)\right) = 0 , \end{aligned}$$

also

$$Q_2(\gamma) = P_2(z_\gamma) + P_1^{\cdot}(z_\gamma) P_1(z_\gamma) - \tfrac{1}{2} z_\gamma P_1^2(z_\gamma) . \tag{3.15}$$

Setzt man die Darstellung (3.4) von P_1 und P_2 in (3.15) ein und ordnet nach Potenzen von z_γ, so liefert das die Behauptung.

<u>Folgerung:</u> Es gilt

$$P(\underline{Y}_n < y_{2,n}(\gamma)) = \gamma + o\left(\frac{1}{n}\right) , \tag{3.16}$$

wobei

$$y_{2,n}(\gamma) = z_\gamma + \frac{1}{\sqrt{n}} Q_1^*(z_\gamma) + \frac{1}{n} Q_2^*(z_\gamma) \tag{3.17}$$

mit $Q_i^* = Q_i \circ \Phi$, $i=1,2$, Q_i gemäß (3.7).

4. Bestimmung des notwendigen Stichprobenumfangs für Konfidenzintervalle vorgegebener Länge

<u>Definition 4.1:</u> Das zufällige Intervall

$$\underline{B} = B(1-\alpha, \underline{\hat{\theta}}_n, d) \subset \mathbb{R}^1 , \quad d>0, \tag{4.1}$$

heißt Konfidenzintervall zum Niveau $1-\alpha$, falls gilt

$$P(\theta_0 \in \underline{B}) = 1-\alpha . \tag{4.2}$$

$\underline{B}$ habe die Form

$$\text{(a)} \quad \underline{B}_o = (\underline{\hat{\theta}}_n - d, +\infty) , \tag{4.3}$$

$$\text{(b)} \quad \underline{B}_u = (-\infty, \underline{\hat{\theta}}_n + d) \tag{4.4}$$

oder

(c) $\underline{B}_2 = (\hat{\underline{\theta}}_n - d_u, \hat{\underline{\theta}}_n + d_o)$, $d_u + d_o = 2d$, (4.5)

sei also ein nach oben offenes, ein nach unten offenes bzw. ein zweiseitiges Konfidenzintervall zum Niveau $1-\alpha$. Für $\underline{B}_2$ sei insbesondere verlangt, daß es symmetrisch bezüglich α ist, das heißt, es soll

$$P(\theta_0 < \hat{\underline{\theta}}_n - d_u) = \frac{\alpha}{2} , \qquad P(\theta_0 > \hat{\underline{\theta}}_n + d_o) = \frac{\alpha}{2} \tag{4.6}$$

gelten.
Auf Grund der Konsistenz von Maximum-Likelihood-Schätzern existiert ein minimales n mit (4.2).

<u>Definition 4.2:</u> $n_q = n(q,B)$, $q = 0,1,2,\dots$, (4.7)

heißt minimaler Stichprobenumfang q-ter Ordnung für das Konfidenzintervall $\underline{B}$, falls in (4.2) bei der Berechnung des notwendigen Stichprobenumfangs n das Quantil $y_n(\gamma)$ der Verteilungsfunktion F_n von $\underline{V}_n$ ersetzt wird durch dessen Cornish-Fisher-Entwicklung q-ter Ordnung $y_{q,n}(\gamma)$, $q \geq 1$. Für $q = 0$ wird $y_n(\gamma)$ durch das Quantil der asymptotischen Normalverteilung von $\underline{V}_n$ ersetzt.

<u>Satz 4.1:</u> Den minimalen Stichprobenumfang q-ter Ordnung n_q, $q = 0,1,2$, erhält man für Konfidenzintervalle der Form (4.3), (4.4) bzw. (4.5) aus den Gleichungen

(i) $q = 0$: $t = u$, $n_0 = t^2$,

(ii) $q = 1$: $t^2 - ut - v = 0$, $t>0$, $n_1 = t^2$,

(iii) $q = 2$: $t^3 - ut^2 - vt - w = 0$, $t>0$, $n_2 = t^2$,

mit

$u = \frac{6}{\Delta} z_\gamma$, $v = \frac{6}{\Delta} Q_1^*(z_\gamma)$ bzw. $v = 0$ in (c),

$w = \frac{6}{\Delta} Q_2^*(z_\gamma)$,

wobei in

(a) $\gamma = 1-\alpha$, $\Delta = d$,

(b) $\gamma = \alpha$, $\Delta = -d$,

(c) $\gamma = 1-\frac{\alpha}{2}$, $\Delta = d$

gilt. In (ii) und (iii) kann es durchaus vorkommen, daß keine Lösungen existieren. Gibt es mehrere, so wähle man die kleinste positive Lösung.

