

ISSN 0138-3248

ROSTOCKER

MATHEMATISCHES KOLLOQUIUM

Heft 44

UNIVERSITÄT



ROSTOCK

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

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

#### **Bezugsmöglichkeiten**

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

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

# ROSTOCKER MATHEMATISCHES KOLLOQUIUM

## Heft 44

---

HANS E. PORST	<i>Zur Struktur freier topologischer Gruppen</i>	5
FRITZ v. HAESLER	<i>Fraktale Geometrie und Juliamengen</i>	21
ULRICH KRAUSE	<i>Über positive diskrete dynamische Systeme</i>	29
GERHARD BECKER	<i>Die Bedeutung von Analogien für das Lehren und Lernen von Mathematik</i>	39
INGO KÖLBL	<i>Stochastik mit 15-jährigen Schülern – Ergebnisse eines Unterrichtsversuchs</i>	53
WOLF-DIETER RICHTER	<i>Eine geometrische Methode in der Stochastik</i>	63
FRIEDRICH LIESE	<i>Aspekte der Statistik zufälliger Prozesse mit unabhängigen Zuwächsen</i>	73
GERHARD OSIUS	<i>Aspekte statistischer Modellierung und Auswertung</i>	81
MANFRED TASCHE	<i>Konstruktion schneller Algorithmen für diskrete Fouriertransformationen</i>	99

---

UNIVERSITÄT ROSTOCK  
FACHBEREICH MATHEMATIK  
1991

**Wissenschaftliche Leitung:** Prof. Dr. Hans-Wolfgang Stolle  
(Sprecher des Fachbereichs Mathematik)  
Dr. Werner Plischke  
**Redaktionelle Bearbeitung:** Dr. Werner Plischke  
**Herstellung der Druckvorlage:** Dr. Andreas Straßburg

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

**Redaktionsschluß:** 20. Juni 1991

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

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

**Zitat-Kurztitel:** Rostock. Math. Kolloq. (1991) 44

---

Universität Rostock  
Abteilung Wissenschaftspublizistik  
Vogelsang 13/14, O-2500 Rostock, Bundesrepublik Deutschland  
Telefon 369577  
Druck: Ostsee-Zeitung, Verlag und Druck GmbH Rostock, BT Ribnitz  
01000

## Vorwort

In diesem Heft des ROSTOCKER MATHEMATISCHEN KOLLOQUIUMS ist eine Reihe schriftlicher Fassungen von Vorträgen Bremer und Rostocker Mathematiker zusammengestellt, die auf dem

1. Bremen-Rostocker Mathematischen Kolloquium, Rostock, 27. und 28. April 1990 sowie auf dem
2. Bremen-Rostocker Mathematischen Kolloquium, Bremen, 29. und 30. Juni 1990 gehalten wurden.

Es war die Absicht dieser Kolloquien, die Kollegen in Bremen und Rostock mit typischen Arbeitsbereichen des jeweils anderen Fachbereichs bekannt zu machen. Insofern zielen die Beiträge in diesem Heft nicht unbedingt auf aktuellste Ergebnisse als vielmehr auf die Illustration der jeweiligen Arbeitsgebiete.

Im Ergebnis dieser Kolloquien entwickelten sich eine Reihe nützlicher wissenschaftlicher Kontakte sowie ein Austausch sehr vielfältiger Art auf dem Gebiet der Lehre.

Es ist geplant, in Auswertung der bisherigen Erfahrungen weitere Veranstaltungen spezieller Themenstellung durchzuführen. Insbesondere der Fachbereich Mathematik der Universität Rostock verspricht sich davon Impulse für die Bewältigung der sich abzeichnenden neuen Aufgaben in Lehre und Forschung.

Wir danken den Universitäten Bremen und Rostock für die Unterstützung bei der Durchführung der gemeinsamen Kolloquien.

März 1991

H.-E. Porst, Bremen  
G. Wildenhain, Rostock



HANS-E. PORST

## Zur Struktur freier topologischer Gruppen

---

Am Beispiel der freien topologischen Gruppen soll die wechselseitige Beeinflussung von konkreten mathematischen Fragen und allgemeinen kategoriellen Methoden exemplarisch aufgezeigt werden. Während die Geschichte der Existenzbeweise freier topologischer Gruppen die Entwicklung gewisser kategorieller Konzepte beeinflusst hat, kann kategorielles Denken andererseits zur Erhellung der inneren Struktur dieser Objekte entscheidend beitragen.

### 1. Einleitung

#### 1.1 Freie topologische Gruppen und Kategorientheorie

Vor etwa einem halben Jahrhundert hat A.A. Markov erstmals die Existenz einer freien Hausdorff-Gruppe  $G(X, \xi)$  über einem gegebenen Tychonoff-Raum  $(X, \xi)$  bewiesen [11]. Er konstruierte auf der abstrakten freien Gruppe  $FX$  über der Trägermenge  $X$  des Raums  $(X, \xi)$  eine geeignete Gruppentopologie  $\tau$ , so daß in der Tat die Inklusion der Erzeugendenmenge  $X$  in  $FX$  zu einer stetigen Abbildung (sogar zu einer topologischen Einbettung)  $\eta_{(X, \xi)}: (X, \xi) \rightarrow (FX, \tau)$  wird und sich jede stetige Abbildung  $f$  von  $(X, \xi)$  in eine topologische Gruppe  $A$  zu einem stetigen Homomorphismus  $f^\sharp$  auf  $G(X, \xi) = (FX, \tau)$  fortsetzen läßt. Zwar ließ diese Konstruktion eine befriedigende Antwort auf Markov's rein topologische Fragestellung zu, insgesamt jedoch gab die sehr aufwendige Konstruktion wenig Einsicht in die Struktur des erhaltenen mathematischen Objekts.

Sehr bald wurden deshalb einfachere Beweise gefunden, von denen hier insbesondere die Samuel's [18] und Kakutani's [5] genannt seien. Diese waren — obwohl naturgemäß die kategorielle Sprache noch nicht benutzend — rein kategoriell in ihrer Natur und nahmen in bestimmtem Umfang schon den Beweis des später als *General Adjoint Functor Theorem (GAFT)* bekannt gewordenen Satzes vorweg.

Verfeinerte kategorielle Methoden führte dann Wyler mit seinem *Taut Lift Theorem* [22] in diesen Problemkreis ein, das im Gegensatz zum GAFT auch die innere Struktur des freien Objekts einer Untersuchung zugänglich macht.

Auf Nel [13] geht schließlich die Einsicht zurück, daß man gemischte topologisch - algebraische Strukturen besser über *kartesisch abgeschlossenen* Kategorien studiert. Wir verfolgen diese Idee, indem wir im Abschnitt 3 zunächst Gruppen über der kartesisch abgeschlossenen Kategorie **k-Top** der  $k$ -Räume studieren und die hier leicht erhältlichen Strukturaussagen dann — soweit möglich — auf topologische Gruppen übertragen.

## 1.2 Präliminarien

In dieser Note beschäftigen wir uns also mit dem Problem, ob der Funktor  $V$  in dem folgenden kommutativen Diagramm von Funktoren einen Linksadjungierten  $G: \mathbf{Top} \rightarrow \mathbf{TopGrp}$  hat, und wie dieser aussieht.

$$\begin{array}{ccc} \mathbf{TopGrp} & \xrightarrow{V} & \mathbf{Top} \\ S \downarrow & & \downarrow T \\ \mathbf{Grp} & \xrightarrow{U} & \mathbf{Set} \end{array}$$

Hier bezeichnen **Set** (bzw. **Grp**, **Top**, **TopGrp**) die Kategorien der Mengen (bzw. Gruppen, topologischen Räume, topologischen Gruppen) und Abbildungen (bzw. Homomorphismen, stetigen Abbildungen, stetigen Homomorphismen), während die angedeuteten Funktoren die üblichen Vergißfunktoren sind. Das Symbol  $V$  für den *Unterliegender Raum Funktor* benutzen wir auch für seine möglichen (Co-)Restriktionen wie z. B. in  $V: \mathbf{Top}_2\mathbf{Grp} \rightarrow \mathbf{Tych}$ , wo **Tych** die volle Unterkategorie von **Top** der Tychonoff-Räume bezeichnet (Wir setzen als bekannt voraus, daß eine Hausdorff'sche topologische Gruppe wegen ihrer Uniformisierbarkeit automatisch Tychonoff'sch ist (vgl. [6])). Der Index 2 an **Top** deutet wie üblich auf die Hausdorff-Eigenschaft hin.

Die folgenden vollen Unterkategorien von **Top** werden eine besondere Rolle als Basiskategorien für  $V$  spielen: **Creg**, die Kategorie der vollständig regulären Räume (ohne  $T_1$ ) und **FHaus**, die Kategorie der funktional Hausdorff'schen Räume; hier heißt ein Raum  $(X, \xi)$  *funktional Hausdorff'sch*, falls sich je zwei Punkte von  $(X, \xi)$  durch eine stetige reellwertige Abbildung trennen lassen. Es liegen dann offenbar folgende Einbettungen vor:

$$\mathbf{Tych} \subset \mathbf{FHaus} \subset \mathbf{Top}_2 \subset \mathbf{Top} \quad \text{und} \quad \mathbf{Tych} \subset \mathbf{Creg} \subset \mathbf{Top}.$$

Jede dieser Einbettungen ist reflexiv mit surjektiven Reflektionsabbildungen; die erste und letzte sind sogar von identischen Abbildungen getragen (vgl. [4, 1.3.5] und [1, 5.22]).

Die Kategorie **Grp** wird als Unterkategorie von  $\Gamma\text{-Alg}$ , der Kategorie der assoziativen  $\Gamma$ -Algebren, betrachtet;  $\Gamma$  bezeichnet hier den *Gruppentyp* bestehend aus einer nullstelligen, einer einstelligen und einer zweistelligen Operation (vgl. [10]).  $F: \mathbf{Set} \rightarrow \mathbf{Grp}$  ist der *Freie Gruppe Funktor* und  $\eta_M: M \rightarrow UFM$  die *Einbettung der Generatoren*; insbesondere sei erwähnt, daß  $\eta_M$  über die freie  $\Gamma$ -Algebra  $\Gamma M$  über der Menge  $M$  faktorisiert (vgl. [10, 1.2.5]). Für

eine Abbildung  $f: M \rightarrow G$  von einer Menge  $M$  in eine Gruppe  $G$  bezeichnet  $f^\#$  stets die homomorphe Fortsetzung auf  $FM$ .

Auf folgenden einfachen, aber grundlegenden Sachverhalt möchten wir besonders hinweisen:

**Lemma 1.1** Sei  $(f_i: G \rightarrow (H_i, \tau_i))_{i \in I}$  eine Familie von Gruppenhomomorphismen in topologische Gruppen  $(H_i, \tau_i)$ . Dann ist die initiale Topologie  $\tau_{\text{init}}$  auf  $G$  bezüglich dieser Familie eine Gruppentopologie.

Insbesondere ist der Funktor  $S: \mathbf{TopGrp} \rightarrow \mathbf{Grp}$  ein topologischer Funktor, und der Funktor  $V: \mathbf{TopGrp} \rightarrow \mathbf{Top}$  bewahrt Initialität (d.h.  $V$  angewandt auf eine  $S$ -initiale Quelle liefert eine  $T$ -initiale Quelle).

**Beweis.** Da die  $f_i$ 's Homomorphismen sind, liegen folgende kommutativen Diagramme vor:

$$\begin{array}{ccc} (G, \tau_{\text{init}}) & \xrightarrow{f_i} & (H_i, \tau_i) & & (G, \tau_{\text{init}})^2 & \xrightarrow{f_i^2} & (H_i, \tau_i)^2 \\ & & \downarrow i & & m \downarrow & & \downarrow m \\ (G, \tau_{\text{init}}) & \xrightarrow{f_i} & (H_i, \tau_i) & & (G, \tau_{\text{init}}) & \xrightarrow{f_i} & (H_i, \tau_i) \end{array}$$

wo  $i$  (bzw.  $m$ ) die Gruppeninversionen (bzw. Multiplikationen) bezeichnet. Die Stetigkeit der Operationen auf den Gruppen  $(H_i, \tau_i)$  zusammen mit der Initialität der Familie der  $f_i$  zeigt nun die Stetigkeit der Gruppenoperationen auf  $(G, \tau_{\text{init}})$ . ■

Als eine einfache Folgerung hieraus erzeugt jede Teilmenge  $A$  einer topologischen Gruppe  $(G, \tau)$  eine topologische Untergruppe von  $(G, \tau)$ , indem die von  $A$  erzeugte Untergruppe von  $G$  mit der Relativtopologie von  $(G, \tau)$  versehen wird.

Zur gewählten Notation sei noch bemerkt, daß wir einen topologischen Raum mit  $(X, \xi)$  bezeichnen, wobei  $X$  seine unterliegende Menge und  $\xi$  seine Topologie ist; entsprechend wird die topologische Gruppe mit der unterliegenden Gruppe  $G$  und der (Gruppen-)Topologie  $\tau$  mit  $(G, \tau)$  bezeichnet. Wir unterscheiden notationsmäßig nicht zwischen einer Gruppe und ihrer unterliegenden Menge (insbesondere kann also  $(G, \tau)$  auch der der topologischen Gruppe unterliegende Raum sein) und auch nicht zwischen einem Morphismus in einer konkreten Kategorie und seinem unterliegenden Morphismus in der Basiskategorie.

## 2. Existenzbeweise

Wir rekapitulieren in diesem Abschnitt kategoriell argumentierende Existenzbeweise für freie topologische Gruppen. Samuels Argument ist zweistufig: zunächst wird — in heutiger Terminologie — das Erfülltsein der sogenannten *Lösungsmengenbedingung* überprüft (2.1), um dann, fast wie im Beweis des GAFT, das freie Objekt zu konstruieren (2.2).

**Lemma 2.1** Für jeden topologischen Raum  $(X, \xi)$  gibt es nur eine Menge  $\mathcal{I}$  stetiger Abbildungen  $i: (X, \xi) \rightarrow VI$  mit topologischen Gruppen  $I$ , so daß es für jede stetige Abbildung  $f: (X, \xi) \rightarrow VA$  ein geeignetes  $i_f: (X, \xi) \rightarrow VI_f \in \mathcal{I}$  zusammen mit einem Isomorphismus  $\varphi: I_f \rightarrow A_f$  gibt derart, daß  $f = \varphi \circ i_f$  gilt.

**Beweis.** Der Beweis ist naheliegend: aus einer Lösungsmenge für die Menge  $X$  bezüglich  $U: \mathbf{Grp} \rightarrow \mathbf{Set}$  (vgl. [8]) erhält man eine Lösungsmenge für den Raum  $(X, \xi)$  bezüglich  $V$  aufgrund der Beobachtung, daß  $X$  nur eine Menge von Topologien tragen kann. ■

**Satz 2.2 (Samuel)** Der Funktor  $V: \mathbf{Top}_2\mathbf{Grp} \rightarrow \mathbf{Top}$  erfüllt die Voraussetzungen des GAFT und hat deshalb einen Linksadjungierten.

**Beweis.** Betrachte für einen Raum  $(X, \xi)$  die Klasse aller stetigen Abbildungen  $f: (X, \xi) \rightarrow VA$  von  $(X, \xi)$  in (die unterliegenden Räume) alle(r) möglichen topologischen Gruppen  $A$ . Für jedes  $f$  erzeugt das Bild  $f[X]$  eine topologische Untergruppe  $A_f$  der entsprechenden topologischen Gruppe  $A$ . Wähle nun entsprechend 2.1 einen Isomorphismus  $\varphi_f: A_f \rightarrow I_f$  mit  $I_f \in \mathcal{I}$  und forme das Produkt  $\prod_{I \in \mathcal{I}} I$  mit Projektionen  $\pi_I$ . Es resultiert (beachte, daß  $V$  Produkte bewahrt!) eine induzierte stetige Abbildung  $\Phi$  derart, daß das folgende Diagramm kommutiert:

$$\begin{array}{ccc} (X, \xi) & \xrightarrow{\Phi} & \prod_{I \in \mathcal{I}} I \\ f \downarrow & & \downarrow \pi_I \\ A_f & \xrightarrow{\varphi_f} & I_f \end{array}$$

Es ist nun leicht zu überprüfen, daß die Corestriktion  $\rho_{(X, \xi)}: (X, \xi) \rightarrow VG(X, \xi)$  von  $\Phi$  auf den unterliegenden Raum der topologischen Untergruppe von  $\prod_{I \in \mathcal{I}} I$ , die von  $\Phi[X]$  erzeugt wird,  $V$ -universell für  $(X, \xi)$  ist. ■

Im Gegensatz zu dieser direkten Konstruktion des universellen Objekts geht der im folgenden dargestellte Wyler'sche Ansatz von der Idee aus, die Adjunktion für die untere Zeile unseres kommutativen Quadrats zu einer Adjunktion für  $V$  anzuheben. Wir formulieren das Resultat als Spezialfall des wichtigeren Teils von *Wyler's Taut-Lift-Theorem* [22].