<u>Beweis:</u> (a) Der Stichprobenumfang n muß so gewählt werden, daß

$$P(\theta_0 \in \underline{B}_o) = P(\theta_0 > \hat{\underline{\theta}}_n - d) = P(\hat{\underline{\theta}}_n - \theta_0 < d) = F_n(\tfrac{d}{\sigma}\sqrt{n}) = 1-\alpha ,$$

also

$$\frac{d}{\sigma}\sqrt{n} = y_n(1-\alpha) \tag{4.8}$$

gilt. Gemäß Definition 4.2 bestimmen wir n_q aus der Gleichung

$$\sqrt{n} - \frac{\sigma}{d} y_{q,n}(1-\alpha) = 0 . \tag{4.9}$$

Sei jetzt $t = \sqrt{n}$. Setzt man für $y_{q,n}$ nach (3.5) bzw. (3.17) ein und multipliziert jeweils mit t^q, so erhält man die im Satz genannten Gleichungen ersten, zweiten und dritten Grades.

(b) Analog sind die Umformungen für $\underline{B}_u$. Man erhält

$$\sqrt{n} + \frac{\sigma}{d} y_{q,n}(\alpha) = 0 . \tag{4.10}$$

Die weiteren Überlegungen sind entsprechend.

(c) Aus der Forderung (4.6) erhält man die Gleichungen

$$\frac{d_u}{\sigma}\sqrt{n} = y_n(1-\tfrac{\alpha}{2}) , \qquad \frac{d_o}{\sigma}\sqrt{n} = y_n(\tfrac{\alpha}{2}) , \tag{4.11}$$

wobei

$$y_n(1-\tfrac{\alpha}{2}) - y_n(\tfrac{\alpha}{2}) = \frac{d_u+d_o}{\sigma} = \frac{2d}{\sigma}\sqrt{n}$$

gelten soll. n_q wird also aus der Gleichung

$$\sqrt{n} - \frac{\sigma}{2d}\left(y_{q,n}(1-\tfrac{\alpha}{2}) - y_{q,n}(\tfrac{\alpha}{2})\right) = 0 \tag{4.12}$$

ermittelt. Man beachte, daß $z_\gamma = -z_{1-\gamma}$, $Q_1^*(-z) = Q_1^*(z)$ und $Q_2^*(-z) = -Q_2^*(z)$ gilt. Das weitere Vorgehen ist analog.

<u>Folgerung:</u> Wegen $v = 0$ gilt in (c) $n_1 = n_o$. Die Werte für d_u und d_o gewinnt man aus (4.11), wobei für n das zuvor berechnete n_q eingesetzt wird.

Bemerkung 4.1: Die asymptotische Varianz habe die spezielle Struktur $\sigma = h\theta_0$, wobei h eine Konstante ist, die für $p > 1$ auch von den anderen $p-1$ Parametern abhängen kann. Die Größe $v = d/\theta_0$ kann als relative Abweichung vom wahren Parameter θ_0 angesehen werden. Es ist also $\sigma/d = h/v$. Man kann demnach σ/d durch h/v in allen Gleichungen zur Bestimmung von n_q ersetzen. Somit ist der minimale Stichprobenumfang q-ter Ordnung nicht mehr von d und θ_0 einzeln abhängig, sondern nur noch von ihrem Verhältnis v.

Schlußbemerkung: Weil nach dem minimalen n gesucht wird, kommt man zwangsläufig in die Bereiche, wo die Edgeworth- und die Cornish-Fisher-Entwicklungen auch schlechte Approximationen liefern. Trotzdem zeigen Vergleiche mit Simulationsergebnissen bei der Weibullverteilung (bisher unveröffentlicht), daß diese Art der Stichprobenplanung zu besseren Ergebnissen führt, als wenn man nur die asymptotische Normalverteilung benutzt.

Literatur:

1. Bhattacharya, R.N. ,and Ghosh, J.K.: On the validity of the formal Edgeworth expansion. Ann. of Statist. 6, 434-451 (1978)

2. Gnedenko, B.W., und Kolmogorov, A.N.: Grenzverteilungen von Summen unabhängiger Zufallsgrößen. Berlin 1959

eingegangen: 22. 2. 1989 revidierte Fassung: 3. 9. 1989

Verfasser:

Dipl.-Math. Hartmut Fischer
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
Rostock
DDR-2500

Rostock. Math. Kolloq. 40, 83 - 94 (1990) 90C30 90C50 62K05 62P10

Petra Duchrau; Kurt Frischmuth

A numerical search algorithm for experimental design

1. Introduction

In the recent years there has been paid much attention to algorithms for the construction of experimental designs which are optimal for the estimation of the parameter ϑ in models of the type

$$y_i = g(t_i,\vartheta) + \varepsilon_i, \quad i = 1 \ldots n . \tag{1}$$

Here is $\vartheta \in \Omega \subseteq \mathbb{R}^p$, $t \in I$ and ε_i is some error term. Most work has been done for the case of

$$\vartheta \longrightarrow g(t_i,\vartheta)$$

being linear, say

$$g(t_i,\vartheta) = X(t_i)\vartheta$$

and the D-criterion

$$f(t) = 1/\det\left(\sum_{i=1}^{n} X^T(t_i)X(t_i)\right),$$

where $t = (t_1 \ldots t_n)$.

Even for this appearently simple case severe difficulties occur in the minimization. As pointed out by Mitchell [1], the main problem is to get trapped by one of the numerously existing local minima of the objective function f. In [2] Cook and Nachtsheim compared several algorithms for the linear case and the D-criterion. Most of them consist in some exchanging strategy which adds and subtracts points to the current design using properties of the determinant and the linearity of g. None of the considered algorithms was superior to all others, especially none of them did always find the best known design.

For nonlinear models an analog of the D-criterion was introduced by Chernoff [3], first analytical and numerical results on optimal designs were obtained by Box and Lucas [4]. A further sophistication of the D-criterion basing on curvature measures of the solution locus as introduced by Bates and Watts [5] was given by Hamilton and Watts [6], but little attention was paid to the problem of minimizing the computed objective function.

Rasch, Rudolph and Schimke published in [7] their algorithm OPREG together with applications to a special growth model and a variety of optimality criterions derived from the asymptotic co-variance matrix. OPREG proved to be very succesful in finding the (probably) global minimizer in small sized problems ($n \leq 15$, $p = 3$), but it requires excessively high computation time even for medium problems.

Our present work is mainly aimed to provide a better understanding of the general problems with searching for optimal experimental designs, independent of the underlying statistics, and to propose a specific tool for overcoming local minima.

It turns out (see Sec. 2) that for objective functions in experimental design typical (local) minimizers exhibit only few different components. Basing on this fact we define in Sec. 3 a map U from the domain $C = I^n$ of the objective function f into its power set. This map is used lateron in order to enlarge the neighbourhoods of points, especially it selects candidates for a restart if the minimization reaches a local minimum.

The idea is as follows: we start with some minimization algorithm which terminates if we find a $t \in C$ with

$$\exists \ O - \text{open},\ t \in O \ \ \forall\, \tilde{t} \in O \quad f(t) \leq f(\tilde{t}). \tag{2}$$

But now, we strengthen the condition for termination demanding

$$\exists \ O - \text{open},\ t \in O \ \ \forall\, \tilde{t} \in O \cup U(t) \quad f(t) \leq f(\tilde{t}). \tag{3}$$

If (2) is valid but (3) is not then we swich from our initial minimization algorithm to a supplementary algorithm for the minimization of $f_{|U(t)}$ and restart at the minimizer. Hence the risk of getting trapped in a local minimizer decreases at the cost of a more expensive search.

It turned out that for our map $U\colon C \longrightarrow 2^C$ the above algorithm renders the search for optimal experimental designs very economic and removes the problem of finding "wrong" minima to a practically sufficient degree. Furthermore, for certain problems it finds even the global minimizer (see Sec. 5).

2. The objective function for experimental designs

For most applications the optimality criterion is based on the so called asymptotic co-variance matrix. Let $t \in C = I^n$ be an experimental design, then

$$F(t) = \sum_{i=1}^{n} [\mathrm{grad}_{\vartheta}\, g(t_i, \vartheta_o)]^T \,\mathrm{grad}_{\vartheta}\, g(t_i, \vartheta_o) \in \mathbb{R}^{P \times P} \tag{4}$$

with a given a priori information ϑ_o is called the Fisher information matrix of t and its inverse

$$V(t) = F^{-1}(t), \tag{5}$$

if existing, is the asymptotic co-variance matrix.
Commonly used criterions are $\det(V)$, $\mathrm{tr}(V)$, and V_{ii}, $(i = 1 \ldots p)$, (cf. [8]).
Hamilton and Watts modify $\det(V)$ by adding a term containing second derivatives of the model function at t and ϑ_o (cf. [6]).
Note that $t \longrightarrow F(t)$ is fully symmetric, hence all criterions basing on $V(t)$ are so, the criterion from [6] is fully symmetric as well. Further, smooth model functions yield smooth objective functions. Denoting $T(t) = \{t_1, \ldots, t_n\}$, we see that $F(t)$ becomes singular if $|T(t)| < p$. For such designs the objective function becomes infinite. We may assume the model function to yield finite values of the criterion otherwise.
Hence, in general, we deal with the following problem

$$f(t) = \mathrm{Min} \quad ! \tag{6}$$

with a smooth $f: C \longrightarrow \overline{\mathbb{R}^+}$ satisfying

$$f(t) = f(\pi(t)) \tag{7}$$

for any permutation $\pi \in S(n)$ and

$$f(t) = +\infty \quad \text{iff} \quad |T(t)| < p \tag{8}$$

with some p, $1 < p \leq n$.

Of course, f cannot be convex on C since it then would take its minimum on the diagonal

$$D = \{ (\tau, \ldots, \tau) \mid \tau \in I \} \subset C$$

which is obviously not true. Further, f is in general not convex even on $S \subset C$,

$$S = \{ t \in C \mid t_1 \leq t_2 \leq \ldots \leq t_n \}, \tag{9}$$

consequently the possibility of restricting f to S is not to much advantageous.