**Theorem 2.3 (Wyler)** *Da  $S: \mathbf{TopGrp} \rightarrow \mathbf{Grp}$  ein topologischer Funktor ist und der Funktor  $V: \mathbf{TopGrp} \rightarrow \mathbf{Top}$  Initialität bewahrt (1.1), gibt es zu jedem topologischen Raum  $(X, \xi)$  eine  $V$ -universelle stetige Abbildung  $\rho_{(X, \xi)}: (X, \xi) \rightarrow VG(X, \xi)$ , die eine Liftung der  $U$ -universellen Abbildung  $\eta_X: X \rightarrow UFX$  in dem Sinne ist, daß  $T\rho_{(X, \xi)} = \eta_X$  und  $SG(X, \xi) = FX$  gilt.*

Die Topologie von  $G(X, \xi)$  ist die initiale (Gruppen-)Topologie auf  $FX$  bezüglich der Familie

$$\{f^\sharp: FX \rightarrow G \mid G = S(G, \tau), (G, \tau) \in \mathbf{TopGrp}, f: (X, \xi) \rightarrow S(G, \tau)\}^1.$$

**Beweis.** Da  $S$  topologisch ist, existiert die im Theorem beschriebene initiale Gruppen topologie  $\tau_{(X, \xi)}$  auf  $FX$ ; da  $V$  Initialität bewahrt, ist sie auch initial in  $\mathbf{Top}$  bezüglich der Familie aller Abbildungen  $f^\sharp$ . Nun kann alles aus dem kommutativen Dreieck

$$\begin{array}{ccc} (X, \xi) & \xrightarrow{\eta_X} & (FX, \tau_{(X, \xi)}) \\ & \searrow f & \downarrow f^\sharp \\ & & (G, \tau) \end{array}$$

abgelesen werden. ■

Die folgende Verallgemeinerung von 2.2 erhält man, indem man zunächst die reguläre Faktorisierung der Familie der  $f^\sharp$ 's durchführt und dann die initiale Topologie bezüglich der sich ergebenden punktgetrennten Familie bildet. Auf diese Weise wird der Taut-Lift-Ansatz auch unter Annahme von Trennungsaxiomen anwendbar:

**Theorem 2.4 ([2, 21])** *Sei in dem kommutativen Quadrat von Funktoren*

$$\begin{array}{ccc} A & \xrightarrow{V} & B \\ S \downarrow & & \downarrow T \\ \mathbf{Grp} & \xrightarrow{U} & \mathbf{Set} \end{array}$$

*$S$  ein monotopologischer Funktor und bewahre  $V$  Initialität von punktgetrennten Familien; dann gibt es zu jedem  $B$ -Object  $B$  einen  $V$ -universellen Morphismus  $\rho_B: B \rightarrow VGB$ , der eine Liftung des Kompositums  $e \circ \eta_{TB}$  mit einer geeigneten Surjektion  $e$  ist. ■*

par

**Bemerkung.** Im Gegensatz zu 2.3 enthält 2.4 jedoch keine Information mehr über die algebraische Struktur des  $V$ -freien Objekts, es sei denn, man kann zeigen, daß die Familie

<sup>1</sup>Grob gesprochen ist dies die Klasse aller homomorphen Fortsetzungen von stetigen Abbildungen auf  $(X, \xi)$  mit Werten in irgendwelchen topologischen Gruppen  $(G, \tau)$ .

der  $f^{\#}$ 's punktstetig ist und deshalb die Surjektion  $e$  in 2.4 als Identität gewählt werden kann. Dies ist sicherlich der Fall, wenn  $FX$  eine  $\mathbf{A}$ -Struktur trägt (d.h.  $FX = SA$ ), so daß  $\eta_X$  ein  $\mathbf{B}$ -Morphismus ist (d.h.  $\eta_X = T(g: (X, \xi) \rightarrow VA)$ ), da dann  $\eta_X^{\#} = id_{FX}$ . Hieraus ergibt sich (vgl. Abschnitt 3.4) die Bedeutung eines schönen Resultats von Świrczkowski [20].

### 3. Die Struktur freier topologischer Gruppen

#### 3.1 Die primitive Topologie

Die Richtigkeit nachfolgenden Lemmas ist unmittelbar einzusehen. Wir formulieren es hier, da es der Ausgangspunkt einer für das folgende entscheidenden Begriffsbildung ist.

**Lemma 3.1** *Es sei die topologische Gruppe  $(G, \tau)$  die freie (Hausdorff'sche) topologische Gruppe über dem Raum  $(X, \xi)$  vermöge der  $V$ -universellen Abbildung  $\rho_{(X, \xi)}$ . Dann ist  $\tau$  die größte —d.h. feinste— (Hausdorff'sche) Gruppentopologie, bezüglich der die Abbildung  $\rho_{(X, \xi)}$  stetig ist. ■*

Für jede topologische Gruppe  $(G, \tau)$  sind nicht nur die Gruppenoperationen stetig, sondern auch alle Abbildungen von endlichen Potenzen von  $(G, \tau)$  nach  $(G, \tau)$ , die sich aus diesen ableiten lassen, wie z. B.  $(x, y, z) \mapsto xy^{-1}z$  oder  $(x, y, z, u) \mapsto z^{-1}yx^2u$ . Wir nennen diese Abbildungen *Termabbildungen* (für eine präzise Definition verweisen wir auf die Universelle Algebra) und bezeichnen mit  $\mathcal{T}_G$  die Menge aller dieser Termabbildungen für die Gruppe  $G$ . Um die Topologie der freien topologischen Gruppe  $G(X, \xi)$  über dem Raum  $(X, \xi)$  zu beschreiben, müssen wir gemäß 2.3 und 3.1 eine möglichst feine Topologie auf  $FX$  suchen, bezüglich der die Einschränkungen aller Termabbildungen aus  $\mathcal{T}_{FX}$  auf die entsprechenden Potenzen von  $(X, \xi)$  stetig sind. Diese Beobachtung führt unmittelbar auf den folgenden auf Mal'cev [9] zurückgehenden Begriff.

**Definition 3.2 (Primitive Topologie)** *Für einen topologischen Raum  $(X, \xi)$  wird die finale Topologie auf  $FX$  bezüglich der Einschränkungen aller Termabbildungen aus  $\mathcal{T}_{FX}$  auf die entsprechenden topologischen Potenzen des Raums  $(X, \xi)$  die primitive Topologie auf  $FX$  bezüglich  $(X, \xi)$  genannt. Wir bezeichnen sie mit  $\tau_{(X, \xi)}^p$ .*

Unglücklicherweise ist die primitive Topologie auf  $FX$  – wie bereits Mal'cev bemerkte – im allgemeinen keine Gruppentopologie; falls sie dies jedoch ist, ist  $(FX, \tau_{(X, \xi)}^p)$  offenbar die freie topologische Gruppe über  $(X, \xi)$ ; für diesen Fall hätten wir dann eine zufriedenstellend explizite Beschreibung der Topologie der freien topologischen Gruppe zur Hand (die

Beschreibung gemäß 2.3 ist weit davon entfernt, eine solche zu sein, da sie auf die gesamte Kategorie der betrachteten topologischen Gruppen zurückgreift, insbesondere also auf deren freie Objekte!).

Der Grund für diese unbefriedigende Situation liegt nun nicht im Begriff der primitiven Topologie, sondern an einem Manko der Kategorie **Top**, wie wir im folgenden aufzeigen wollen.

### 3.2 Freie $k$ -Gruppen

Anstelle der Kategorie **Top** werden wir nun mit der Kategorie **k-Top** der  $k$ -Räume arbeiten. **k-Top** kann definiert werden als die mono-coreflexive Hülle in **Top** aller kompakten Hausdorff-Räume.  $k$ -Räume sind also die Quotienten von Hausdorff'schen lokalkompakten Räumen. Für uns wird es ausreichen, mit folgenden kategoriellen Eigenschaften der Kategorie **k-Top** vertraut zu sein. Beweise sind zum Teil durch die Referenzen skizziert; für Details und für einen Beweis der letzten Aussage verweisen wir auf [7].

- **k-Top** ist eine volle konkret coreflexive Unterkategorie von **Top** (vgl. [1, 16.5 (1)]). Insbesondere folgt:
- **k-Top** ist eine topologische Kategorie, deren finale Strukturen auch final in **Top** sind, (vgl. [1, 21.30 – 21.35]).
- **k-Top** ist kartesisch abgeschlossen.

Als bi-coreflexive Unterkategorie von **Top** hat die Kategorie **k-Top** Produkte; jedoch sind diese  $k$ -Produkte (im allgemeinen echte) Verfeinerungen der entsprechenden topologischen Produkte. Insbesondere ist ein Paar  $(G, \tau)$  mit einer Gruppe  $G$  und einer  $k$ -Topologie  $\tau$  auf  $G$  im allgemeinen keine topologische Gruppe, wenn die Gruppenoperationen bezüglich des  $k$ -Produkts stetig sind. Wir werden ein solches Objekt eine  $k$ -Gruppe nennen.  $k$ -Gruppen zusammen mit den stetigen Gruppenhomomorphismen bilden die Kategorie **k-Grp**. Obwohl  $k$ -Gruppen im allgemeinen keine topologischen Gruppen sind, werden wir in Abschnitt 3.4 sehen, daß das Studium der freien Objekte in **k-Grp** beträchtliche Einsicht in die topologische Struktur freier topologischer Gruppen gestattet. Zunächst bemerken wir, daß man für einen  $k$ -Raum  $(X, \xi)$  in offensichtlicher Weise eine primitive  $k$ -Struktur  $\kappa_{(X, \xi)}^p$  auf  $FX$  definieren kann: in 3.2 muß man nur *topologische Potenz* durch  $k$ -Potenz ersetzen (man beachte, daß finale Strukturen in **k-Top** auch final in **Top** sind).

Unser erstes Resultat zeigt nun, daß – dank der kartesischen Abgeschlossenheit von **k-Top** (einer in diesem Zusammenhang fundamentalen Eigenschaft, an der es der Kategorie **Top** mangelt) – das Studium freier  $k$ -Gruppen viel einfacher ist als das freier topologischer Gruppen.

**Theorem 3.3** Für jeden  $k$ -Raum  $(X, \xi)$  ist die primitive  $k$ -Struktur kompatibel, d.h.,  $(FX, \kappa_{(X, \xi)}^p)$  ist eine  $k$ -Gruppe und somit die freie  $k$ -Gruppe  $G^k(X, \xi)$  über dem  $k$ -Raum  $(X, \xi)$ .

**Beweis.** Um die Stetigkeit der Gruppeninversion  $i$  zu beweisen, muß man nur bemerken, daß für jede Termabbildung  $t \in \mathcal{T}_{FX}$  die Abbildung  $i \circ t$  wieder eine Termabbildung und somit stetig bezüglich  $\kappa_{(X, \xi)}^p$  ist; aus der Finalität dieser Topologie bezüglich  $\mathcal{T}_{FX}$  folgt dann die Behauptung. Ähnlich folgt die Stetigkeit der Multiplikation  $m$ , da für jedes Paar  $(t_1, t_2)$  von Termabbildungen auch  $m \circ (t_1 \times t_2)$  eine Termabbildung und  $\{t_1 \times t_2 \mid t_1, t_2 \in \mathcal{T}_{FX}\}$  nach dem folgenden Lemma eine finale Familie ist. ■

**Lemma 3.4** Es sei  $(f_i: C_i \rightarrow C)_{i \in I}$  eine finale Epi-Senke (überdeckende Familie) in der kartesisch abgeschlossenen Kategorie  $\mathcal{C}$ . Dann ist auch  $(f_i \times f_j: C_i \times C_j \rightarrow C \times C)_{i, j \in I}$  eine finale Epi-Senke.

**Beweis.** Benutze [1, 27.22] und die Tatsache, daß finale Epi-Senken kompositiv sind. ■

Wir wollen einige Konsequenzen aus dem vorhergehenden Theorem festhalten. Hierzu benutzen wir folgende Begriffe und Bezeichnungen:

- Für eine Menge  $X$  wird das freie Monoid über der disjunkten Vereinigung von zwei Exemplaren von  $X$  mit  $\Gamma X$  bezeichnet.  $can_X: \Gamma X \rightarrow FX$  ist dann die kanonische Darstellung der freien Gruppe über  $X$  als ein (Monoid-) Quotient der freien  $\Gamma$ -Algebra  $\Gamma X$  über  $X$ . Es sei daran erinnert, daß die Elemente von  $\Gamma X$  Terme sind, wie z.B.  $xy^{-1}yz$  mit  $x, y, z \in X$ , wobei jedoch  $xy^{-1}yz$  verschieden ist von  $xz$ ; diese beiden Terme werden dann jedoch durch  $can_X$  identifiziert. Die Terme entsprechen wieder Termabbildungen; die Zahl  $a(t)$  von Variablen, die im Term  $t$  wirklich vorkommen, wird die *Arität* oder *Stelligkeit* von  $t$  genannt.
- Für einen  $k$ -Raum  $(X, \xi)$  kann der Menge  $\Gamma X$  (vermöge ihrer  $\Gamma$ -Terme) eine  $k$ -Topologie in derselben Weise wie in 3.2 aufgeprägt werden. Der so entstehende Raum wird mit  $\Gamma(\widehat{X}, \xi)$  bezeichnet. In vollständiger Analogie zu 3.3 ist dies dann die freie  $k$ - $\Gamma$ -Algebra über  $(X, \xi)$ .
- Mit  $\Gamma(\widehat{X}, \xi)$  bezeichnen wir den Raum  $\coprod_{t \in \Gamma X} (X, \xi)^{a(t)}$ . Offensichtlich gibt es eine kanonische Surjektion  $quot_{(X, \xi)}: \coprod_{t \in \Gamma X} (X, \xi)^{a(t)} \rightarrow \Gamma(\widehat{X}, \xi)$ , die charakterisiert ist durch  $quot_{(X, \xi)} \circ \iota_t = t$  für jede Termabbildung  $t: (X, \xi)^{a(t)} \rightarrow \Gamma(\widehat{X}, \xi)$  (die  $\iota_t$  sind die Coprodukt-Injektionen).
- Ausgehend von einer beliebigen Abzählung der Termabbildungen von  $FX$  (!), startend mit  $t_0 = \eta_X$ , bezeichnen wir mit  $F_k X$  die Menge  $\bigcup_{l \leq k} t_l[X^{a(t_l)}]$ . Offenbar gilt  $FX = \bigcup_{k=0}^{\infty} F_k X$ .  $F_k X$  kann dann mit der finalen Topologie in Bezug auf die Familie  $(t_l: (X, \xi)^{a(t_l)} \rightarrow F_k X)_{l \leq k}$  versehen werden; der entstehende Raum sei  $F_k(\widehat{X}, \xi)$ .

**Korollar 3.5** Für einen  $k$ -Raum  $(X, \xi)$  gelten die folgenden Aussagen:

1. Die kanonische Darstellung ist, betrachtet als Abbildung  $can_X: \Gamma(\widetilde{X}, \xi) \rightarrow (FX, \kappa_{(X, \xi)}^p)$ , eine topologische Quotientenabbildung.
2. Die kanonische Surjektion  $quot_{(X, \xi)}: \Gamma(\widetilde{X}, \xi) \rightarrow \Gamma(\widetilde{X}, \xi)$  ist eine topologische Quotientenabbildung.
3. Der  $k$ -Raum  $(FX, \kappa_{(X, \xi)}^p)$  ist ein Colimes der Kette  $(\widetilde{F_k X})_{k \in \mathbb{N}}$  in  $k\text{-Top}$ .

**Beweis.** 1. folgt aus der Tatsache, daß  $FX$ , versehen mit der finalen Topologie bezüglich  $can_X: \Gamma(\widetilde{X}, \xi) \rightarrow FX$ , die freie  $k$ -Gruppe über  $(X, \xi)$  ist. Dies ist eine Konsequenz aus der Kompatibilität dieser Topologie (benutze 3.4 und die Tatsache, daß  $can_X$  ein  $\Gamma$ -Algmorphismus ist) und der Bemerkung, daß die stetige homomorphe Fortsetzung jeder stetigen Abbildung  $f$  von einem  $k$ -Raum  $(X, \xi)$  in eine  $k$ -Gruppe  $(G, \tau)$  über  $\Gamma(\widetilde{X}, \xi)$  faktorisiert (benutze die entsprechende Beobachtung für abstrakte Algebren und die Finalität von  $can_{(X, \xi)}$ ).

2. ist offensichtlich, da  $quot_{(X, \xi)}$  der erste Faktor einer finalen Senke ist (vg. [1, 8.13]).

3. gilt offenbar für die unterliegenden Mengen; benutze nun 3.3 und die Kompositionsabgeschlossenheit finaler Senken. ■

### 3.3 Freie $k$ -Gruppen über $t_2$ - und $k_\omega$ -Räumen

Wir wollen nun die verschiedenen Beschreibungen der Topologie der freien  $k$ -Gruppe  $G^k(X, \xi)$  über einem  $k$ -Raum  $(X, \xi)$ , die wir im letzten Abschnitt erhalten haben, anwenden. Ziel ist hier das Studium topologischer Eigenschaften, die  $G^k(X, \xi)$  von  $(X, \xi)$  ererbt. Es handelt sich dabei insbesondere um die beiden folgenden:

**Definition 3.6** ( $t_2$ -Räume und  $k_\omega$ -Räume) Ein  $k$ -Raum  $(X, \xi)$  heißt

- schwach Hausdorff'sch oder ein  $t_2$ -Raum, falls seine Diagonale  $\Delta_X$  abgeschlossen im  $k$ -Produkt  $(X, \xi) \times (X, \xi)$  ist.  $k\text{-Top}_2$  bezeichnet die Kategorie der  $t_2$ -Räume.
- $k_\omega$ -Raum, falls er homöomorph ist zum Colimes einer abzählbaren Kette kompakter Hausdorff Räume.  $k\text{-Top}_\omega$  bezeichnet die entsprechende volle Unterkategorie von  $k\text{-Top}$ .

Da  $k$ -Produkte Verfeinerungen topologischer Produkte sind, ist jeder Hausdorff-Raum schwach Hausdorff'sch. Die für das folgende wichtigsten topologischen Eigenschaften der gerade beschriebenen Raumklassen fassen wir in dem nächsten Satz zusammen. Für Beweise verweisen wir auf die angegebene Literatur.

**Satz 3.7** ([7, 12, 14, 16])

1.  $k\text{-Top}_2$  ist eine reflexive Unterkategorie von  $k\text{-Top}$ ; ihre Egalisatoren sind genau die abgeschlossenen Einbettungen.
2. Ein gerichteter Colimes einer Kette von  $t_2$ -Räumen ist wieder ein  $t_2$ -Raum.
3. Ist  $q: (X, \xi) \rightarrow (Y, \nu)$  ein Quotient in  $k\text{-Top}$  und ist  $(X, \xi)$  ein  $t_2$ -Raum, so ist  $(Y, \nu)$   $t_2$  genau dann, wenn  $(q \times q)^{-1}[\Delta_Y]$  abgeschlossen in  $(X, \xi)$  ist.
4. Ein  $t_2$ -Raum ist  $k_\omega$ , falls er Quotient eines  $k_\omega$ -Raums oder Colimes einer abzählbaren Kette kompakter Räume ist.
5. Produkte von  $k_\omega$ -Räumen stimmen in  $\text{Top}$  und  $k\text{-Top}$  überein.
6.  $k_\omega$ -Räume sind normal.

■

Um zu beweisen, daß  $G^k(X, \xi)$  die  $t_2$ -Eigenschaft vom Raum  $(X, \xi)$  ererbt, werden wir folgendes Kriterium benutzen.

**Satz 3.8** Für einen schwach Hausdorff'schen  $k$ -Raum  $(X, \xi)$  ist die freie  $k$ -Gruppe  $G^k(X, \xi)$  ebenfalls schwach Hausdorff'sch, falls für jedes Paar  $(s, t)$  von Termabbildungen von  $\Gamma X$  die Menge

$$M(s, t) = \{(x, y) \in X^{a(s)} \times X^{a(t)} \mid \text{can}_X s(x) = \text{can}_X t(y)\}$$

eine abgeschlossene Teilmenge im  $k$ -Produkt  $(X, \xi)^{a(s)} \times (X, \xi)^{a(t)}$  ist.