For the following assertion we assume I to be the interval $[x_1, x_u] \subset \mathbb{R}$ and $p = 3$. Further let D_{n_1,n_2}, $n_1 + n_2 < n$, be the hyperdiagonal

$$\{ t \in C \mid t_1 = \ldots = t_{n_1} = x_1, \; t_{n_1+1} = \ldots = t_{n_1+n_2} = x_u,$$
$$t_{n_1+n_2+1} = \ldots = t_n \}.$$

<u>Assertion 1:</u> For $p = 3$, $I = [x_1, x_u] \subset \mathbb{R}$ and each pair of positive integers n_1, n_2 with $n_1 + n_2 < n$ there exists a critical point of f on D_{n_1,n_2}.

<u>Proof:</u> If we introduce in the actual tangent space to C at some inner point of D_{n_1,n_2} the basis

$$\begin{aligned}
e_1 &= (0,0,\ldots 0,\underset{n_1+n_2}{1}, 1, 1, \ldots, 1) \\
e_2 &= (0,0,\ldots 0,1,-1, 0, \ldots, 0) \\
e_3 &= (0,0,\ldots 0,0, 1,-1, 0, \ldots, 0) \\
\\
e_{n-n_1-n_2} &= (0,0,\ldots 0,0, 0, \ldots 0, 1,-1)
\end{aligned}$$

then $df/de_i = 0$, $i = 2, \ldots, n-n_1-n_2$ due to the symmetry condition (7). Further we have

$$\lim_{\tau \to x_1(x_u)} f(x_1 \ldots x_1, x_u \ldots x_u, \tau, \tau, \ldots, \tau) = \infty$$

hence due to Rolle's theorem df/de_1 vanishes at some point

$$t = (x_1 \ldots x_1, x_u \ldots x_u, x_m \ldots x_m) \in D_{n_1,n_2}. \tag{10}$$

We infer that the projection of grad f(t) onto the tangent space to C at t vanishes, which completes the proof. ■

<u>Remark:</u> Note that the mentioned tangent space is constant along D_{n_1,n_2}.

<u>Assertion 2:</u> Let $\hat{n} < n$ and $t = (t_1 \ldots t_{\hat{n}}, 0, 0 \ldots 0)$ and

$$B_{\hat{n},\hat{t}} = \{ t \mid t = t + (\underbrace{0,0,\ldots 0}_{n}, \tau, \tau \ldots \tau), \tau \in I \}.$$

Assume a) f is convex in some neighbourhood V of $B_{\hat{n},\hat{t}}$ and

b) V contains a local minimizer of f.

Then $B_{\hat{n},\hat{t}}$ contains a local minimzer of f.

The proof is a direct consequence of (7) and Jensens's inequality, here no assumptions about I and p are needed.

There is obviously little practical use to be made from the above assertions, in Assertion 1 it remains open whether the gradient points inside or outside C, while the assumptions of Assertion 2 are practically not to be verified. But nevertheless, they give some insight why the minimization of design criterions yields usually minimizers with a minimum number (i.e., equal to the number of unknown parameters p) of different components (with exceptions!). The situation studied by Rasch, Rudolph and Schimke [7] meets the assumptions of Assertion 1, and indeed most of the results reported by them are critical points just of the form (10). Moreover, for the D-criterion and a certain class of growth functions, the existence of local minimizers of the type (10) can be proven (cf. [9]). On the other hand, there exist special cases where the assumptions of Assertion 1 are fulfilled, but not both x_1 and x_u belong to the optimal design, e.g. for the growth function (14) below. Further "experimental" results and hypotheses on the properties of optimal experimental designs based on numerical experience are contained in Rasch a.o. [7].

The above results together with some experience suggest that for large n and small p the terminal values of minimization algorithms are more economically represented by ordered tables

$$t \sim \left(\frac{x_1 \; \dots \; x_m}{n_1 \; \dots \; n_m} \right), \; x_i \in I, \; n_i \in \mathbb{N}^+ \tag{11}$$

containing only the different components of t and their numbers of occurence. Now, it is quite natural to treat two such tables as neighbouring if they have a common first row and small differences in the second. This motivates the considerations of the next section.

3. The mapping U

Let us consider two families of sets

$$U_\varepsilon^1(t) = \{ \tilde{t} \in C \mid T(\tilde{t}) \subseteq T(t) \wedge \| n_{\tilde{t}} - n_t \| \leq \varepsilon \},$$

$$U_\varepsilon^2(t) = \{ \tilde{t} \in C \mid T(\tilde{t}) \subseteq T(t) \wedge ||| n_{\tilde{t}} - n_t ||| \leq \varepsilon \},$$

with $n_t(\tau) = |\{ i : t_i = \tau \}|$,

$$\| n_{\tilde{t}} - n_t \| = 1/2 \sum_{\tau \in T(t)} | n_{\tilde{t}}(\tau) - n_t(\tau) |$$

and

$$||| n_{\tilde{t}} - n_t ||| = \max_{\tau \in T(t)} | n_{\tilde{t}}(\tau) - n_t(\tau) | .$$

It is easy to see :

a) For $\varepsilon < 1$ is $U_\varepsilon^1(t) = U_\varepsilon^2(t) = \{t\}$.

b) We have $|U_1^1(t)| = |T(t)| \; (\; |T(t)| - 1 \;) + 1,$

$$|U_1^2(t)| \leq 3^{|T(t)|} .$$

c) For each ε there exists a bound b_E independent of t such that

$$|U_\varepsilon^1(t)| \leq |U_\varepsilon^2(t)| < b_\varepsilon \; .$$

d) $U_\varepsilon^1(t) \subseteq U_\varepsilon^2(t)$.

We interpret the sets $U^i(t)$ as some "neighbourhoods" of the point t. Nevertheless, it should be pointed out that from the topological point of view the families $U^i_\varepsilon(t)$, $\varepsilon \geq 1$, are no bases of neighbourhoods, further the topology generated by their sum is the discrete one.

Now we define

Definition: A point $t \in C$ is called a semi-local minimizer (SLM) of f on C if

$$\exists\, \varepsilon_1 > 0,\ \varepsilon_2 \geq 1\ \ \forall\, \tilde{t} \in C \quad \tilde{t} \in B_{t,\varepsilon_1} \cup U^i_{\varepsilon_2}(t) \Longrightarrow f(\tilde{t}) \geq f(t),$$

where B_{t,ε_1} is the ball with radius ε_1 and center at t and $i = 1$ or $i = 2$.

Of course, each SLM is a local minimizer, but not contrarily. In fact, the following assertion is true:

Assertion 3: A point $t \in C$ is a SLM of f on C iff

a) it is a local minimizer and
b) $f(t) = \min f_{|U^i_1(t)}$.

From now on we fix i and set $\varepsilon = 1$, defining $U(t) = U^i_1(t)$. Since $U(t)$ is for each $t \in C$ a rather small finite set the above assertion suggests a very simple algorithm for the search of SLM which will be discussed in the next section.

4. The search algorithm

Basing on Assertion 3 we propose the following procedure

1. $k = 0$,
2. choose a starting point t^o,
3. calculate an (approximative) local minimizer t^{k+1}, using a descent method starting at t,
4. calculate $\min f_{|U(t^{k+1})}$ $(=: f(t^{k+2}))$,
5. if $t^{k+2} = t^k$ then stop, else $k = k+2$ and go to 3.

Some details of steps 3 and 4 may be of great importance for the overall performance of the algorithm :

a) Since for $I = [x_l, x_u]$ the domain C is a cube we deal here with so called box constraints. It is either possible to transform the problem into an unconstraint minimization problem (cf. [10]) or to use some backtracking and projection method in step 3. In our implementation we used the latter approach, basing on a secant method of Broyden type (cf.[11]). In each step of the descent method the gradient of the objective function as well as the search direction are reduced, i.e. projected onto the actual tangent space of C.

b) The first application of 3. provides usually a remarkable reduction of $|T(t)|$, while the minimizer very often obeys $|T(t)| = p$. Consequently, with respect to the cardinality of $U(t)$, we recommend to use in 4. the norm $\|\cdot\|$ until $|T(t)|$ becomes small enough and then to swich to the more expensive norm $|\!|\!|\cdot|\!|\!|$.

c) In the representation of designs as well as in the calculation of derivatives and updating of quasi-Hesse matrices the symmetry (7) should be used. This provides huge savings in the final stage of minimization. On the other hand, attention must be paid to the numerical method of calculating $\det(V)$, $\mathrm{tr}(V)$ or V_{ii}, respectively, for the vectors $\mathrm{grad}_\vartheta g(t_i, \vartheta_o)$, $i = 1,\dots,n$, are usually almost linearly dependent.

It is quite obvious that the above procedure terminates always at a SLM, provided the number of local minimizers is finite. Of course, step 3 is possible to be carried out only approximately. Step 4 may be substituted by 4a :

4a. search for t^{k+2} such that $t^{k+2} = \min f_{|U(t^{k+2})}$.

Again, it is possible to swich between 4. and 4a. in dependence on the efficiency at the last taken step.

5. A convergence result

For the general case of the optimization problem (6)-(8) the convergence of the above algorithm to a global minimizer is not ensured. Indeed, there are examples with the norm $\|\cdot\|$ and $\varepsilon = 1$, where the global minimizer was not found for certain starting points. Nevertheless, for the special case of the criterion $f(t) = \det(V(t))$ and $p = 3$ we are able to prove the following assertion, which applies to several models of practical interest.

<u>Assertion 4:</u> Let $p = 3$ and the graph of the function

$$z : [x_l, x_u] \longrightarrow \mathbb{R}^3,$$

$$z(\tau) = \text{grad}_\vartheta g(\tau, \vartheta_o)$$

be a regular curve contained in some plane π with $O \notin \pi$. In π there exists a basis $\{b_1, b_2\}$ such that the b_2-component of z is a convex function of the b_1-component. Then the algorithm from the previous section solves the problem

$$f(t) = \det(V(t)) = \text{Min!} \quad , \; t \in C_p$$

for each norm $\|\cdot\|$ or $|||\cdot|||$ each $\varepsilon \geqslant 1$ and each starting point $t_o \in C_p$ in not more then n steps.
Here is $C_p = \{ t \in C : |T(t)| = p \}$.