**Beweis.** In einem ersten Schritt zeigen wir, daß  $\Gamma(\widehat{X}, \xi)$  ein  $t_2$ -Raum ist. Nach Definition ist dieser Raum der gerichtete abzählbare Colimes (bezüglich einer vorgegebenen Abzählung der Termabbildungen von  $\Gamma X$ )

$$\Gamma(\widehat{X}, \xi) = \text{colim}_{k \rightarrow \infty} \prod_{l \leq k} (X, \xi)^{a(l)}.$$

Da die  $t_2$ -Eigenschaft offensichtlich stabil unter endlichen topologischen Summen ist, liefern 3.7 (1.) und (2.) das gewünschte Resultat.

Mit der Abkürzung  $q = \text{can}_X \circ \text{quot}_{(X, \xi)}$  gilt die Mengengleichheit

$$M(s, t) = (q \times q)^{-1}[\Delta_{G^k(X, \xi)}] \cap ((X, \xi)^{a(s)} \times (X, \xi)^{a(t)}).$$

Unsere Voraussetzung garantiert nun, daß  $(q \times q)^{-1}[\Delta_{G^k(X, \xi)}]$  einen abgeschlossenen Durchschnitt mit jedem Summanden von

$$\Gamma(\widehat{X}, \xi) \times \Gamma(\widehat{X}, \xi) = \prod_{(s, t)} (X, \xi)^{a(s)} \times (X, \xi)^{a(t)}$$

hat und somit abgeschlossen in  $\Gamma(\widehat{X}, \xi)^2$  ist. Nun folgt alles mit 3.7 (3.) und 3.5. ■

Unter Anwendung des Kriteriums 3.8 folgt

**Theorem 3.9 (Lamartin)** Für jeden schwach Hausdorff'schen  $k$ -Raum  $(X, \xi)$  ist die freie  $k$ -Gruppe  $G^k(X, \xi)$  schwach Hausdorff'sch.

An Stelle eines vollständigen Beweises, der recht technisch wäre (vgl. [17]), illustrieren wir das allgemeine Argument an einem typischen Beispiel: Wir betrachten die beiden  $\Gamma X$ -Terme  $s = xy^{-1}r^{-1}rzw$  und  $t = uw$ . Die (Einschränkungen der) entsprechenden Termabbildungen sind

$$\begin{aligned} s: X^5 &\longrightarrow FX & \text{and } t: X^2 &\longrightarrow FX \\ (a_1, \dots, a_5) &\mapsto a_1 a_2^{-1} a_3^{-1} a_3 a_4 a_5 & (b_1, b_2) &\mapsto b_1 b_2 \end{aligned}$$

Es ist nun leicht nachzurechnen, daß  $M(s, t) = M_1 \cup M_2$  gilt mit

$$M_1 = \{(a_1, a_1, a_3, a_4, a_5, a_4, a_5) \mid a_i \in X\} \quad \text{und} \quad M_2 = \{(a_1, a_2, a_3, a_2, a_5, a_1, a_5) \mid a_i \in X\}.$$

Wegen  $M_1 = X^3 \times \Delta_{X^2}$  und  $M_2 \simeq X \times \Delta_X^3$  ist  $M(s, t)$  also Vereinigung zweier abgeschlossener Teilmengen von  $(X, \xi)^7 = (X, \xi)^{a(s)} \times (X, \xi)^{a(t)}$  und somit abgeschlossen. ■

Nun ist es nicht mehr schwierig nachzuweisen, daß  $G^k(X, \xi)$  auch die  $k_\omega$ -Eigenschaft von  $(X, \xi)$  ererbt.

**Theorem 3.10** Für jeden  $k_\omega$ -Raum  $(X, \xi)$  ist die freie  $k$ -Gruppe  $G^k(X, \xi)$  wieder  $k_\omega$  und enthält  $(X, \xi)$  als einen abgeschlossenen Teilraum.

**Beweis.** Sei zunächst der Raum  $(X, \xi)$  sogar kompakt Hausdorff'sch. Dann ist  $\Gamma(\widehat{X}, \xi)$  ein  $k_\omega$ -Raum vermöge der auch im Beweis von 3.8 benutzten Darstellung, da endliche Summen von Potenzen kompakter Hausdorff-Räume kompakt Hausdorff'sch sind. Nach 3.9 und 3.7 (4.) ist nun alles bewiesen. Alternativ: wegen 3.9 erfüllt die Darstellung aus 3.5 (3.) die zweite Voraussetzung von 3.7 (4.).

Sei nun  $(X, \xi)$  als Colimes der Kette  $(X_n, \xi_n)_{n \in \mathbb{N}}$  kompakter Hausdorff-Räume. Für jedes  $n \in \mathbb{N}$  gibt es nach dem ersten Schritt dieses Beweises eine Darstellung (des unterliegenden Raums) von  $G^k(X, \xi)_n$  als Colimes einer abzählbaren Kette kompakter Hausdorff-Räume  $(C_k^n)_{k \in \mathbb{N}}$ . Aber dann ist  $G^k(X, \xi)$  als Raum auch Colimes der Kette  $(C_k^n)_{n \in \mathbb{N}}$  und somit ein  $k_\omega$ -Raum. Dieses Argument stützt sich auf die Tatsachen, daß (i)  $G^k$  als Linksadjungierter und (ii) der Funktor  $V$  in die kartesisch abgeschlossene topologische Kategorie  $\mathbf{k-Top}$  (vgl. [16, 3.1]) gerichtete Colimiten bewahrt, so daß  $VG^k(X, \xi) = \text{colim}_{n \rightarrow \infty} VG^k(X, \xi)_n$ , und daß (iii) eine spezielle Version des Satzes über Limiten mit Parametern (vgl. [8]) in topologischen Kategorien gilt (vgl. [16, 4.1]).

Ausgehend von der Darstellung in 3.5 (3.) zeigt dieselbe Überlegung auch die Abgeschlossenheit von  $(X, \xi)$  in  $VG(X, \xi)$ . ■

### 3.4 Zur Struktur freier topologischer Gruppen

Wir werden nun mit Hilfe der Ergebnisse des letzten Abschnitts Aussagen über die algebraische und topologische Struktur der freien *topologischen* Gruppe  $G(X, \xi)$  über einem topologischen Raum  $(X, \xi)$  herleiten. Der Schlüssel hierfür ist der folgende Satz.

**Satz 3.11** Für jeden  $k_\omega$ -Raum  $(X, \xi)$  stimmt die freie topologische Gruppe  $G(X, \xi)$  überein mit der freien  $k$ -Gruppe  $G^k(X, \xi)$ . Insbesondere ist die primitive Topologie eine Gruppentopologie, so daß  $G(X, \xi) = (FX, \tau_{(X, \xi)}^p)$  gilt.

**Beweis.** Da  $G^k(X, \xi)$  nach 3.10 ein  $k_\omega$ -Raum ist, folgt aus 3.7 (5.), daß  $G^k(X, \xi)$  auch eine topologische Gruppe ist. Darüberhinaus zeigt 3.7 (5.) in Verbindung mit der Tatsache, daß finale Strukturen in **k-Top** auch final in **Top** sind, daß die primitive  $k$ -Struktur  $\kappa_{(X, \xi)}^p$  die primitive Topologie  $\tau_{(X, \xi)}^p$  ist. ■

Eine unmittelbare Folge hieraus ist

**Theorem 3.12 (Mal'cev)** Die freie topologische Gruppe  $G(X, \xi)$  über einem  $k_\omega$ -Raum  $(X, \xi)$  ist normal und enthält  $(X, \xi)$  als einen abgeschlossenen Teilraum.

**Beweis.** Benutze 3.11 und 3.10 zusammen mit 3.7 (6.). ■

Auf sehr einfache Weise kann man nun von kompakten Hausdorff-Räumen auf Tychonoff-Räume verallgemeinern. Wir beweisen zunächst den wichtigsten Teil von Świrczkowski's schon zu Beginn zitierten Resultat 3.14, wobei unser Beweis zusätzlich eine konzeptionelle Interpretation der konstruierten Topologie zuläßt. Wir möchten jedoch darauf hinweisen, daß wir nicht wissen, ob beide Konstruktionen dieselbe Topologie beschreiben.

**Satz 3.13** Für jeden Tychonoff-Raum  $(X, \xi)$  trägt die freie abstrakte Gruppe  $FX$  eine Tychonoff'sche Gruppentopologie  $\sigma_{(X, \xi)}$ . Die topologische Gruppe  $(FX, \sigma_{(X, \xi)})$  ist die freie vollständig reguläre Gruppe über  $(X, \xi)$ , die sich zusätzlich als Hausdorff'sch erweist und somit auch die freie Tychonoff Gruppe<sup>2</sup> über  $(X, \xi)$  ist. Sie enthält den Raum  $(X, \xi)$  als einen (abgeschlossen<sup>3</sup>) Teilraum.

**Beweis.** Ersetzt man im Beweis von 2.3 die Kategorie **Top** durch die Kategorie **Creg** (die — als konkrete reflexive Unterkategorie von **Top** (vgl. [1, 21.35]) — ebenfalls beliebige initiale Strukturen zuläßt), so erhält man eine topologische Gruppe  $G^c(X, \xi) = (FX, \sigma_{(X, \xi)})$ , die frei über dem Raum  $(X, \xi)$  in der Kategorie **CregGrp** der vollständig regulären Gruppen ist. Sei nun  $\beta: (X, \xi) \rightarrow \beta(X, \xi)$  die Stone-Čech Kompaktifizierung von  $(X, \xi)$ . Da  $G\beta(X, \xi)$

<sup>2</sup>Wir hätten hier natürlich Hausdorff Gruppe sagen können; der Grund für unsere Terminologie ergibt sich aus der abschließenden Bemerkung dieser Arbeit.

<sup>3</sup>Die Frage der Abgeschlossenheit diskutieren wir hier nicht — vgl. etwa [19].

nach 3.12 zu **CregGrp** gehört, liefert die universelle Eigenschaft von  $G^c(X, \xi)$  eine stetige Injektion  $F\beta: G^c(X, \xi) \rightarrow G\beta(X, \xi)$ , die das folgende Diagramm kommutieren läßt.

$$\begin{array}{ccc} (X, \xi) & \xrightarrow{\eta_{(X, \xi)}} & G^c(X, \xi) \\ \beta \downarrow & & \downarrow F\beta \\ \beta(X, \xi) & \xrightarrow{\eta_{\beta(X, \xi)}} & G\beta(X, \xi) \end{array}$$

Weil  $G\beta(X, \xi)$  nach 3.12 Hausdorff'sch ist, ist dann auch  $\sigma_{(X, \xi)}$  Hausdorff'sch. Also ist  $\sigma_{(X, \xi)}$  Tychonoff'sch. Das Diagramm zeigt zusätzlich, daß  $\eta_{(X, \xi)}$  eine topologische Einbettung ist, denn die Einbettung  $\eta_{\beta(X, \xi)} \circ \beta$  faktorisiert über  $\eta_{(X, \xi)}$ . ■

Świrczkowski's vollständiges Resultat ergibt sich nun als einfaches Korollar:

**Satz 3.14** (Świrczkowski) *Für jeden funktional Hausdorff'schen Raum  $(X, \xi)$  trägt die abstrakte freie Gruppe  $FX$  über der unterliegenden Menge von  $(X, \xi)$  eine Tychonoff'sche Gruppentopologie, die  $\eta_X$  stetig sein läßt. Ist  $(X, \xi)$  sogar ein Tychonoff-Raum, so ist  $\eta_X$  eine Einbettung.*

**Beweis.** Ist  $(X, \xi)$  funktional Hausdorff'sch, so bildet man zunächst die Tych-Reflexion  $\rho_{(X, \xi)}: (X, \xi) \rightarrow \rho(X, \xi)$ ; da  $\rho_{(X, \xi)}$  als Abbildung die Identität ist, folgt  $FT(X, \xi) = FT\rho(X, \xi)$  und damit alles weiter aus 3.13. Beachte, daß  $\eta_{(X, \xi)}$  and  $\eta_{\rho(X, \xi)}$  zwar als Abbildungen übereinstimmen,  $\eta_{\rho(X, \xi)}$  jedoch nur dann eine Einbettung ist, wenn  $(X, \xi)$  schon ein Tychonoff-Raum war. ■

Unter Benutzung der Anmerkung im Anschluß an die Taut-Lift-Sätze ergibt sich unmittelbar folgende Aussage über die algebraische Struktur der freien Gruppen.

**Theorem 3.15** *Die freie Tychonoff'sche<sup>4</sup> Gruppe (= freie (funktional) Hausdorff'sche Gruppe = freie topologische Gruppe!) über einem funktional Hausdorff'schen Raum  $(X, \xi)$  ist algebraisch die freie Gruppe  $FX$  über der unterliegenden Menge von  $(X, \xi)$ .* ■

**Bemerkung.** Für einen nicht funktional Hausdorff'schen Raum  $(X, \xi)$  erhält man die freie Hausdorff Gruppe als freie (Hausdorff-) topologische Gruppe über der **FHaus**-Reflexion  $\chi(X, \xi)$  von  $(X, \xi)$ ; sie ist dann also algebraisch die freie Gruppe über der unterliegenden Menge von  $\chi(X, \xi)$ , die von echt kleinerer Mächtigkeit als  $X$  ist.

Als Aussagen über die topologische Struktur erhalten wir schließlich

<sup>4</sup>siehe Fußnote 2

**Theorem 3.16** Sei  $(X, \xi)$  ein topologischer Raum und  $G(X, \xi)$  die freie topologische Gruppe über  $(X, \xi)$ . Dann gilt  $G(X, \xi) = (FX, \tau_f)$  mit der Einbettung der Generatoren als universeller stetiger Abbildung  $\eta_{(X, \xi)}: (X, \xi) \rightarrow G(X, \xi)$ .

Darüber hinaus gilt

1.  $(X, \xi)$  ist Tychonoff'sch  $\iff G(X, \xi)$  ist Tychonoff'sch<sup>5</sup> und  $\eta_{(X, \xi)}$  ist eine Einbettung
2.  $(X, \xi)$  ist funktional Hausdorff'sch  $\iff G(X, \xi)$  ist funktional Hausdorff'sch<sup>5</sup>.

In beiden Fällen ist  $G(X, \xi)$  auch die freie Hausdorff-Gruppe über  $(X, \xi)$ .

**Beweis.** Ist  $(X, \xi)$  ein funktional Hausdorff'scher Raum, so ist  $id_{FX}: G(X, \xi) \rightarrow (FX, \sigma_{(X, \xi)})$  die stetige (!) homomorphe Fortsetzung von  $\eta_{(X, \xi)}: (X, \xi) \rightarrow (FX, \sigma_{(X, \xi)})$ ; also verfeinert die Topologie von  $G(X, \xi)$  die Tychonoff-Topologie  $\sigma_{(X, \xi)}$  und ist somit funktional Hausdorff'sch, also auch Tychonoff'sch.

Ist  $(X, \xi)$  sogar Tychonoff'sch, so ist  $\eta_{(X, \xi)}: (X, \xi) \rightarrow G(X, \xi)$  eine Einbettung, wie man mit demselben Argument wie im Beweis zu 3.13 sieht; man muß nur  $G^c(X, \xi)$  durch  $G(X, \xi)$  ersetzen.

Der Rest ist offensichtlich. ■

**Bemerkung.** In dieser Arbeit kann man durchgängig *Gruppe* durch *gleichungsdefinierte Algebra* (bezüglich eines gegebenen endlichen Typs  $\Omega$ ) ersetzen. Wenn dann  $\Omega$  an Stelle des Gruppentyps  $\Gamma$  und die entsprechende Varietät anstatt **Grp** gesetzt wird, bleiben alle Ergebnisse ohne jede Änderung der Beweise richtig, mit Ausnahme des Resultats (1.) in 3.16, das dann nur wie folgt lautet

- $(X, \xi)$  ist Tychonoff'sch  $\implies G(X, \xi)$  ist funktional Hausdorff'sch und  $\eta_{(X, \xi)}$  ist eine Einbettung

(der Beweis von 3.16 machte hier von der Tatsache Gebrauch, daß eine funktional Hausdorff'sche Gruppe Tychonoff'sch ist; er kann hier deshalb nur auf Algebren wie Abelsche Gruppen oder Ringe verallgemeinert werden).

Es ist ein offenes Problem, ob 3.16 (1.) für beliebige Algebren gilt, d.h., ob die freie topologische Algebra über einem Tychonoff-Raum stets Tychonoff'sch ist, auch wenn die Algebra keine unterliegende Gruppenstruktur hat.

---

<sup>5</sup>siehe Fußnote 2

## Literatur

- [1] Adámek, J. Herrlich, H., and Strecker, G.E.: *Abstract and Concrete Categories*. New York 1990
- [2] Fay, T.H.: *An axiomatic approach to categories of topological algebras*. *Quaestiones Math.* **2**, 113–137 (1977)
- [3] Freyd, P.: *Functor theory*. PhD Dissertation, Princeton University, Princeton 1960
- [4] Herrlich, H.: *Topologische Reflexionen und Coreflexionen*. *Lecture Notes in Mathematics* **78**, Berlin 1968
- [5] Kakutani, S.: *Free topological groups and finite discrete product topological groups*. *Proc. Imp. Acad. Tokyo* **20**, 595–598 (1944)
- [6] Kelley, J.L.: *General Topology*. New York 1975
- [7] Lamartin, W.F.: *On the Foundations of  $k$ -group Theory*. *Dissertationes Math.* (Rozprawy Mat.) **CXLVI**, Warsaw 1977
- [8] Mac Lane, S.: *Categories for the Working Mathematician*. New York 1972
- [9] Mal'cev, A.I.: *Free topological algebras*. *Izv. Akad. Nauk SSSR Ser. Mat.* **21**, 171–198 (1957) [ *Amer. Math. Soc. Transl. Ser. 2* **17**, 173–200 (1961)]
- [10] Manes, E.G.: *Algebraic Theories*. New York 1976
- [11] Markov, A.A.: *On free topological groups*. *Izv. Akad. Nauk. SSSR Ser. Mat.* **9**, 3–64 (1945) [ *Amer. Math. Soc. Transl.* **30**, 11–88 (1950); Reprint: *Amer. Math. Soc. Transl. Ser. 1* **8**, 195–272 (1962)]
- [12] McCord, M.C.: *Classifying spaces and infinite symmetric products*. *Trans. Amer. Math. Soc.* **146**, 273–298 (1969)
- [13] Nel, L.D.: *Universal topological algebra needs closed topological categories*. *Topology Appl.* **12**, 321–330 (1981)
- [14] Ordman, E.T.: *Free  $k$ -groups and free topological groups*. (General) *Topology Appl.* **5**, 205–219 (1975)
- [15] Porst, H.-E.: *On free topological algebras*. *Cahiers Topologie Géom. Différentielles Catégoriques* **28**, 235–253 (1987)
- [16] Porst, H.-E.: *Free algebras over cartesian closed topological categories*. In: Frolík, Z. (Ed.): *General Topology and its Relations to Modern Analysis and Algebra*. VI. Proc. Sixth Prague Topological Symposium 1986, pp. 437–450, Berlin 1988

- [17] **Porst, H.-E.:** *Separation in free algebras over Top and k-Top.* In: Adámek, J., and MacLane, S. (Eds.): *Categorical Topology and its Relations to Analysis, Algebra and Combinatorics.* Proc. Conf. Prague 1988, pp. 238–245, Singapore 1989
- [18] **Samuel, P.:** *On universal mappings and free topological groups.* Bull. Amer. Math. Soc. **54**, 591–598 (1948)
- [19] **Smith-Thomas, B.:** *Free topological groups.* (General) Topology Appl. **4**, 51–72 (1974)
- [20] **Świrczkowski, S.:** *Topologies in free algebras.* Proc. London Math. Soc. (3) **14**, 566–576 (1964)
- [21] **Tholen, W.:** *On Wyler's Taut Lift Theorem.* (General) Topology Appl. **8**, 197–206 (1978)
- [22] **Wyler, O.:** *On the categories of general topology and topological algebra.* Arch. Math. **XXII**, 7–17 (1971)

eingegangen: Januar 1991

**Verfasser:**

Prof. Dr. H.-E. Porst  
Fachbereich Mathematik und Informatik  
Universität Bremen  
Postfach 33 04 40  
W-2800 Bremen 33  
Germany

FRITZ V. HAESELER

# Fraktale Geometrie und Juliamengen

---

## 1. Einleitung

Der Begriff *fraktale Geometrie* wurde 1977 von B.B.Mandelbrot [11] geprägt. In einer umgangssprachlichen Formulierung können alle Objekte fraktal genannt werden, die unter noch so starker Vergrößerung niemals glatt erscheinen, also niemals gerade Linien zeigen.