<u>Proof:</u> In [9] it is proven that under the assumptions of the assertion the objective function obeys

$$f(t) = \varphi(x_t) \, \psi(n_t)$$

with x_t being the first row of the representation (11) and a function ψ depending only on the image of n_t. Further, φ takes a unique minimum on I^P and ψ is a convex symmetric function of the nonvanishing values of n. Hence steps 3 and 4 work independently.
Consequently, the algorithm is finite, i.e., it finds a SLM. Due to the above properties of f, φ and ψ the conditions for a SML and for a global minimizer are the same. The bound for the number of steps follows from a simple examination of the worst case $\varepsilon = 1$. ■

The assumptions of Assertion 4 are fulfilled for the models (12) and (14) of the following section. However, we applied the algorithm succesfully in several other situations, too.

6. Numerical experience

We performed numerical experiments for $n = p$ (= 3 or 4), $n = 10$ and $n = 100$, using $I = [0,65]$ and objective functions calculated from the growth models

$$y = \vartheta_1 + \vartheta_2 \exp(\vartheta_3 x) \ , \ x \in I, \quad \vartheta = (\vartheta_1, \vartheta_2, \vartheta_3) \tag{12}$$

with $\vartheta_3 = -0.05$ and $\vartheta_3 = -0.03$,

$$y = (\vartheta_1 + \vartheta_2 \exp(\vartheta_3 x))^3, \quad x \in I, \quad \vartheta = (\vartheta_1, \vartheta_2, \vartheta_3) \tag{13}$$

with $\vartheta_1 = 5$, $\vartheta_2 = -5$, $\vartheta_3 = -0.05$,

$$y = \vartheta_1 + \vartheta_2 \exp(\exp(\vartheta_3 x)), \quad x \in I, \ \vartheta = (\vartheta_1, \vartheta_2, \vartheta_3) \tag{14}$$

with $\vartheta_1 = 100$, $\vartheta_2 = -3$, $\vartheta_3 = 0.1$,

$$y = \vartheta_1 + \vartheta_2 \tanh(\vartheta_3 x + \vartheta_4), \ x \in I \ , \ \vartheta = (\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4) \tag{15}$$

with $\vartheta_2 = 50$, $\vartheta_3 = -3$, $\vartheta_4 = 0.1$.

For the function (12) we chose the V-criterion, for functions (13)-(15) we chose the D-criterion.
Further experiments similar to that from [2] are planned by Rasch et al. in forthcoming papers.
In order to check the danger of getting trapped in SLM's being not a global minimizer we started our algorithm from different starting points choosen to posses $m := |T(t_o)|$ different components with an uniform distribution over I. The numbers of equal components were choosen to be "almost equal" (cf.[8]). We run the algorithm with $m = 3$, $m = 4$, $m = 6$, $m = 8$ and $m = 10$. The result was for all m a design with the same value of the criterion. $|T(t)|$ was p with only two exceptions: $|T(t)|$ was $p+1$ for the growth model (13) with $n = 10$, $m = 6$ and for the growth model (15) with $n = 10$, $m = 8$ and $m = 10$. There was only a small difference between two components of t, but the value of the criterion was also the same like for the other m, where $|T(t)|$ was equal to p. Our experience suggests that for small values of n the starting point t with $t_i = x_l + (x_u - x_l)(i-1)/(n-1)$ is most succesful, while for larger n the choice $m = p$ seems to be

sufficient. Note that usual implementations of step 3 cannot increase $|T(t)|$ with respect to the symmetry of f.

7. Concluding remarks

The presented concept of SLM released in an obvious manner the danger of finding local minimizers that are not global. On the other hand, the problem of constructing mappings U that ensure the SLM to be a global minimizer and result in reasonable algorithms, remains open in general. Here we mean by reasonable that $U(t)$ should be much smaller than C ($U(t) = C$ would of course make each SLM a global minimizer). It seems that the choice of a useful mapping U requires always some deeper understanding of the global behaviour of the objective function under consideration as provided here by Assertions 1 and 2.

References:

1. Mitchell, T.: An algorithm for the construction of "D-optimal" experimental designs. Technometrics 16, 2, 203-210 (1974)

2. Cook, R.D., and Nachtsheim, Ch.J.: A comparison of algorithms for constructing exact D-optimal designs. Technometrics 22, 3, 315-324 (1980)

3. Chernoff, H.: Locally optimal designs for estimating parameters. Ann. Math. Statist. 24, 586-602 (1953)

4. Box, G.E.P., and Lucas, H. L.: Design of experiments in non-linear estimations. Biometrica 49, 77-99 (1959)

5. Bates, D.M., and Watts, D. G.: Relative curvature measures of nonlinearity. J. Roy. Statist. Soc. Ser. B 42, 1, 1-25 (1980)

6. Hamilton, D.C., and Watts, D.G.: Quadratic design criterion for precise estimation in non-linear regression models. Technometrics 27, 3, 241-250 (1985)

7. Rasch, D., Rudolph, P.R., und Schimke, E.: Optimale Versuchspläne in der nichtlinearen Regression. In: Rasch, D. (Hrgb.): Robustheit IV. Probleme der angew. Statist. 13, 93-121 (1985)