Mathematisch präziser ist die Definition fraktaler Objekte durch den Begriff der Hausdorffdimension möglich. Ein fraktales Objekt in einem metrischen Raum ist eine Menge, deren Hausdorffdimension keine natürliche Zahl ist. Dies ist Mandelbrots Definition für fraktale Mengen. Die Hausdorffdimension wurde 1915 von F. Hausdorff [8] eingeführt. Wir wollen uns jedoch mit der anschaulichen Sichtweise begnügen.

Beginnen werden wir mit zwei klassischen fraktalen Objekten, die zunächst nach klassischem Vorbild konstruiert werden. Anschließend werden wir zwei Charakterisierungen dieser Objekte vorstellen, die durch die Theorie der dynamischen Systeme motiviert sind.

Diese dynamische Charakterisierung fraktaler Objekte wird uns bei der Iteration rationaler Abbildungen auf der Riemannschen Sphäre erneut begegnen. Die Theorie der Iteration rationaler Abbildungen wurde um 1920 von den französischen Mathematikern G. Julia [10] und P. Fatou [6] unabhängig voneinander entwickelt. Heute wird diese Theorie meist als Theorie der Juliamengen oder als Theorie der holomorphen dynamischen Systeme bezeichnet. Wir werden uns ausschließlich mit der Iteration des Polynoms  $p_c(z) = z^2 + c$ ,  $z \in \hat{C}$  und  $c \in C$  befassen.

## 2. Zwei klassische Fraktale

Wir stellen nun zwei klassische Fraktale und ihre Konstruktion vor. Es sind die Cantormenge, die von G. Cantor 1878 als Beispiel einer perfekten, in  $\mathbb{R}$  nirgends dichten Menge erfunden wurde, und das Sierpińskidreieck, welches 1915 von W. Sierpiński als Beispiel einer Kurve, die nur Verzweigungspunkte hat, entwickelt wurde. Beide Mengen sind Paradebeispiele für Fraktale.

Beginnen wir mit dem Konstruktionsverfahren der Cantormenge. Es sei  $I_0 = [0, 1]$  und  $I_1 = [0, 1/3] \cup [2/3, 1]$ , also der Teil von  $I_0$ , der nach Herausnahme des offenen mittleren Drittels von  $I_0$  bleibt. Die beiden Teilintervalle von  $I_1$  werden nun genauso behandelt, wie das Intervall  $I_0$ ; d.h., die offenen mittleren Drittel werden erneut entfernt. Somit haben wir eine Menge  $I_2$  erhalten, die aus 4 Intervallen der Länge  $1/9$  zusammengesetzt ist. Iterieren wir diesen Prozeß, so erhalten wir im  $n$ -ten Schritt eine kompakte Menge  $I_n$ , die aus  $2^n$  Intervallen der Länge  $3^{-n}$  besteht. Die Menge

$$C = \bigcap_{n=0}^{\infty} I_n$$

heißt Cantormenge. Die Cantormenge ist ein Fraktal mit der Hausdorffdimension  $\ln 2 / \ln 3$ . Wenden wir uns nun dem Konstruktionsverfahren des Sierpińskidreiecks zu. Wie bei der Cantormenge ist der Ausgangspunkt eine kompakte Menge, aus der iterativ bestimmte offene Teile entfernt werden. Es sei  $D_0$  das Dreieck mit den Eckpunkten  $(0, 0)$ ,  $(1, 0)$ ,  $(1/2, \sqrt{3}/2)$ . Dann ist  $D_1$  der Teil von  $D_0$ , der nach der Herausnahme des offenen, von den Mittelpunkten der Seiten von  $D_0$  erzeugten Dreiecks verbleibt.  $D_1$  ist also aus 3 gleichseitigen Dreiecken zusammengesetzt. Im nächsten Schritt werden diese Teildreiecke ebenso behandelt wie das Ausgangsdreieck  $D_0$ . Wir erhalten eine Menge  $D_2$ , die aus 9 kleinen gleichseitigen Dreiecken besteht. Diesen Prozeß wiederholen wir. Im  $n$ -ten Schritt haben wir eine Menge  $D_n$ , die aus  $3^n$  gleichseitigen Dreiecken besteht. Die Menge

$$S = \bigcap_{n=0}^{\infty} D_n$$

heißt Sierpińskidreieck (Abbildung 1). Das Sierpińskidreieck ist ein Fraktal mit der Hausdorffdimension  $\ln 3 / \ln 2$ .

Wir charakterisieren die Cantormenge bzw. das Sierpińskidreieck durch ein dynamisches System. Wieder beginnen wir mit der Cantormenge. Durch

$$f(x) = \begin{cases} 3x & x \leq .5 \\ -3x + 3 & x > .5 \end{cases} \quad x \in \mathbb{R}$$

ist eine reellwertige stetige Abbildung definiert. Mit  $f^n(x) = f(f^{n-1}(x))$  bezeichnen wir die  $n$ -te Iterierte von  $f$ . Jedem Punkt  $x \in \mathbb{R}$  können wir seinen Orbit

$$\mathcal{O}(x) = \{f^n(x) \mid n \in \mathbb{N}\}$$

zuordnen. Es existieren genau zwei Typen von Orbits. Der erste ist unbeschränkt, wir sagen  $\lim_{n \rightarrow \infty} f^n(x) = -\infty$ , und der zweite ist beschränkt. Wie man leicht sieht, gilt für die beschränkten Orbits  $\mathcal{O}(x) \subset [0, 1]$ . Offenbar ist die Menge der Punkte  $x \in \mathbb{R}$  mit  $\lim_{n \rightarrow \infty} f^n(x) = -\infty$  offen. Wir bezeichnen sie mit  $A(-\infty)$ . Damit ist die Menge

$$\mathcal{B} = \{x \in \mathbb{R} \mid \mathcal{O}(x) \subset [0, 1]\}$$

der Punkte  $x \in \mathbb{R}$  mit beschränktem Orbit kompakt. Eine Charakterisierung von  $\mathcal{B}$  ist durch

$$\mathcal{B} = \bigcap_{n=0}^{\infty} f^{-n}([0, 1])$$

gegeben. Die Menge der Punkte  $x \in \mathbb{R}$  mit beschränktem Orbit ist somit die Cantormenge; also  $\mathcal{B} = \mathcal{C}$ . Anders gesehen, die Cantormenge ist der Rand des Attraktionsgebiets  $A(-\infty)$ .

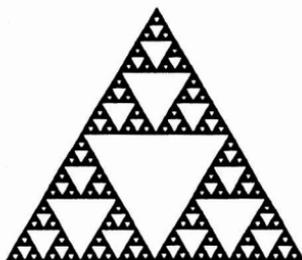


Abbildung 1 Das Sierpińskidreieck

Die dynamische Charakterisierung des Sierpińskidreiecks benutzt die Theorie der iterierten Funktionensysteme. Diese Theorie wurde um 1981 von J. E. Hutchinson [9] entwickelt und von M. Barnsley [1] fortgeführt. Wir beginnen mit der Definition dreier Abbildungen  $f_j : C \rightarrow C$ ,  $j = 1, 2, 3$ ,

$$f_1(z) = z/2$$

$$f_2(z) = (z + 1)/2$$

$$f_3(x) = (z + .5 + i\sqrt{3}/2)/2.$$

Jede Abbildung  $f_j$ ,  $j = 1, 2, 3$ , ist eine Kontraktion. Die Fixpunkte dieser Kontraktionen sind genau die äußeren Ecken des Sierpińskidreiecks. Das Sierpińskidreieck ist nun durch folgende Eigenschaft bestimmt. Für jedes  $x \in \mathbb{R}$  gilt

$$S = \text{Häufungspunkte } \{f_{i_n} \circ \dots \circ f_{i_2} \circ f_{i_1} \mid n \in \mathbb{N} \text{ und } i_j \in \{1, 2, 3\}\}.$$

Die Idee, verschiedene Abbildungen gemeinsam zu iterieren, wird uns bei der Theorie der Juliamengen erneut begegnen.

### 3. Theorie der Juliamengen

Wir wollen nun diskutieren, wie die Betrachtungen des vorigen Abschnitts mit der Theorie der Juliamengen zusammenhängen. Dies soll am Beispiel der Iteration des Polynoms  $p_c(z) = z^2 + c$  auf der Riemannschen Sphäre  $\hat{C}$  erläutert werden. Dabei ist  $c \in C$  ein komplexer Parameter.

Es ist leicht einzusehen, daß für hinreichend große Werte von  $|z|$  der Orbit  $\mathcal{O}(z) = \{p_c^n(z) \mid n \in \mathbb{N}\}$  unbeschränkt ist, d.h.  $\lim_{n \rightarrow \infty} p_c^n(z) = \infty$ . Die Menge der Punkte  $z \in \hat{C}$  mit unbeschränktem Orbit ist offen und heißt Attraktionsgebiet,  $A_c(\infty)$ , von  $\infty$ . Wie bei der dynamischen Definition der Cantormenge ist auch hier der Rand des Attraktionsgebiets interessant.

Das Polynom  $p_c$  besitzt zwei Umkehrfunktionen  $f_1, f_2$ . Analog zur Charakterisierung des Sierpińskidreiecks interessiert uns die Menge der Häufungspunkte aller Kombinationen von Iterierten der Umkehrfunktionen von  $p_c$ .

Beide Probleme wollen wir für das Polynom  $p_c$  diskutieren. Das erste Problem ist für die Iteration beliebiger rationaler Abbildungen auf  $\hat{C}$  nicht sehr sinnvoll, wohl aber das zweite. Wir beginnen mit der Untersuchung von  $A_c(\infty)$ . Da  $p_c^{-1}(\infty) = \{\infty\}$  ist, ist  $A_c(\infty)$  stets zusammenhängend.

**Definition 1** Der Rand von  $A_c(\infty)$  heißt *Juliamenge*,  $J_c$ , von  $p_c$ .

Für beliebige rationale Abbildungen läßt sich unter Benutzung des Begriffs der normalen Familie ebenfalls eine Juliamenge definieren; insbesondere ist die Juliamenge einer rationalen Abbildung niemals die leere Menge. Für unsere Zwecke genügt die obige Definition, die äquivalent zu der allgemeinen Definition ist. In Abbildung 2 sind einige Juliamengen zu sehen. Wir sehen, daß Juliamengen fraktal sind. Dieser Sachverhalt wird durch einen Satz von Fatou (vgl. Brolin [3]) erhärtet. Nach diesem Satz besitzt die Juliamenge  $J_c$  für  $c \in C \setminus \{-2, 0\}$  keine Tangenten. Für  $c \in \{0, -2\}$  sind die Juliamengen leicht anzugeben. Es gilt  $J_0 = \{z \in C \mid |z| = 1\}$  und  $J_{-2} = [-2, 2]$ .

Betrachten wir nun das Problem der Rückwärtsiteration von  $p_c$ , so gilt:

**Theorem 1** Für alle  $z \in A_c(\infty) \setminus \{\infty\}$  gilt

$$J_c = \text{Häufungspunkte } \{w \in C \mid p_c^n(w) = z \text{ für ein } n \in \mathbb{N}\}.$$

Auch dieses Theorem besitzt eine Verallgemeinerung für rationale Abbildungen. Im allgemeinen gilt obige Aussage jedoch nicht für alle  $z \in \hat{C}$ . Hier ist also ein Unterschied zur dynamischen Charakterisierung des Sierpińskidreiecks festzustellen.

Zum Abschluß wollen wir die Juliamengen  $J_c$  noch etwas genauer betrachten. Diese Betrachtung wird dann zu einem Fraktal führen, welches in der Theorie der Juliamengen von überragender Bedeutung ist, der Mandelbrotmenge.

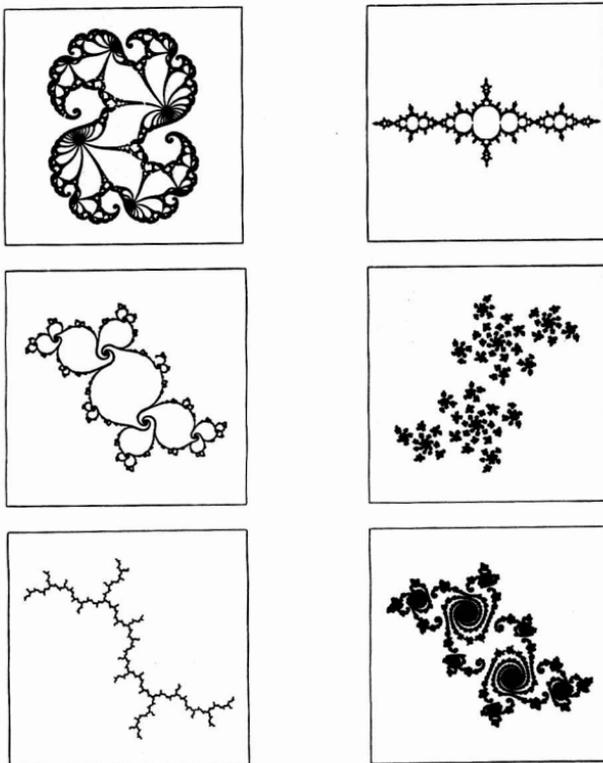


Abbildung 2 Einige Juliamengen des Polynoms  $p_c$

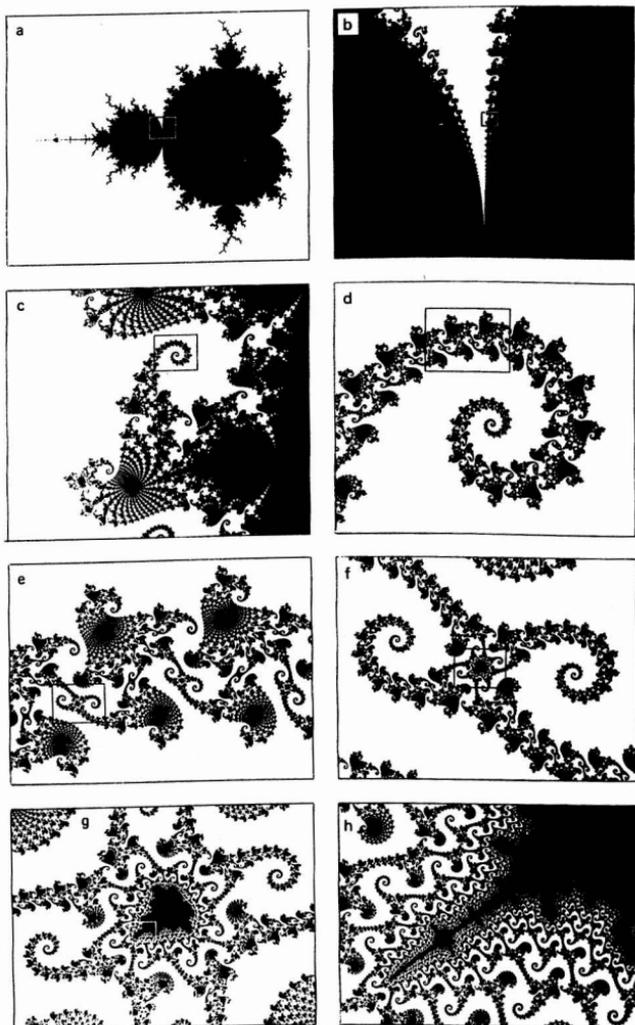


Abbildung 3 Die Mandelbrotmenge und einige Vergrößerungen

**Theorem 2**  $A_c(\infty)$  ist entweder einfach zusammenhängend oder von unendlich hohem Zusammenhang.  $A_c(\infty)$  ist genau dann einfach zusammenhängend, wenn  $0$ , der kritische Punkt von  $p_c$ , nicht in  $A_c(\infty)$  enthalten ist.

Ist  $A_c(\infty)$  einfach zusammenhängend, so ist  $J_c$  zusammenhängend. Ist  $A_c(\infty)$  von unendlich hohem Zusammenhang, so ist  $J_c$  eine Cantormenge, d.h. eine perfekte, in  $C$  nirgends dichte Menge. Diese Dychotomie führt zu

**Definition 2** Die Menge

$$M = \{c \in C \mid J_c \text{ ist zusammenhängend}\}$$

heißt Mandelbrotmenge.

Die Mandelbrotmenge scheint auch ein Fraktal zu sein (vgl. Abbildung 3). Sie ist eines der wichtigsten Fraktale in der Theorie der Juliamengen, da sie in einem sehr präzisen Sinn ein universelles Objekt ist, vgl. die Arbeit von Douady und Hubbard [4].

Dieser Aufsatz kann natürlich keine vollständige Einführung in die fraktale Geometrie und die Theorie der Juliamengen sein. In den Arbeiten von P. Blanchard [2], H. Brolin [3] und anderen über holomorphe dynamische Systeme sind die fehlenden Beweise zu finden. Dasselbe gilt auch für die unbewiesenen Aussagen aus dem Bereich der fraktalen Geometrie. Hier sind besonders die Bücher von K. Falconer [5] und M. Barnsley [1] zu nennen.