8. Duchrau, P., Frischmuth, K., Rasch, D., and Schimke, E.: Optimum experimental designs in growth curve analysis. Rostock. Math. Kolloq. 30, 93-104 (1986)

9. Frischmuth, K.: On locally D-optimal experimental designs. Rostock Math. Kolloq. 36, 51-64 (1989)

10. Sisser, F.S.: Elimination of bounds in optimization by transforming variables. Math. Programming 20, 110-121 (1981)

11. Dennis, J.E., Jr., and Schnabel, R.B.: Numerical Methods for Unconstrained Optimization and Non-linear Equations. Englewood Cliffs, New Jersey 1983

12. Frischmuth, K.: An analytical result in experimental design. Rostock. Math. Kolloq. 31, 95-101 (1987)

eingegangen: 19. 7. 1989

Authors

Dr. Kurt Frischmuth
W.-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
Rostock
DDR-2500

Dipl.-Math. Petra Duchrau
W.-Pieck-Universität Rostock
Sektion Informatik
Universitätsplatz 1
Rostock
DDR-2500

Rostock. Math. Kolloq. 40, 95-96 (1990) 60G15
62M05

Carsten Stade

Gaußsche Markovprozesse höherer Ordnung als Zeitreihenmodelle und ihre numerische Behandlung - mit Anwendung auf Daten aus einem Wachstumsversuch mit Tieren

Autorreferat zur Dissertation A

Die Dissertation behandelt eine spezielle Klasse instationärer Zeitreihen, die der Gaußschen Markovprozesse (GMP) höherer Ordnung. Sie werden auf der Grundlage der Charakterisierung ihrer Kovarianzfunktion und der bedingten Verteilungen beschrieben. Eine wesentliche Rolle spielt dabei die Tatsache, daß die Inverse der Kovarianzmatrix eines GMP der Ordnung τ eine Bandmatrix der Bandbreite $(2\tau + 1)$ ist. Es werden notwendige und hinreichende Bedingungen dafür angegeben, wann ein vorliegender Gaußscher Prozeß auch ein Markovprozeß der Ordnung τ ist.

Die stationären GMP lassen sich in die allgemeine Theorie stationärer stochastischer Prozesse einordnen. Dabei wird der Zusammenhang zwischen der bekannten Stationaritätsbedingung für autoregressive Prozesse, die auf Box und Jenkins zurückgeht, und den Voraussetzungen über die Kovarianzmatrix eines GMP aufgezeigt. Die stationären GMP (als der einfachste Fall Gaußscher Markovprozesse) dienen wiederholt als Beispiel zur Veranschaulichung erzielter Resultate.

Ausgehend von der Tatsache, daß der bedingte Erwartungswert die beste Prognose einer unbekannten Prozeßrealisierung liefert, behandelt ein Kapitel die Auswirkungen der Markovbedingung auf bedingte Momente mit unterschiedlichen Realisierungen in den Bedingungen und somit die Konsequenzen für Prognose und Interpolation. Insbesondere erfolgt eine Diskussion der Vorhersage zeitlich weit entfernt liegender Realisierungen (Mehrschrittprognose).

Überlegungen zu Transformationen, die aus einem GMP einen anderen derselben Ordnung erzeugen, führen zu einfachen notwendigen und hinreichenden Bedingungen, unter denen ein GMP in einen stationären GMP übergeführt werden kann.

Unter dem Gesichtspunkt der Beschreibung Gaußscher Markovprozesse durch charakteristische Größen werden numerische Algorithmen zu deren Berechnung vorgeschlagen. Die Demonstration der Algo-

rithmen erfolgt am Beispiel mehrerer Simulationsexperimente, die auf einem EC 1056 durchgeführt wurden.

Für Datenmaterial aus einem Wachstumsversuch mit Tieren werden Möglichkeiten zur Beschreibung durch das Modell eines GMP untersucht. Die zugehörigen Momentschätzungen sind in Abbildungen veranschaulicht, die Parameterschätzungen können Tabellen entnommen werden. Eine weiter Tabelle gibt einen Überblick über erzielte Prognoseergebnisse. Ein abschließendes Beispiel behandelt das Verhalten der vorgestellten Algorithmen an simulierten Gauß-Prozessen mit einer den Daten ähnlichen Kovarianzstruktur.

eingereicht: 26. 8. 1987 verteidigt: 6. 5. 1988

Gutachter:

Prof. Dr. rer. nat. habil. L. Berg (Rostock)
Prof. Dr. sc. nat. F. Liese (Rostock)
Prof. Dr. rer. nat. habil. P.H. Müller (Dresden)

Verfasser:

Dr. rer. nat. Carsten Stade
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
Rostock
DDR-2500