## Literatur

- [1] Barnsley, M.F.: *Fractals Everywhere*. New York 1988
- [2] Blanchard, P.: *Complex analytic dynamics on the Riemann sphere*. Bull. Amer. Math. Soc. **11**, 1, 85–141 (1984)
- [3] Brolin, H.: *Invariant sets under iteration of rational functions*. Ark. Mat. **6**, 103–141 (1966)
- [4] Douady, A, et Hubbard, J.H.: *Etude dynamique des polynômes complexes I,II*. Publ. Mat. Orsay 84-02 (1984), 85-02 (1985)
- [5] Falconer, K.: *Fractal Geometry*. Chichester 1990
- [6] Fatou, P.: *Sur les équations fonctionnelles*. Bull. Soc. Math. France **47**, 161–271 (1919), Bull. Soc. Math. France **48**, 33–94 et 208–304 (1920)

- [7] Haeseler, F.v., and Peitgen, H.-O.: *Newton's method and complex dynamical systems*. Acta Appl. Math. **13**, 3-58 (1988)
- [8] Hausdorff, F.: *Dimension und äußeres Maß*. Math. Ann. **79**, 157-178 (1919)
- [9] Hutchinson, J.E.: *Fractals and self-similarity*. Indiana Univ. Math. J. **30**, 713-747 (1981)
- [10] Julia, G.: *Mémoire sur l'itération des fonctions rationnelles*. J. Math. Pures Appl. **8**, 1, 47-245 (1918)
- [11] Mandelbrot, B.B.: *Fractals: Form, Chance, and Dimension*. San Francisco 1977
- [12] Mandelbrot, B.B.: *Fractal aspects of the iteration of  $z \mapsto \lambda z(1-z)$  for complex  $\lambda$  and  $z$* . Ann. New York Acad. Sci **375**, 249-259 (1980)
- [13] Peitgen, H.-O., Saupe, D., and Haeseler, F.v.: *Cayley's problem and Julia sets*. Math. Intelligencer **6**, 2, 11-20 (1984)
- [14] Peitgen, H.-O., and Richter, P.: *The Beauty of Fractals*. Berlin 1985
- [15] Peitgen, H.-O., and Saupe, D.: *The Science of Fractal Images*. Berlin 1988

eingegangen: Januar 1991

**Verfasser:**

Prof. Dr. F. v. Haeseler  
Institut für Dynamische Systeme  
Universität Bremen  
Postfach 33 04 40  
W-2800 Bremen 33  
Germany

ULRICH KRAUSE

# Über positive diskrete dynamische Systeme

---

## 1. Einleitung

Die Definition eines diskreten dynamischen Systems ist äußerst simpel: eine (nicht-leere) Menge  $M$  zusammen mit einer (nicht-trivialen) Selbstabbildung  $T : M \rightarrow M$ . Auch eine der zentralen Aufgaben bei der Beschäftigung mit diskreten dynamischen Systemen läßt sich sofort formulieren, nämlich die Untersuchung des asymptotischen Verhaltens der Bahnen  $\{T^n x \mid n \geq 0\}$  für  $x \in M$ .

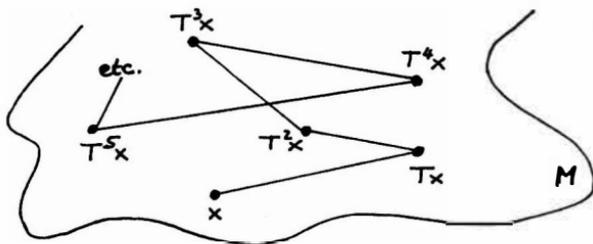


Abbildung 1 Diskretes dynamisches System

Tatsächlich aber ist die mathematische Analyse diskreter Systeme äußerst schwierig. Nur wenige Systeme sind bisher erforscht, und selbst trivial anmutende Fälle haben es in sich. Ein Beispiel eines solchen Systems, dessen Verhalten erst in letzter Zeit aufgeklärt wurde, ist das logistische System mit  $M = [0, 1]$  und  $T_a x = ax(1 - x)$  für einen Parameter  $0 \leq a \leq 4$ . Für gewisse Werte von  $a$  zeigt das System ein einfaches Verhalten, für andere Werte von  $a$  jedoch ein sehr komplexes Verhalten, sogenanntes chaotisches Verhalten. Für Teilmengen  $M$

in höheren Dimensionen, nicht zu reden von unendlichen Dimensionen, kann die analytische Behandlung zu unüberwindlichen Schwierigkeiten führen, falls  $T$  nicht eine lineare Abbildung ist. Angesichts dieser Situation empfiehlt es sich, solche diskreten dynamischen Systeme zu betrachten, die zusätzliche Eigenschaften besitzen. Diese zusätzlichen Eigenschaften sollten einerseits erlauben, die damit ausgestatteten Systeme analytisch befriedigend zu behandeln, sie sollten aber auch andererseits nicht nur mathematische Erfindungen sein, sondern auch tatsächlichen Systemen in Natur, Technik und Gesellschaft zukommen. Aufgrund der letzteren Anforderung genügt es nicht, lineare Systeme zu betrachten. Positive diskrete dynamische Systeme nun, der Gegenstand dieses Artikels, sind solche nichtlinearen diskreten Systeme, deren zusätzliche Eigenschaft mit dem Begriff der Positivität verknüpft sind. Das heißt insbesondere, daß  $M$  ein konvexer Kegel in einem reellen Vektorraum ist und  $T$  ein monotoner Operator hinsichtlich der durch den Kegel induzierten Halbordnung. Es zeigt sich, daß dies allein noch nicht ausreicht, weswegen die Monotonieeigenschaft noch verschärft wird (vgl. Sätze 1 und 2). Daß Positivitätseigenschaften tatsächlich einen Einfluß haben, zeigt bereits das Beispiel der Matrizen Theorie. Die Theorie positiver (nichtnegativer) Matrizen, die sogenannte Perron–Frobenius Theorie (vgl. [14]), erlaubt viel spezifischere Aussagen als die Theorie der Matrizen allgemein, und sie benutzt besondere Hilfsmittel. Außerdem hat die Theorie positiver Matrizen zahlreiche Anwendungen, z.B. in Biologie, Demographie, Wahrscheinlichkeitstheorie, Ökonomie. Ähnlich steht es mit den positiven diskreten dynamischen Systemen, deren Anwendungsbereich noch größer ist, da nichtlineare Beziehungen zugelassen werden. Der Grund für die breite Anwendbarkeit liegt darin, daß viele Variablen wie z.B. die Größe einer Population oder der Preis eines Gutes auf natürliche Weise positiv sind und Anlaß zu monotonen Beziehungen geben. (Vgl. die später behandelten Beispiele.)

Eine Theorie positiver diskreter dynamischer Systeme im eigentlichen Sinne liegt nicht vor, ist jedoch im Entstehen. Das Kernstück einer solchen Theorie wäre die Untersuchung des asymptotischen Verhaltens der Iterierten positiver nichtlinearer Operatoren. Dazu gibt es eine ganze Reihe von Arbeiten, insbesondere das Werk von Krasnosel'skij und seiner Mitarbeiter ([6, 7, 8, 9]). Als weitere und ganz unterschiedliche Quellen einer Theorie positiver diskreter dynamischer Systeme lassen sich Arbeiten von Hirsch und Smith ([4, 5], [17]), von Nussbaum ([15, 16]), sowie die Arbeiten [1, 2, 3] [10, 11, 12, 13] ansehen.

Auf den folgenden Seiten werden als Kostprobe zwei Sätze über positive diskrete dynamische Systeme vorgestellt. Die Beweise sind fortgelassen, auf die Beweisidee jedoch wird kurz eingegangen. Weiterhin wird zu jedem der Sätze ein Beispiel aus der Biologie bzw. der Ökonomie behandelt. Für weitere Aussagen, z.B. über die für Anwendungen wichtige inhomogene Iteration ([3], [13]), sei auf die angegebene Literatur verwiesen.

## 2. Eine konkave Version eines Satzes von Perron

Bereits 1907 bewies (obwohl in etwas anderer Formulierung) Perron den folgenden Satz über positive Matrizen, der allerdings weniger bekannt ist als die übrigen Aussagen der Perron-Frobenius Theorie.

**Satz:** Jede strikt positive Matrix  $A$  besitzt einen eindeutig bestimmten normierten positiven Eigenvektor  $x^*$ , und die normalisierten Iterierten  $\frac{A^k x}{\|A^k x\|}$  konvergieren für jeden von 0 verschiedenen Startvektor  $x$  gegen  $x^*$ , wenn  $k$  gegen  $\infty$  strebt.

Man überlegt sich schnell, daß der Satz von Perron nicht mehr gilt, falls  $A$  nur als irreduzibel vorausgesetzt wird oder falls die Iterierten ohne Normalisierung betrachtet werden. Offensichtlich gilt der Satz von Perron auch für nichtnegative Matrizen  $A$ , falls nur eine Potenz  $A^p$  von  $A$  strikt positiv ist, d.h. für sogenannte *primitive Matrizen*. (Es gibt jedoch Matrizen, die nicht primitiv sind und für die der Satz dennoch gilt, z.B.  $\begin{bmatrix} 1/2 & 0 \\ 1/2 & 1 \end{bmatrix}$ .)

Die Konvergenz der normalisierten Iterierten  $\frac{A^k x}{\|A^k x\|}$  (die übrigens gleich der iterierten Normalisierung  $(\frac{Ax}{\|Ax\|})^k$  ist) wird aus naheliegenden Gründen auch als *relative Stabilität* bzw. *Richtungsstabilität* des zu  $A$  gehörigen dynamischen Systems bezeichnet.

Für das folgende wollen wir einige Bezeichnungen festhalten.  $\mathbb{R}^n = \{x = (x_1, \dots, x_n) \mid x_i \in \mathbb{R} \text{ für alle } i = 1, \dots, n\}$  bezeichne den  $n$ -dimensionalen Euklidischen Raum. Für  $x, y \in \mathbb{R}^n$  schreiben wir  $x \leq y$ , falls  $x_i \leq y_i$ , gilt für alle  $i$ , und  $x < y$ , falls  $x_i < y_i$ , gilt für alle  $i$ . Gilt  $x \leq y$  und ist  $x \neq y$ , so schreiben wir  $x \not\geq y$ . Entsprechende Schreibweisen werden für Matrizen benutzt, z.B. für  $A = (a_{ij})$ ,  $A > 0$  falls  $a_{ij} > 0$  für alle  $i$  und  $j$ .  $K = \mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x \geq 0\}$  bezeichnet den Standardkegel im  $\mathbb{R}^n$ . Bevor wir den Perron'schen Satz auf gewisse nichtlineare Abbildungen ausdehnen werden, hier noch ein Beispiel zum Satz.

### Beispiel: Fibonacci's Kaninchen

Eines der ältesten biomathematischen Modelle geht auf Fibonacci im 13. Jahrhundert zurück und handelt von der Vermehrung der Kaninchen. Dabei wird unterstellt, daß jedes Kaninchenpaar vom zweiten Monat an jeden Monat ein neues Kaninchenpaar hervorbringt. Bezeichnet  $f_n$  die Anzahl der Kaninchenpaare im Monat  $n$  für  $n = 0, 1, 2, \dots$ , so überlegt man sich sofort, daß folgende Rekursion gelten muß:  $f_{n+1} = f_n + f_{n-1}$  für  $n = 1, 2, 3, \dots$  Mit  $f_0 = f_1 = 1$  lassen sich daraus alle  $f_n$  berechnen, die sogenannten *Fibonaccizahlen*  $1, 1, 2, 3, 5, 8, \dots$ . Offensichtlich streben die  $f_n$  gegen  $\infty$ . Wir interessieren uns dafür, wie sich das Verhältnis von Alten (gebärfähigen Kaninchen) zu Jungen (noch nicht gebärfähigen Kaninchen) im Laufe der Zeit entwickelt, d.h., wir interessieren uns für das Verhalten

der Folge der  $\frac{f_k}{f_{k-1}}$ . Dazu ziehen wir den Perron'schen Satz heran und formulieren die Rekursionsgleichung um. Sei  $x^{(k)}$  der (Spalten-) Vektor  $(f_k, f_{k-1})$  und  $A = \begin{bmatrix} 1 & 1 \\ .1 & 0 \end{bmatrix}$ . Die Rekursionsgleichung ist äquivalent mit  $x^{(k+1)} = Ax^{(k)}$ . Es ist  $A^2 > 0$ , also  $A$  primitiv. Wählen wir als Norm  $\|x\| = |x_1| + |x_2|$ , so folgt aus dem Perron'schen Satz, daß

$$\frac{A^{k-1}x^{(1)}}{\|A^{k-1}x^{(1)}\|} = \frac{x^{(k)}}{\|x^{(k)}\|} = \frac{(f_k, f_{k-1})}{f_k + f_{k-1}}$$

gegen  $x^*$  konvergiert für  $k \rightarrow \infty$ . Da  $x^* = \left(\frac{1+\sqrt{5}}{3+\sqrt{5}}, \frac{2}{3+\sqrt{5}}\right)$ , so ergibt sich  $\lim_{k \rightarrow \infty} \frac{f_k}{f_{k-1}} = \frac{1+\sqrt{5}}{2}$ , d.h., das Verhältnis von Alten und Jungen strebt einem festen Wert zu, und zwar dem sogenannten Goldenen Mittel.

Der Satz von Perron läßt sich auf nichtlineare Abbildungen verallgemeinern. Eine besonders einfache Form ist die folgende *konkave Version des Perron'schen Satzes*. Ein Operator  $T: K \rightarrow K$  heißt *konkav*, wenn für alle  $x, y \in K$  und alle  $\lambda \in [0, 1]$  gilt

$$T(\lambda x + (1 - \lambda)y) \geq \lambda Tx + (1 - \lambda)Ty.$$

Ein konkaver Operator ist stets *monoton*, d.h. aus  $0 \leq x \leq y$  folgt  $Tx \leq Ty$ . Es sei  $\|\cdot\|$  eine Norm auf  $\mathbb{R}^n$  die (auf  $K$ ) monoton ist. Der *renormierte Operator*  $\tilde{T}$  ist definiert durch

$$\tilde{T}x = \frac{Tx}{\|Tx\|} \quad \text{für } x \in K \text{ mit } Tx \neq 0.$$

**Satz 1** ([11, 12]) *Es sei  $T: K \rightarrow K$  ein konkaver Operator mit  $Tx > 0$  für  $x \gneq 0$ . Dann hat die Gleichung  $Tx = \lambda x$  für  $x \gneq 0$ ,  $\|x\| = 1$ ,  $\lambda \in \mathbb{R}$  genau eine Lösung  $x^*$ ,  $\lambda^*$ , und für den renormierten Operator gilt  $\lim_{k \rightarrow \infty} \tilde{T}^k x = x^*$  für alle  $x \gneq 0$  (bez. der Euklidischen Topologie).*

**Bemerkung.** Der obige Satz, der natürlich den Satz von Perron als Spezialfall enthält, läßt sich auch auf gewisse nichtlineare positive Operatoren ausdehnen, die einen unendlichdimensionalen Kegel in sich abbilden ([10], [18]). Der Satz läßt sich auch für primitive Operatoren aussprechen.

Wir können Satz 1 sofort auf konkave Verallgemeinerungen von Fibonacci's Populationsmodell anwenden. Später kommen wir darauf zurück. Jetzt wollen wir eine Anwendung von Satz 1 auf ein ökonomisches Problem skizzieren.

**Beispiel: Interdependente Preissetzung** ([1, 11, 12, 13])

Man betrachte eine endliche Menge  $\{1, \dots, n\}$  von Produzenten, von denen Produzent  $i$  genau eine Ware, ebenfalls mit  $i$  bezeichnet, mit Hilfe der von den anderen Produzenten hergestellten Waren und mittels (homogener) Arbeitskraft produziert. Dem Produzenten  $i$  stehe die Technologiemenge  $A_i$  zur Verfügung, die diejenigen inputs  $(a, l)$  angibt, mit deren Hilfe Produzent  $i$  eine Einheit von  $i$  herstellen kann; dabei besteht  $a \in \mathbb{R}_+^n$  aus den materiellen Aufwendungen an Gütern  $1, \dots, n$  und  $l \in \mathbb{R}_+$  ist der Arbeitsinput. Wir wollen annehmen, daß eine vollständige Automatisierung der Technik nicht möglich ist in dem Sinne, daß sich alle Arbeitsaufwendungen durch eine eventuelle Erhöhung der materiellen inputs ersetzen lassen. Genauer, wir nehmen an, daß  $\inf\{l \mid (a, l) \in A_i\} > 0$  ist für jedes  $i$ .

Jeder Produzent wird bei gegebenem Preisvektor  $p \in \mathbb{R}_+^n$  aufgrund der Kongruenz seine Kosten zur Produktion einer Einheit minimieren. Die *Minimalkosten* des Produzenten  $i$  sind gegeben durch

$$T_i(p, w) = \inf\{pa + wl \mid (a, l) \in A_i\}$$

( $pa$  = inneres Produkt der Vektoren  $p, a \in \mathbb{R}_+^n$ ). Dabei bezeichnet  $w$  den Lohnsatz, von dem wir der Einfachheit annehmen, daß er vollständig für einen Standardkonsumkorb  $c \in \mathbb{R}_+^n$ ,  $c > 0$ , ausgegeben wird, also  $w = pc$ .  $p \mapsto T_i(p) = T_i(p, pc)$  ist eine konkave Funktion und daher ist  $T: K \rightarrow K$ , definiert durch  $(Tx)_i = T_i(x)$ , ein konkaver Operator. Aufgrund der Annahme nicht vollständiger Automatisierung gilt  $Tp > 0$  für  $p \geq 0$ . Zur Beschreibung der Dynamik in der Zeit  $t = 0, 1, 2, \dots$  ist es naheliegend anzunehmen, daß jeder Produzent für die nächste Periode einen Preis festlegt, der proportional zu seinen gegenwärtigen Kosten ist, also  $p(t+1) = \lambda(t)Tp(t)$ . Die Preise können wir normieren,  $\tilde{p} = \frac{p}{\|p\|}$  mit  $\|p\| = pc$  ( $p \in \mathbb{R}_+^n$ ). In normierten Preisen (relativen Preisen) erhalten wir daher

$$\tilde{p}(t+1) = \frac{p(t+1)}{\|p(t+1)\|} = \frac{Tp(t)}{\|Tp(t)\|} = \frac{T\tilde{p}(t)}{\|T\tilde{p}(t)\|} = \tilde{T}\tilde{p}(t).$$

Aus Satz 1 folgt nun, daß die normierten Preise sich im Laufe der Zeit stabilisieren, sich nämlich dem durch  $Tp^* = \lambda p^*$ ,  $p^* \in \mathbb{R}_+^n$  mit  $p^*c = 1$ , eindeutig bestimmten Gleichgewichtspreis annähern. Obwohl die einzelnen Produzenten unabhängig voneinander ihre Preise festlegen, führt das zu der obigen Stabilität. Diese basiert im wesentlichen darauf, daß die Produzenten via Technologie interdependent sind und sich bei der Preissetzung einheitlich an ihren Kosten orientieren.

Das Stabilitätsresultat, das wir im obigen Beispiel vermöge der konkaven Version von Perrons Satz erhalten haben, hätten wir mit dem ursprünglichen Satz von Perron nicht erhalten können.

### 3. Absolute Konvergenz der Iterierten

Es liegt in der Natur linearer Abbildungen, daß im Satz von Perron nur die relative Konvergenz der Iterierten, nicht aber die absolute Konvergenz, d.h. die Konvergenz der Iterierten selbst, von Interesse ist. In der nichtlinearen Situation von Satz 1 jedoch ist es interessant zu wissen, unter welchen zusätzlichen Bedingungen sogar die Iterierten selbst konvergieren. Der folgende Satz, bei dem es sich um einen Spezialfall eines Ergebnisses in [10] handelt, gibt darauf eine einfache Antwort. (Siehe auch [18].)

**Satz 2** *Es sei  $T : K \rightarrow K$  ein monotoner Operator mit  $Tx > 0$  für  $x \gneq 0$  und derart, daß für ein  $0 \leq r < 1$ , alle  $x \in K$  und alle  $\lambda \in [0, 1]$  gilt  $\lambda^T Tx \leq T(\lambda x)$ . Dann hat  $T$  genau einen Fixpunkt  $x^* \gneq 0$ , und es gilt  $\lim_{k \rightarrow \infty} T^k x = x^*$  für alle  $x \gneq 0$  (bez. der Euklidischen Topologie).*

**Bemerkung.** Da konkave Abbildungen monoton sind, liegt die Einschränkung, die in Satz 2 gegenüber Satz 1 gemacht wird, in der Forderung, daß  $r < 1$  sein soll.

#### Beispiel: Fibonacci's Kaninchen unter Populationsdruck

In Fibonacci's Modell strebt die Anzahl der Kaninchen im Laufe der Zeit über alle Grenzen. Aber in der Natur wachsen die Bäume nicht in den Himmel. (Warum?) Bei anwachsender Population macht sich schließlich der Populationsdruck bemerkbar, z.B. weil das vorhandene Terrain oder die Nahrungsmittel nicht mehr ausreichen. Bei Wachstumsprozessen dieser Art ist eine Wachstumsrate, die sich mit steigendem Niveau verringert, realistischer als eine konstante Wachstumsrate. Das aber bedeutet, daß der zugrundeliegende Prozeß durch ein nichtlineares Modell zu beschreiben ist. Das lineare Fibonacci-Modell konnten wir auch beschreiben durch die Abbildung  $T : K \rightarrow K$ ,  $Tx = (x_1 + x_2, x_1)$ , nämlich  $Tx = Ax = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ . Zur Beschreibung des Populationsdrucks sind konkave Funktionen sehr geeignet, da sie zwar Wachstum beschreiben, aber ein solches, dessen Zuwachsrate immer geringer wird. Eine einfache Beschreibung des Populationsdrucks für Fibonacci's Kaninchen könnte etwa lauten  $Tx = (\sqrt{x_1} + \sqrt{x_2}, \sqrt{x_1})$ . (Für die Anzahl der Kaninchenpaare in Periode  $k+1$  bedeutet dies  $f_{k+1} = \sqrt{f_k} + \sqrt{f_{k-1}}$ ; das heißt, daß jetzt ein Kaninchenpaar vom zweiten Monat an durchschnittlich weniger als ein Kaninchenpaar hervorbringt, und gleichzeitig, daß nicht jedes Paar das gebärfähige Alter erlebt.) Für den obigen Operator  $T$  ist die zweite Iterierte monoton mit  $T^2x > 0$  für  $x \gneq 0$ . Außerdem gilt  $\sqrt{\lambda}Tx \leq T(\lambda x)$ , und daher  $\sqrt{\lambda}T^2x \leq T^2(\lambda x)$ . Aus Satz 2 folgt, daß  $T^2$  genau dann einen Fixpunkt  $x^* \gneq 0$  hat mit  $\lim_{k \rightarrow \infty} T^k x = x^*$  für alle  $x \gneq 0$ . Für  $Tx$  statt  $x$  folgt  $\lim_{k \rightarrow \infty} T^{2k+1}x = x^*$ ; andererseits

folgt aufgrund der Stetigkeit von  $T$ , daß  $\lim_{k \rightarrow \infty} T^{2k+1}x = Tx^*$  ist. Also ist sogar  $Tx^* = x^*$  und  $\lim_{k \rightarrow \infty} T^k x = x^*$  für alle  $x \geq 0$ . Also führt der Populationsdruck dazu, daß die Kaninchenpopulation schließlich einer festen Anzahl zustrebt (die sich aus der Fixpunktgleichung unschwer ermitteln läßt).

#### 4. Zum Beweis

Die Idee beim Beweis der Sätze 1 und 2 ist, daß in beiden Fällen unter den Voraussetzungen an den Operator  $T$  dieser zu einer Kontraktion bezüglich einer gewissen Metrik wird. Alles übrige erledigt dann der Banachsche Fixpunktsatz oder eine seiner Varianten. (Die Beweise unterscheiden sich sehr von Perron's ursprünglichem Beweis; unter Benutzung von Metriken läßt sich ein einfacher Beweis des Perron'schen Satzes geben.) Doch welches sind die Metriken, und wieso wird  $T$  oder der renormierte Operator  $\tilde{T}$  zu einer Kontraktion?

Die folgende Figur illustriert am Beispiel des Operators  $Tx = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} x$  zum linearen Fibonacci-Modell den kontraktiven Charakter von  $T$  bzw.  $\tilde{T}$ , indem die Bahnen unter  $T$  bzw.  $\tilde{T}$  von ausgewählten Punkten angegeben sind.

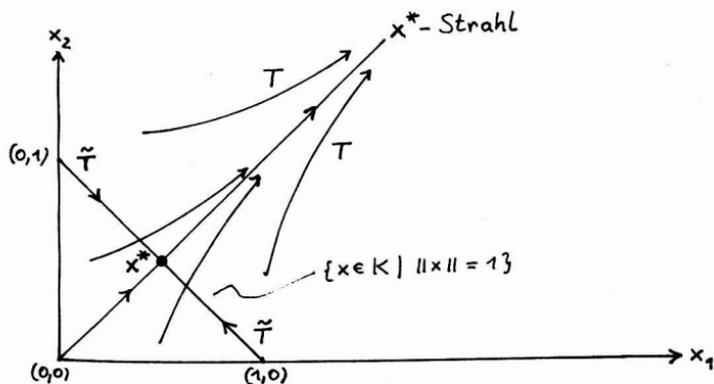


Abbildung 2  $Tx = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} x$

Der durch das (normierte) Gleichgewicht und durch 0 definierte Halbstrahl wird durch  $T$  in sich überführt, nämlich  $T(sx^*) = s(\frac{1+\sqrt{5}}{2})x^*$ . Bahnen von Punkten, die sich auf diesem Halbstrahl befinden, nähern sich ihm asymptotisch an. Es sieht also so aus, als ob der

$x^*$ -Strahl alle T-Bahnen ( $\neq 0$ ) asymptotisch anzieht. Analog zieht  $\tilde{T}$  die Menge der normierten Vektoren auf  $x^*$  zusammen. Die Metrik, die im Falle von Satz 1 ein gutes Maß für das Zusammenziehen ist, ist eine von Hilbert in seinen Untersuchungen zu den Grundlagen der Geometrie verwendete projektive Metrik. Eine geeignete Metrik im Falle von Satz 2 ist eine auf Thompson (vgl. [10]) zurückgehende Metrik, die heute vielfach als Part-Metrik oder auch als Birkhoff-Metrik (vgl. [7]) bezeichnet wird. (Die üblicherweise in Vektorräumen gebräuchlichen Metriken sind nicht geeignet, Operatoren  $T$  der betrachteten Art zu Kontraktionen zu machen. Da hat sich in dieser Hinsicht in das schöne Buch „Fractals Everywhere“ von M. Barnsley auf S. 79, Problem 6.13, ein Fehler eingeschlichen. Die dort betrachtete Abbildung ist im allgemeinen nicht kontraktiv bez. der Euklidischen Metrik, wohl aber bez. Hilbert's projektiver Metrik.)

Beide in den Beweisen der Sätze 1 und 2 benutzte Metriken lassen sich ganz allgemein für einen beliebigen konvexen und abgeschlossenen Kegel (in einem lokalkonvexen Vektorraum), der keine affine Halbgerade enthält, folgendermaßen definieren (Abbildung 3):

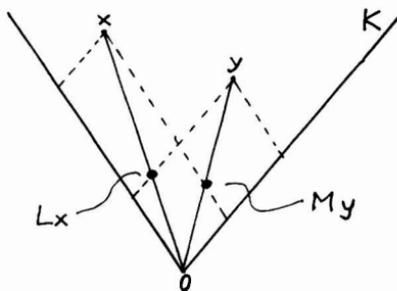


Abbildung 3

Für  $x, y \in K$  schreiben wir  $x \leq y$ , falls  $y - x \in K$ . Definiere für  $x, y \in K \setminus \{0\}$   $L = \sup\{\lambda \in \mathbb{R}_+ \mid \lambda x \leq y\}$ ,  $M = \sup\{\lambda \in \mathbb{R}_+ \mid \lambda y \leq x\}$ . ( $L, M \in \mathbb{R}_+$ , da  $K$  keine affine Gerade enthält). Die projektive Metrik ist dann gegeben durch  $d(x, y) = -\log(L \cdot M)$  und die Part-Metrik durch  $p(x, y) = -\log \min\{L, M\}$ . (Da  $+\infty$  als Wert möglich ist, sind  $d$  und  $p$  tatsächlich nur Pseudo-Metriken.) „Kugeln“ um  $x$  mit dem Radius  $r$  sind im Falle von  $d$  durch die konvexen Kegel  $\{sy \mid s \in \mathbb{R}_+, x \leq y \leq e^r x\}$  und im Falle von  $p$  durch die Intervalle  $\{y \mid e^{-r} x \leq y \leq e^r x\}$  gegeben. Im Falle des Standardkegels  $K = \mathbb{R}_+^n$  lassen sich die Metriken explizit angeben als

$$d(x, y) = \max\{\log x_i - \log y_i \mid 1 \leq i \leq n, x_i > 0\} - \min\{\log x_i - \log y_i \mid 1 \leq i \leq n, y_i > 0\}$$

$$p(x, y) = \max\{|\log y_i - \log x_i| \mid 1 \leq i \leq n, x_i > 0, y_i > 0\}.$$

(Für die Durchführung des Beweises siehe [10, 11, 12].)

## Literatur

- [1] Fujimoto, T., and Krause, U.: *An ergodic property for certain classes of nonlinear operators*. In: Nagel, R., Schlotterbeck, U., and Wolff, M.P.H. (Eds.): *Aspects of Positivity in Functional Analysis*. Proc. of the Conference held on the Occasion of H.H.Schaefer's 60th Birthday, Tübingen, June 24–28, 1985. North-Holland Math. Studies **122**, 191–197, Amsterdam 1986
- [2] Fujimoto, T., and Krause, U.: *Ergodic price setting with technical progress*. In: Semmler, W. (Ed.): *Competition, Instability, and Nonlinear Cycles*. Lecture Notes in Econom. and Math. Systems **275**, 115–124, Berlin 1986
- [3] Fujimoto, T., and Krause, U.: *Asymptotic properties for inhomogeneous iterations of nonlinear operators*. *SIAM J. Math. Anal.* **19**, 841–853 (1988)
- [4] Hirsch, M.: *Attractors for discrete-time monotone dynamical systems in strongly ordered spaces*. In: Alexander, J., and Harer, J. (Eds.): *Geometry and Topology*. Lecture Notes in Math. **1167**, 141–153, Berlin 1985
- [5] Hirsch, M.: *The chaos of dynamical systems*. In: Fischer, P., and Smith, W.R. (Eds.): *Chaos, Fractals and Dynamics*. Lecture Notes in Pure and Appl. Math. **98**, 189–196, New York 1985
- [6] Krasnosel'skij, M.A.: *Positive Solutions of Operator Equations*. Groningen 1964
- [7] Krasnosel'skij, M.A., Vainikko, G.M., Zabreiko, P.P., and Rutitskij, Y.B.: *Approximate Solution of Operator Equations*. Groningen 1972
- [8] Krasnosel'skij, M.A., and Zabreiko, P.P.: *Geometrical Methods of Nonlinear Analysis*. Berlin 1984
- [9] Krasnosel'skij, M.A., Lifshits, J.A., and Sobolev, A.V.: *Positive Linear Systems*. Berlin 1989

- [10] **Krause, U.:** *An nonlinear extension of the Birkhoff-Jentzsch theorem.* J. Math. Anal. Appl. **114**, 552-568 (1986)
- [11] **Krause, U.:** *Perron's stability theorem for non-linear mappings.* J. Math. Econom. **15**, 275-282 (1986)
- [12] **Krause, U.:** *Iteration of concave positive operators with examples from economics and biology.* Universität Tübingen, Semesterberichte Funktionalanalysis **12**, 43-54 (1987)
- [13] **Krause, U.:** *Path stability for prices in a nonlinear Leontief model.* Manuskript 1990 (erscheint demnächst)
- [14] **Minc, H.:** *Nonnegative Matrices.* New York 1988
- [15] **Nussbaum, R.D.:** *Hilbert's projective metric and iterated nonlinear maps.* Mem. Amer. Math. Soc. **391**, Providence, R.I., 1988
- [16] **Nissbaum, R.D.:** *Hilbert's projective metric and iterated nonlinear maps II.* Mem. Amer. Math. Soc. **401**, Providence, R.I., 1989
- [17] **Smith, H.L.:** *Cooperative systems of differential equations with concave nonlinearities.* Nonlinear Anal. **10**, 1037-1052 (1986)
- [18] **Weller, D.:** *Hilbert's metric, part metric and self mappings of a cone.* PhD. Dissertation, Universität Bremen 1987

**eingegangen:** Januar 1991

Prof. Dr. U. Krause  
Fachbereich Mathematik und Informatik  
Universität Bremen  
Postfach 33 04 40  
W-2800 Bremen 33  
Germany

GERHARD BECKER

## Die Bedeutung von Analogien für das Lehren und Lernen von Mathematik

---

Mein Vortrag gliedert sich wie folgt:

1. Zunächst soll an einige Phänomene, Sachverhalte, Zusammenhänge erinnert werden, die uns geläufig sind, in der Absicht, uns darauf aufmerksam zu machen, wie häufig wir in sehr verschiedenen Bereichen geistiger Betätigung Analogien verwenden;
2. soll speziell auf das Lernen von Mathematik eingegangen werden, um insbesondere in diesem Bereich die Bedeutung von Analogiebildungen aufzuzeigen;
3. wird eine Beispielerie die allgemeineren Ausführungen illustrieren.

### 1. Die Bedeutung von Analogien

Unbestritten gehören Analogien zu den fruchtbarsten Ideen und Prinzipien im menschlichen Denken, insbesondere in der Entwicklung der Wissenschaften. Die Ausweitung oder die Bestätigung unseres Wissens, die Überprüfung auf Konsistenz, die Einordnung neu erworbenen Wissens, das Verstehen uns zunächst schwer begreiflicher Phänomene stützen sich häufig auf Analogien.

Hinsichtlich der Mathematik und des mathematischen Unterrichts sei auf Autoren hingewiesen, die in besonderer Weise die Untersuchung des Problemlösens und der mathematischen Entdeckung zum Gegenstand von Untersuchungen gemacht haben: Insbesondere widmet G. Polya ein Kapitel seines Buches *Mathematik und plausible Schließen* [6] der Analyse von Analogien, die Bedeutung von Analogien wird in J. Hadamards ungewöhnlicher Schrift *The Psychology of Invention in the Mathematical Field* immer wieder deutlich; und viele Autoren, die sich auf Publikationen der beiden genannten stützen, führen deren Überlegungen weiter.

Eine Analogie ist eine Entsprechung, eine Korrespondenz zwischen Verhältnissen, Beziehungen. Diese Begriffsbestimmung findet man sinngemäß schon bei Thomas von Aquin. Der

Vagheit, Mehrdeutigkeit dieser Umschreibung entspricht die vielseitige Aufweisbarkeit und Verwendbarkeit von Analogien. Bei einer Analogie liegen zwei inhaltliche Bereiche vor, die im Hinblick auf in ihnen bestehende Beziehungen miteinander verglichen werden. Sucht man eine mathematische Präzisierung für den umgangssprachlichen Begriff Analogie, so wäre wohl am ehesten an den Begriff des Funktors zu denken. Üblicherweise stellt man sich den einen Bereich als bekannt vor, und auf Grund einer vorliegenden Analogie wird der zunächst weniger vertraute zweite Bereich gedanklich erschlossen. Damit erweist sich die Analogie als heuristisches Prinzip von großer Tragweite.

Herausragende Beispiele für die erfolgreiche Verwendung einer Analogiebetrachtung sind etwa die „Entdeckung“ des Benzolrings (Keulè), Newtons Vergleich des um die Erde umlaufenden Mondes mit einem geworfenen Stein; auch die Unterscheidung historischer Epochen, der die Vorstellung in wesentlichen Zügen sich wiederholender Ereignisse zugrundeliegt, die korrespondierende Funktion von Organen in verschiedenen Tierklassen, die Funktion verschiedener Teile in inhaltlich insgesamt disparaten Bereichen (Flügel eines Vogels, eines Gebäudes, einer Partei usw.), Gleichnisse in der Dichtung.

In der Mathematik benutzen wir global die Vorzüge von Entsprechungen bei der Analogie zwischen arithmetischen und geometrischen Beziehungen (Zahlengerade, Zahlenstrahl; Cartesische Ebene) von Mengensprache und Aussagen- bzw. Prädikatenlogik usw. Für die Fachmathematik wie für die Fachdidaktik drückt der Gruppenbegriff als Ordnungsinstrument die Entsprechung zwischen inhaltlich verschiedenartigen Bereichen aus. Abbildungen, Funktionen, Diagramme liefern je nach Zusammenhang mehr globale oder lokale Analogien. Lokal ausgewertete Analogien treten beispielsweise in der Entsprechung zwischen Addition und Multiplikation entgegen, mit der Konsequenz, daß wir Gesetzmäßigkeiten wie Kommutativität, Assoziativität von einer Verknüpfung auf die andere übertragen können. Wir sprechen von Rechenarten 1., 2., 3. Stufe und heben damit die Gemeinsamkeiten und Unterschiede hervor. Punktuelle Analogie kann bestehen zwischen Einzelschritten in Beweisen, Algorithmen, Konstruktionen, Begriffsbildungen.