Hinweise für Autoren

Manuskripte (in deutscher, ggf. auch in russischer oder englischer Sprache) bitten wir, an die Schriftleitung zu schicken. Die gesamte Arbeit ist linksbündig zu schreiben. Eine Ausnahme hiervon bilden hervorzuhebende Formeln und das Literaturverzeichnis. Der Kopf der Arbeit soll folgende Form haben: Rostock. Math. Kolloq./ Leerzeile/ Vorname Name/ Leerzeile/ Titel der Arbeit/ 1 Zeilenumschaltung/ Unterstreichung/ Leerzeile. Der Text der Arbeit ist einsinhalbzeilig (= 3 Zeilenumschaltungen) zu schreiben mit maximal 63 Anschlägen je Zeile und maximal 37 Zeilen je Seite. Zwischenüberschriften sind wie folgt einzuordnen: 6 Zeilenumschaltungen/ Zwischenüberschrift/ Unterstreichung (ohne Zeilenumschaltung)/ 5 Zeilenumschaltungen. Hervorhebungen sind durch Unterstreichen und Sperren möglich. Ankündigungen wie Satz, Definition, Bemerkung, Beweis u. a. sind zu unterstreichen und mit einem Doppelpunkt abzuschließen. Vor und nach Sätzen, Definitionen u. ä. ist ein Zeilenabstand von 5 Umschaltungen zu lassen. Fußnoten sind möglichst zu vermeiden. Sollte doch davon Gebrauch gemacht werden, so sind sie durch eine hochgestellte Ziffer im Text zu kennzeichnen und innerhalb des oben angegebenen Satzspiegels unten auf der gleichen Seite anzugeben. Formeln und Bezeichnungen sollen möglichst mit der Schreibmaschine zu schreiben sein. Hervorzuhebende Formeln sind drei Leerzeichen einzurücken und mit 6 Umschaltungen zum übrigen Text zu schreiben. Formelzähler sollen am rechten Rand stehen. Der Platz für Abbildungen ist beim Schreiben auszusparen; die Abbildungen selbst sind in der dem ausgesparten Platz entsprechenden Größe gesondert nach TGL-Vorschrift auf Transparentpapier beizufügen. Der zugehörige Begleittext ist im Manuskript mitzuschreiben. Sein Abstand nach unten beträgt 5 Umschaltungen. Literaturzitate im Text sind durch laufende Nummern in Schrägstrichen (vgl. /8/, /9/ und /10/) zu kennzeichnen und am Schluß der Arbeit unter der Zwischenüberschrift Literatur zusammenzustellen.

Beispiele: (Zeitschriftenabkürzungen nach Math. Reviews)

/8/ Zariski, O., and Samuel, P.: Commutative Algebra. Priceton 1958

/9/ Steinitz, E.: Algebraische Theorie der Körper. J. Reine Angew. Math. 137, 167 - 309 (1920)

/10/ Gnedenko, B. W.: Über die Arbeiten von C. F. Gauß zur Wahrscheinlichkeitsrechnung. In: Reichard, H. (Ed.): C. F. Gauß, Gedenkband anläßlich des 100. Todestages. S. 193 - 204, Leipzig 1967

Die Angaben sollen in Originalsprache erfolgen; bei kyrillischen Buchstaben soll die bibliothekarische Transkription (Duden) verwendet werden.

Am Ende der Arbeit stehen folgende Angaben zum Autor und zur Arbeit: eingegangen: Datum/ Leerzeile/ Anschrift des Verfassers:/ Titel Initialen der Vornamen Name/ Institution/ Struktureinheit/ Straße Hausnummer/ Land Postleitzahl Ort.

Der Autor wird gebeten, eine Korrektur des Durchschlags vom Offsetmanuskript zu lesen und dabei die mathematischen Symbole einzutragen. Ferner sollte er 1 - 2 Klassifizierungsnummern (entsprechend der "1980 Mathematics Subject Classification" der Math. Reviews) zur inhaltlichen Einordnung seiner Arbeit angeben.

Schriftenreihen der Wilhelm-Pieck-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/Semesterbericht der Sektion Lateinamerikawissenschaften	ISSN 0458-7944
– Erziehungswissenschaftliche Beiträge	ISSN 0138-2373
– Beiträge zur Geschichte der Wilhelm-Pieck-Universität Rostock	ISSN 0232-539X
– Beiträge zur Geschichte der FDJ	ISSN 0233-0830
– Probleme der Agrargeschichte des Feudalismus und des Kapitalismus	ISSN 0233-0636
– Rostocker Beiträge zur Hoch- und Fachschulpädagogik	ISSN 0233-0539
– Rostocker Informatik-Berichte	ISSN 0233-0784
– Studien zur Geschichte der deutsch-polnischen Beziehungen	ISSN 0233-0687
– Rostocker Forschungen zur Sprach- und Literaturwissenschaft	ISSN 0233-0644
– Rostocker Universitätsreden	
– Migrationsforschung	ISSN 0863-1735
– Manuskripte zur Rostocker Universitätsgeschichte	ISSN 0863-1727

Bezugsmöglichkeiten

- Bestellungen aus der DDR über die Wilhelm-Pieck-Universität Rostock, Abt. Wissenschaftspublizistik, Vogelsang 13/14, Rostock, DDR-2500.
- Bestellungen aus dem Ausland über die Firma Buchexport, Volkseigener Außenhandelsbetrieb der DDR, Leninstr. 16, DDR-7010.

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