Auch durchgängige Leitgedanken in der mathematischen Vorgehensweise, in (anlogen) Begriffsbildungen, beispielsweise Geradenspiegelung, Punktspiegelung, Spiegelung an einer Ebene, an einem Kreis, der Funktionsbegriff und das funktionale Prinzip im Mathematikunterricht, Transformationen und Invarianten (Morphismen), Begriffserweiterungen, -verallgemeinerungen, insbesondere bei Zahlenbereichserweiterungen (Permanenzprinzip), Fallunterscheidungen (mit dem Hinweis auf analoge Fälle, deren explizite Behandlung man sich spart), die Berufung auf analoge Schritte in Beweisen, belegen die Bedeutung von Analogien als heuristisches Prinzip und die darauf beruhende Denkökonomie.

## 2. Lernen von Mathematik

Hier sollen nur 2 Aspekte aufgegriffen werden:

- das Neu-Lernen und Verstehen eines neuen Sachverhaltes
- Memorieren, Behalten, Reproduzieren-Können des Verstandenen, das Einordnen in einen vorhandenen Wissensbestand.

Neu-Lernen im Sinne des Problemlösens bedeutet: Neu-Organisieren, Umstrukturieren von Wissen, Zusammenkombinieren zu neuen gedanklichen Gebilden, (subjektiv) neue gedankliche Linien ziehen. Dabei spielt partielle Identität im Lernmaterial, die Übereinstimmung in Teilen eine große Rolle; sie ermöglicht Transfer von Lerninhalten und die Anwendung von Lernprinzipien ebenso wie das Behalten und Abrufen gelernter Inhalte.

Psychologische Untersuchungen haben gezeigt, daß mathematisch besonders Begabte bei der Organisation und Verarbeitung ihres Wissens sich in ungewöhnlicher Weise analoger Denkformen bedienen. (Klix [5], van der Meer [8]).

Wie die Beispiele aus vielfältigen Bereichen des geistigen Lebens zeigen, werden durch Analogien Anstöße und Erklärungsmuster angeboten, zum Teil sogar gewagte; eine Analogiebetrachtung garantiert jedoch keineswegs die Richtigkeit einer so gewonnenen Eigenschaft eines Erkenntnisgegenstandes. Die sachliche Richtigkeit, die Brauchbarkeit einer Analogieübertragung muß in jedem Einzelfall überprüft werden. Analogien eignen sich zur Illustration, sie können kühn, originell, aber auch handfest, bieder, bescheiden sein, sind jedoch keineswegs zwingend. In der Anwendung einer Analogie liegen daher immer zwei gedankliche Leistungen, die begrifflich und im Vollzug unterschieden werden müssen:

1. das Bilden der Analogie, also das Herausfinden von Entsprechungen, Formulieren und Übertragen
2. der Nachweis der Gültigkeit und Tragweite; hierbei können eventuell wieder analoge Betrachtungen einfließen (siehe nachfolgende Beispiele).

Zur Verdeutlichung der beiden unterschiedlichen Schritte und zum Zweck einer erleichterten Übung sollten Schülern im Unterricht auch Gelegenheit gegeben werden, beide Komponenten getrennt zu üben. Neben dem Aufsuchen und dem Formulieren von analogen Übertragungen sollten eigenständige Übungsphasen möglich sein, in denen es an Hand nicht allzu anspruchsvoller Inhalte um die Überprüfung geht. Hierbei hat sich insbesondere die „Homogenität“ eines inhaltlichen Bereiches als vorteilhaft erwiesen. Das bedeutet, daß Gleichartigkeit der Beweistechnik, also Abgegrenztheit der zum Nachweis erforderlichen Mittel, und Klarheit

der Zielkriterien (vgl. Dörner [2], S. 11–14) zu thematischen Einheiten führen, die durch eine vergleichsweise große Zahl miteinander im Zusammenhang stehender Inhalte gekennzeichnet sind; das nachfolgende Beispiel stellt eine solche „Insel“ thematisch gleichartiger Inhalte dar.

### 3. Erläuterungen zu einem Beispiel

Eine Analogie kann bestehen in:

- Reduzierung oder Ausweitung der Komplexität eines Gegenstandes der Betrachtung (Begriff, Satz, Argument, Aufgabe) oder einer Aufgabenstellung
- Verringerung oder Vergrößerung der Anzahl von Ecken, Kanten, Flächen oder anderen Teilen einer geometrischen Figur, der Bestimmungsstücke in einer Aufgabe
- Erhöhung oder Verringerung der Dimension von Teilen oder der ganzen Figur
- Veränderung von Anzahl oder Art der vorkommenden Parameter, demzufolge auch von Konstruktions- oder Beweisschritten
- Vertauschung von Voraussetzung und Behauptung in einem Satz, eventuell die Bildung neuer Kombinationen von Teilvoraussetzungen oder Teilbehauptungen.

Gegenbeispiele zu finden gehört mit zu den Fertigkeiten des Begründens, die Schüler an der vorgestellten thematischen Einheit erwerben sollen. Die vier Einheiten stellen thematische Untereinheiten dar, die auseinander durch Analogieübertragungen hervorgehen. Teil 1 beginnt mit einer Ausgangsfrage, die auf die Formulierung der Dreiecksungleichung und der Strecken-Additivitätsbeziehung bez. der üblichen euklidischen Metrik führt, aus beiden Grundeigenschaften gezogene Schlußfolgerungen für elementare geometrische Figuren schließen sich an; die Begründungen sind teilweise hinzugefügt. In anderem Zusammenhang ist diese thematische Einheit publiziert in Holland [4] oder Schupp [7]. Die nachfolgenden Teile beruhen im wesentlichen auf Änderungen (Erhöhungen) der Dimension(en); dabei liefern Analogiebetrachtungen nicht nur Aussagen über geometrische Objekte, sondern auch – mit sinngemäßen Abänderungen – Begründungen. Das vorgestellte Beispiel zeichnet sich zudem dadurch aus, daß hier Eigenschaften räumlicher Figuren als unmittelbare Weiterführung von Eigenschaften ebener Figuren erkennbar werden; in der Unterrichtspraxis wird ja häufig der Mangel an passenden Beispielen beklagt, die Sachverhalte aus der Ebene mit solchen aus dem Raum in engen Zusammenhang bringen.

Eine analoge Übertragung der drei Ausgangszeilen (Gleichung (1.2), Ungleichungen (1.1) und (1.3)), auf Grund der Erhöhung des Exponenten von 1 auf 2, liefert (Teil 4) eine Charakterisierung von Dreieckstypen, als deren Zusammenfassung der Cosinussatz verstanden werden kann, mit dem Sonderfall des Satzes von Pythagoras. Auch zu diesem Satz und den mit ihm verbundenen Themen gibt es räumliche Analoga, von denen einige im Teil 4 aufgeführt sind.

## Teil 1

Ausgangsfrage ist die Aufgabe, zu gegebenen Streckenlängen  $a, b, c$  und gegebenen Punkten  $A, B$  mit  $L(\overline{AB}) = c$  einen Punkt  $C$  zu konstruieren, für den  $L(\overline{AC}) = b$ ,  $L(\overline{BC}) = a$  sein soll: Ein solcher Punkt  $C$

- existiert nicht für  $c > a + b$ , (1.1)

- existiert und liegt auf der Strecke  $\overline{AB}$  für  $c = a + b$ , (1.2)

- existiert und liegt nicht auf der Strecke  $\overline{AB}$  für  $c < a + b$ . (1.3)

Daraus ergibt sich als Bedingung dafür, daß  $C$  auf der Strecke  $\overline{AB}$  liegt, die Gleichung

$$L(\overline{AB}) = L(\overline{AC}) + L(\overline{BC}), \quad (1.2')$$

während die Dreiecksungleichung

$$L(\overline{AB}) < L(\overline{AC}) + L(\overline{BC}) \quad (1.3')$$

den Fall beschreibt, daß  $C$  nicht auf  $\overline{AB}$  liegt.

Im folgenden wird die Länge einer Strecke  $\overline{AB}$  mit  $|\overline{AB}|$  bezeichnet. Mit den Ausgangszeilen (1.2') und (1.3') ergibt sich:

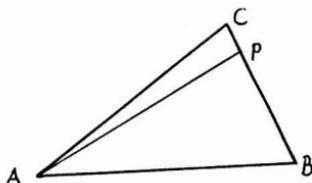


Abbildung 1

Für einen beliebigen Punkt  $P \in \overline{BC}$  ( $P \neq B, C$ ) gilt:

$$\begin{aligned} |\overline{AP}| + |\overline{PB}| &> |\overline{AB}| \\ |\overline{AC}| + |\overline{CP}| &> |\overline{AP}|, & \text{also} \\ |\overline{AC}| + |\overline{CB}| &> |\overline{AP}| + |\overline{PB}| > |\overline{AB}|. \end{aligned} \quad (1.4)$$

Zweimalige Anwendung von (1.4) liefert:

für einen beliebigen Punkt  $Q \in \overline{AP}$  ( $Q \neq A, P$ ) gilt:

$$\begin{aligned} |\overline{AP}| + |\overline{PB}| &> |\overline{AQ}| + |\overline{QB}| \\ |\overline{AC}| + |\overline{CB}| &> |\overline{AP}| + |\overline{PB}|, & \text{also} \\ |\overline{AC}| + |\overline{CB}| &> |\overline{AQ}| + |\overline{QB}| > |\overline{AB}|. \end{aligned} \quad (1.5)$$

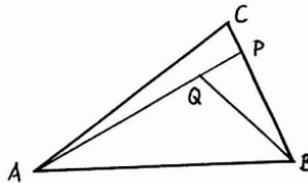


Abbildung 2

Schließlich ergibt dreimalige Anwendung von (1.5):

für einen beliebigen Punkt  $R$  innerhalb des Dreiecks  $ABC$  gilt:

$$\begin{aligned} |\overline{AC}| + |\overline{CB}| &> |\overline{AR}| + |\overline{RB}| > |\overline{AB}| \\ |\overline{BA}| + |\overline{AC}| &> |\overline{BR}| + |\overline{RC}| > |\overline{BC}| \\ |\overline{CB}| + |\overline{BA}| &> |\overline{CR}| + |\overline{RA}| > |\overline{AC}| \quad \text{und} \end{aligned}$$

$$2(|\overline{AB}| + |\overline{BC}| + |\overline{AC}|) > 2(|\overline{AR}| + |\overline{BR}| + |\overline{CR}|) > |\overline{AB}| + |\overline{BC}| + |\overline{AC}|, \quad (1.6)$$

oder, mit den Abkürzungen  $p$  (für den Umfang des Dreiecks) und  $s$  (für die Summe der Abstände des Punktes  $R$  von  $A$ ,  $B$  und  $C$ ):

$$\frac{1}{2}p < s < p \quad (1.6')$$

(vgl. Schupp [7], S. 6-7; Holland [4], S.107-108).

## Teil 2

Analoge Übertragungen auf den Raum beziehen sich auf die Pyramide als die dem Dreieck in der Ebene analoge Figur.

Zunächst werden Aussagen betrachtet, die Summen von Streckenlängen betreffen, wobei die obigen Ergebnisse (1.3) bis (1.6) schrittweise variiert werden.

Trivialerweise gilt (1.3) auch für räumliche Dreiecke.

Betrachten wir drei Strecken anstatt zweier, so gilt für jeden Punkt  $D \notin \Delta ABC$  erst recht:

$$|\overline{AD}| + |\overline{BD}| + |\overline{CD}| > |\overline{AB}| \quad (2.3)$$

Aber, kann (1.3) auch abgeändert werden in

$$|\overline{AD}| + |\overline{BD}| + |\overline{CD}| > |\overline{AB}| + |\overline{AC}| \quad (2.3')$$

für einen beliebigen Punkt  $D \notin \Delta ABC$ ? Hier (wie in den nachfolgenden Überlegungen) braucht  $D$  nicht auf die Ebene von  $\Delta ABC$  beschränkt zu sein, vgl. Abbildung 3.

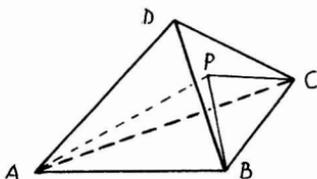


Abbildung 3

(2.3') gilt nicht, wie man durch eine „schmale“ Pyramide zeigt, wobei  $D$  „dicht oberhalb“ des „Grundflächen-Dreiecks“  $ABC$  und „nahe bei“  $B$  und  $C$  liegen.

Hiernach ist

$$|\overline{AD}| + |\overline{BD}| + |\overline{CD}| > |\overline{AB}| + |\overline{BC}| + |\overline{AC}| \quad (2.3'')$$

weniger „wahrscheinlich“; der Gegenbeweis kann durch eine Pyramide mit der Spitze  $D$  „knapp über“  $ABC$  und dem Vergleich mit (1.6) speziell für die Projektion des Punktes  $D'$  von  $D$  in die Ebene des Dreiecks  $ABC$  geführt werden.

Die Absicht, analoge Beziehungen zu (1.4) zu finden, die sich auf die begrenzenden Elemente der betrachteten Figur beziehen und eine Aussage über eine Summe von Streckenlängen darstellen, könnte auf die Aussage führen:

für einen beliebigen Punkt  $P \in \Delta BCD$  ( $P \neq D$ ) ist:

$$|\overline{AD}| + |\overline{BD}| + |\overline{CD}| > |\overline{AP}| + |\overline{BP}| + |\overline{CP}|, \quad (2.4)$$

was durch eine „enge“ Pyramide widerlegt werden kann, deren Spitze  $D$  „knapp über“  $\Delta ABC$  und „weit entfernt“ von  $A$  liegt, während  $B$  und  $D$ , sowie  $P$  „dicht bei“  $C$  liegen.

(1.5), bezogen auf einen inneren Punkt, hat die Entsprechung in:

für einen beliebigen inneren Punkt  $Q$  der Pyramide  $ABCD$  gilt:

$$|\overline{AD}| + |\overline{BD}| + |\overline{CD}| > |\overline{AQ}| + |\overline{BQ}| + |\overline{CQ}|. \quad (2.5')$$

Eine Falsifikation liefert eine „schmale“ Pyramide mit der Spitze  $D$  „dicht über“  $\Delta ABC$ , „nahe bei“  $A$  und  $B$ , und  $Q$  „nahe bei“  $C$ .

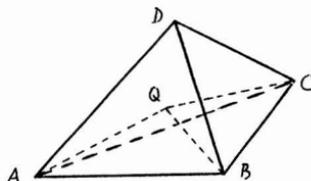


Abbildung 4

Eine andere Eigenschaft der Pyramide, durch analoge Änderungen von (1.5) entstanden, die die Ecken mit einem inneren Punkt in Zusammenhang bringt, ist die folgende:

für einen beliebigen inneren Punkt  $Q$  der Pyramide  $ABCD$  gilt:

$$|\overline{AB}| + |\overline{BC}| + |\overline{AC}| + |\overline{AD}| + |\overline{BD}| + |\overline{CD}| > |\overline{AQ}| + |\overline{BQ}| + |\overline{CQ}| + |\overline{DQ}| \quad (2.5)$$

Dies kann bewiesen werden durch Betrachtung des Schnittpunktes  $S$  der Strecke  $\overline{CD}$  mit der durch  $\Delta ABQ$  bestimmten Ebene:

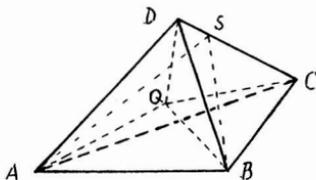


Abbildung 5

Es gilt, gemäß (1.5),  $|\overline{AS}| + |\overline{BS}| > |\overline{AQ}| + |\overline{BQ}|$ , und durch wiederholte Anwendung von (1.3):

$$|\overline{AD}| + |\overline{DS}| > |\overline{AS}| \text{ und } |\overline{SC}| + |\overline{BC}| > |\overline{BS}|, \text{ also } |\overline{AD}| + |\overline{CD}| + |\overline{BC}| > |\overline{AS}| + |\overline{BS}| > |\overline{AQ}| + |\overline{BQ}|, \text{ und entsprechend:}$$

$$|\overline{AC}| + |\overline{AB}| + |\overline{BD}| > |\overline{CQ}| + |\overline{DQ}|; \text{ sowie durch Addition beider Zeilen}$$

$$|\overline{AB}| + |\overline{BC}| + |\overline{AC}| + |\overline{AD}| + |\overline{BD}| + |\overline{CD}| > |\overline{AQ}| + |\overline{BQ}| + |\overline{CQ}| + |\overline{DQ}| \text{ (vgl. (2.5)).}$$

Sechsmalige Anwendung der Dreiecksungleichung (1.3) ergibt

$$|\overline{AQ}| + |\overline{BQ}| > |\overline{AB}|$$

$$|\overline{BQ}| + |\overline{CQ}| > |\overline{BC}|$$

$$|\overline{AQ}| + |\overline{CQ}| > |\overline{AC}|$$

$$|\overline{AQ}| + |\overline{DQ}| > |\overline{AD}|$$

$$|\overline{BQ}| + |\overline{DQ}| > |\overline{BD}|$$

$$|\overline{CQ}| + |\overline{DQ}| > |\overline{CD}|,$$

woraus durch Addition folgt:

$$3(|\overline{AQ}| + |\overline{BQ}| + |\overline{CQ}| + |\overline{DQ}|) > |\overline{AB}| + |\overline{AC}| + |\overline{AD}| + |\overline{BC}| + |\overline{BD}| + |\overline{CD}|, \quad (2.6)$$

dies ist äquivalent zu

$$\frac{1}{3}p < s < p \quad (2.6')$$

wobei mit  $p$  bzw.  $s$  die Summe  $|\overline{AB}| + |\overline{AC}| + |\overline{AD}| + |\overline{BC}| + |\overline{BD}| + |\overline{CD}|$  bzw.  $|\overline{AQ}| + |\overline{BQ}| + |\overline{CQ}| + |\overline{DQ}|$  bezeichnet sind.

### Teil 3

Versuchen wir weiter, analoge Übertragungen von (1.2) und (1.3) zu finden, die sich auf zweidimensionale begrenzte Elemente beziehen, so bietet sich der Flächeninhalt von Dreiecken (im Raum) an.

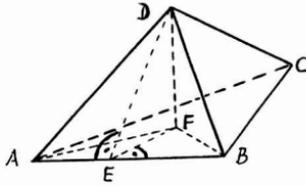


Abbildung 6

Offensichtlich ist wegen  $|\overline{EF}| < |\overline{ED}|$ , und daher auch  $|ABF| < |ABD|$  (wobei  $|ABF|$  den Flächeninhalt von  $\triangle ABF$  bezeichnet usw.):

für einen beliebigen Punkt  $D \in \triangle ABC$ :

$$|ABC| = |ABD| + |BCD| + |ACD| \quad (3.2)$$

und für einen beliebigen Punkt  $D \notin \triangle ABC$ :

$$|ABC| < |ABD| + |BCD| + |ACD|. \quad (3.3)$$

(3.3) könnte man die Pyramidenungleichung nennen.

Auf einem völlig analogen Weg wie von (1.2) und (1.3) zu (1.6) kann nun gefolgert werden:

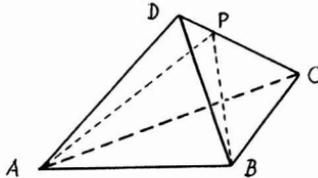


Abbildung 7

für einen beliebigen Punkt  $P \in \overline{CD}$  ( $P \neq C, D$ ):

$|ABD| + |BPD| + |APD| > |ABP|$  und wegen  $|BCP| + |ACP| = |BCP| + |ACP|$  auch

$$|ABD| + |BCD| + |ACD| > |ABP| + |BCP| + |ACP|. \quad (3.4)$$

Anwendung dieses Ergebnisses auf einen Punkt  $Q \in \overline{BP}$  ( $Q \neq B, P$ ), also auch  $Q \in \triangle BCD$ , führt auf die Aussage:

für einen beliebigen Punkt  $Q \in \triangle BCD$  gilt:

$$|ABD| + |BCD| + |ACD| > |ABQ| + |BCQ| + |ACQ|, \quad (3.4')$$

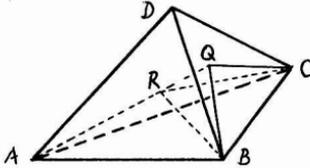


Abbildung 8

und weiter von (3.4') auf einen Punkt  $R \in \overline{AQ}$ :

$$|ABQ| + |BCQ| + |ACQ| > |ABR| + |BCR| + |ACR|;$$

daher, zusammen mit (3.4'):

für einen beliebigen Punkt  $R \in \text{Pyramide } ABCD$  gilt:

$$|ABD| + |BCD| + |ACD| > |ABR| + |BCR| + |ACR| > |ABC|. \quad (3.5)$$

Benutzen wir (3.5) viermal – anstatt nur dreimal, wie oben mit (1.5) – so liefert die Addition dieser vier Zeilen, nämlich

$$|ABD| + |ACD| + |BCD| > |ABR| + |ACR| + |BCR| > |ABC|$$

$$|ABC| + |ACD| + |BCD| > |ABR| + |ADR| + |BDR| > |ABD|$$

$$|ABC| + |ABD| + |BCD| > |ACR| + |ADR| + |CDR| > |ACD|$$

$$|ABC| + |ABD| + |ACD| > |BCR| + |BDR| + |CDR| > |BCD|,$$

$$\begin{aligned} 3(|ABC| + |ABD| + |ACD| + |BCD|) &> 2(|ABR| + |ACR| + |ADR| + \\ &+ |BCR| + |BDR| + |CDR|) \\ &> |ABC| + |ABD| + |ACD| + |BCD|, \end{aligned}$$

was äquivalent ist zu

$$\frac{1}{2}P < S < \frac{3}{2}P, \quad (3.6)$$

wobei  $P$  bzw.  $S$  die Ausdrücke

$|ABC| + |ABD| + |ACD| + |BCD|$  bzw.  $|ABR| + |ACR| + |ADR| + |BCR| + |BDR| + |CDR|$  bezeichnen.

## Teil 4

Schließlich können wir die anfänglichen Zeilen (1.1), (1.2) und (1.3) abändern, unter Berücksichtigung der in den Summanden vorkommenden Exponenten – hier Übergang von 1 auf 2 – und nach der geometrischen Bedeutung der so erhaltenen Zeilen fragen.

Diese, nämlich

$$c^2 > a^2 + b^2, \quad (4.1)$$

$$c^2 = a^2 + b^2, \quad (4.2)$$

$$c^2 < a^2 + b^2 \quad (4.3)$$

charakterisieren ein stumpfwinkliges bzw. ein rechtwinkliges bzw. ein spitzwinkliges Dreieck.

Als eine Art „Zusammenfassung“ von (4.1) bis (4.3) läßt sich der Kosinussatz ansehen:

$$c^2 = a^2 + b^2 - 2ab \cos \gamma \quad (4.4)$$

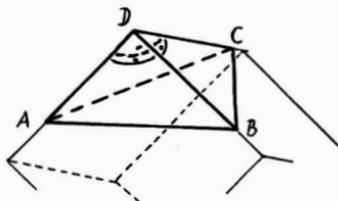


Abbildung 9

Die auf ebene Figuren bezogene Gleichung (4.2) kann auf den Raum übertragen werden, es ergibt sich:

Für eine Pyramide  $ABCD$  mit rechten Winkeln innerhalb der Seitenflächen  $ABD$ ,  $BCD$  und  $ACD$ , jeweils bei der Pyramidenspitze  $D$ , gilt:

$$|ABC|^2 = |ABD|^2 + |BCD|^2 + |ACD|^2 \quad (5.2)$$

Auch eine zu (4.4) entsprechende Gleichung kann wie folgt begründet werden. Ausgehend von der Beziehung

$$F_1 = F_2 \cos(F_1, F_2) + F_3 \cos(F_1, F_3) + F_4 \cos(F_1, F_4)$$

ergibt sich durch einige einfache Zwischenschritte

$$F_1^2 = F_2^2 + F_3^2 + F_4^2 - 2F_1F_2 \cos(F_1, F_2) - 2F_1F_3 \cos(F_1, F_3) - 2F_1F_4 \cos(F_1, F_4). \quad (5.4)$$

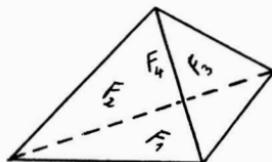


Abbildung 10

Allerdings geht diese Gleichung über das hinaus, was üblicherweise Thema des Mathematikunterrichts in der Sekundarstufe I ist. Analoge Übertragungen stellen, wie auch dieses Beispiel zeigt, häufig Verbindungen zwischen verschiedenen Gegenständen des Mathematikunterrichts her und liefern so Einstiege in weiterführende Themen.

## Literatur

- [1] Becker, G. : *Über den Beitrag des Geometrieunterrichts zum Erwerb heuristischer Strategien.* *mathematica didactica* 10, 123-144 (1987)
- [2] Dörner, D. : *Problemlösen als Informationsverarbeitung.* 2. Auflage, Stuttgart 1979
- [3] Hartkopf, W. : *Erziehung zum heuristisch-methodischen Denken.* *Der Mathematikunterricht* 10, Heft 1, 58-79 (1964)
- [4] Holland, G. : *Vorschläge zur Entwicklung eines Curriculums für den Geometrieunterricht in der Sekundarstufe I.* *Beiträge zum Mathematikunterricht 1972.* Teil I. Hannover 1973, 103-118

- [5] **Klix, F.** : *Analoges Schließen. Kognitive Analyse einer Intelligenzleistung.* In: Ueckert, H., und Rhenius, D. (Ed.): *Komplexe menschliche Informationsverarbeitung.* S. 162–174, Stuttgart 1979
- [6] **Polya, G.** : *Mathematik und plausible Schließen. Bd.I: Induktion und Analogie in der Mathematik.* Stuttgart 1954
- [7] **Schupp, H.** : *Extremwertbestimmungen mit Hilfe der Dreiecksungleichung.* Der Mathematikunterricht **30**, Heft 6, 6–21 (1984)
- [8] **van der Meer, E.** : *Mathematisch-naturwissenschaftliche Hochbegabung.* Zeitschrift für Psychologie **193**, 229–258 (1985)
- [9] **Weihnacht, J.H.** : *Prinzipien zur Lösung mathematischer Probleme.* Braunschweig 1958

**eingegangen:** Januar 1991

**Verfasser**

Prof. Dr. G. Becker  
Universität Bremen  
Fachbereich Mathematik und Informatik  
Postfach 33 04 40  
W-2800 Bremen 33  
Germany

INGO KÖLBL

## Stochastik mit 15-jährigen Schülern – Ergebnisse eines Unterrichtsversuchs

---

Elemente der Wahrscheinlichkeitsrechnung und mathematischen Statistik wurden im obligatorischen Mathematikunterricht in den Schulen der DDR praktisch nicht behandelt, wenn man von den Diagrammen und Mittelwerten absieht. Für den fakultativen Unterricht in der Abiturstufe existierte ein Rahmenprogramm „Elementare Statistik“, das aber in seiner Umsetzung kaum realisiert wurde. Die mathematische (und auch naturwissenschaftliche und technische) Bildung der Schüler war von der Anlage her in den Konzeptionen und Lehrplänen ausgerichtet auf den Grundlagen des mechanischen Determinismus, d.h., Notwendigkeit und Zufall sowie die Beschreibung stochastischer Systeme wurden gelehrt. Das betrifft auch im großen und ganzen den naturwissenschaftlichen Unterricht, wobei im Physik- und Biologieunterricht der Klasse 10 Betrachtungen zu statistischen Gesetzmäßigkeiten bei der Behandlung der Kernphysik und der Mendelschen Gesetze erfolgen sollten. Meines Erachtens ist es aber notwendig, daß für ein besseres Verständnis der Vorgänge in Natur, Technik und Gesellschaft die mathematischen Verfahren der Stochastik herangezogen werden müssen. Es besteht ein großes Defizit bezüglich stochastischer Bildung in den Schulen. Dabei geht es nicht nur darum, im Mathematikunterricht ein Teilgebiet der Mathematik zu behandeln, das z.Z. gänzlich fehlt, von dem man aber weiß, daß es sehr praxisrelevant ist, sondern es geht vielmehr um einen Beitrag zur Allgemeinbildung der Schüler.

Ziel der Behandlung von Elementen der Stochastik im Mathematikunterricht ist, im Zusammenwirken mit weiteren Unterrichtsfächern, die Herausbildung einer „stochastischen Denkweise“ bzw. die Entwicklung des „stochastischen Denkens“ bei den Schülern. Das ist ein Prozeß, der gerichtet ist auf:

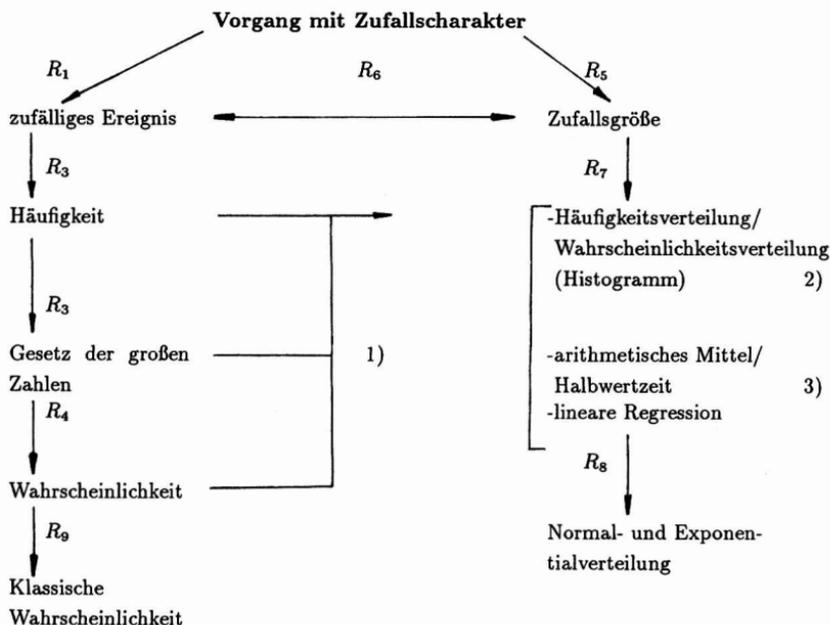
- die Anerkennung der Existenz des Zufalls (durch die Schüler), verbunden mit der Einsicht, daß der Zufall in Natur, Technik und Gesellschaft eine wichtige Rolle spielt (das resultiert daraus, daß der Verlauf von Vorgängen durch statistische Gesetze beschrieben werden kann)

- das Wissen um die Beherrschbarkeit (einschließlich Berechenbarkeit) zufälliger Erscheinungen und das Erkennen von Gesetzmäßigkeiten (das schließt ein, daß Wahrscheinlichkeiten berechnet werden können, ohne daß der Vorgang immer wieder in unabhängigen Wiederholung betrachtet werden muß)
- die Herausbildung von Fähigkeiten, Zufallserscheinungen nach ihren Bedingungen zu analysieren, Interpretationen gewonnener Ergebnisse sowie einen Transfer auf analoge Vorgänge vornehmen zu können (das beinhaltet u.a. auch Vorstellungen über das Stabilwerden relativer Häufigkeiten, das Eintreten sicherer Ereignisse, die Struktur realer Vorgänge und der zugehörigen mathematischen Beschreibung, d.h. die Modellierung solcher Vorgänge auf verschiedenen Stufen der Angepaßtheit, Formalisierung und Komplexität ist zu entwickeln).

Nicht allein der Kalkül der Wahrscheinlichkeitsrechnung und Statistik ist für die Herausbildung einer stochastischen Denkweise bei den Schülern ausschlaggebend, sondern die Betrachtung realer Zufallserscheinungen vor allem in Zusammenhang mit ihrem tatsächlichen Auftreten in den verschiedensten Bereichen unserer Umwelt. Hieraus ergibt sich die Wichtigkeit einer Koordination zwischen dem Mathematikunterricht und insbesondere naturwissenschaftlichen Unterrichtsfächern, damit die Schüler die durch den Mathematikunterricht eingeführten relativ abstrakten Begriffe, Denk- und Arbeitsweisen der Stochastik in ihren Anwendungen erleben können. Das setzt natürlich ein Umdenken im naturwissenschaftlichen Unterricht voraus, denn es müssen Vorgänge mit Zufallscharakter auch als solche behandelt werden.

Unter Beachtung der o. a. Zielbestimmung für die Behandlung von Stochastik im Mathematikunterricht bis zur 10. Klasse entwickelten wir ein Begriffssystem, das als Grundlage für die Erarbeitung eines Lehrplans dient (vgl. Abbildung 1).

Die hier angegebenen Zielvorstellungen für die Klassenstufen 5 bis 10 sind sehr anspruchsvoll und ein „Maximalkatalog“. Weiterführende Untersuchungen müssen Streichungen berücksichtigen. Es besteht z.B. noch keine einheitliche Meinung darüber, ob die Unabhängigkeit von Ereignissen oder die lineare Regression behandelt werden sollen. Viel wird davon abhängen, wie es gelingt, entsprechende, den Schülern auch interessierende, Aufgabenbeispiele zu finden. Schulpraktische Erprobungen müssen hierüber Auskunft geben. Es wird auch auf das methodische Vorgehen beim Unterrichten von Stochastik ankommen.



Rechnen mit Wahrscheinlichkeiten:

- Wahrscheinlichkeit des komplementären Ereignisses
- Additionssatz für unvereinbare Ereignisse
- Multiplikationssatz bei unabhängigen Ereignissen

Schluß von Stichprobe auf die Grundgesamtheit:

- Grundgesamtheit, Element der Grundgesamtheit
- Stichprobe, Stichprobenelement
- Zufallsauswahl (mit gleicher Wahrscheinlichkeit)

Simulation

Abbildung 1

**Legende:**

- R*<sub>1</sub>: Ein Vorgang (mit Zufallscharakter) ist durch Möglichkeiten des Ausgangs gekennzeichnet, diese nennt man zufällige Ereignisse.
- R*<sub>2</sub>: Wiederholung eines Vorgangs – Einstellen von Häufigkeiten der zufälligen Ereignisse.
- R*<sub>3</sub>: Stabilisierung der relativen Häufigkeiten bei unabhängiger Wiederholung des Vorgangs
- R*<sub>4</sub>: Stabiler Wert wird Wahrscheinlichkeit des Vorgangs genannt
- R*<sub>5</sub>: Zahlenmäßige Erfassung der Ausgänge eines Vorgangs, dabei häufig Klasseneinteilung
- R*<sub>6</sub>: Bei einer Zufallsgröße bestehen die zufälligen Ereignisse darin, daß die Zufallsgröße einen bestimmten Zahlenwert bzw. einen Wert aus einer Klasse annimmt.
- R*<sub>7</sub>: Charakterisierung/Beschreibung von Zufallsgrößen
- R*<sub>8</sub>: Bestimmte Arten von Zufallsgrößen haben eine immer wieder auftretende charakteristische Form der Wahrscheinlichkeitsverteilung.
- R*<sub>9</sub>: Sonderfall
  - 1) Ist Voraussetzung für die Behandlung.
  - 2) Beschreibung der Zufallsgröße durch die Wahrscheinlichkeiten der einzelnen Werte/Klassen, die man empirisch aus dem Häufigkeitsdiagramm der Beobachtungsserie erhält. Diese Zuordnung der Häufigkeiten/Wahrscheinlichkeiten zu den entsprechenden Werten/Klassen der Zufallsgrößen nennt man Häufigkeits- bzw. Wahrscheinlichkeitsverteilung der Zufallsgröße.
  - 3) Stabile Werte von arithmetischem Mittel und Halbwertzeit als Kenngrößen der Zufallsgröße.

Die relativ abstrakten Begriffsbildungen, Denk- und Arbeitsweisen der Stochastik erfordern ein abgewogenes und geschicktes Vorgehen durch den Lehrer. Das Arbeiten mit Aufgaben steht im Mittelpunkt, an denen schrittweise und in einem konzentrischen Aufbau (oder Vorgehen) über Schuljahre hinweg theoretische Kenntnisse erreicht werden sollen. Dabei ist eine enge Verbindung zwischen der Behandlung von beschreibender Statistik und Wahrscheinlichkeitsrechnung anzustreben. Die Aufgabenstellungen für die Behandlung von Elementen der Stochastik sind i. allg. in die verschiedenen Stoffgebiete eingebettet, wobei ab Klasse 5 für Systematisierungen zusammenhängende Stunden zur Verfügung stehen müßten.

Für unsere weitere Arbeit war es natürlich von Interesse, wie Schüler auf stochastische Fragestellungen reagieren, wie sie Aufgaben und Probleme mit Hilfe der Stochastik lösen und wie Begriffe, Denk- und Arbeitsweisen der Stochastik von ihnen angeeignet werden. Da keine Möglichkeit bestand, im obligatorischen Mathematikunterricht der polytechnischen Oberschulen Schulversuche durchzuführen, entschlossen wir uns, innerhalb eines fakultativen Kurses mit Schülern der 9. Klassen ausgewählte Themen der Stochastik zu behandeln. In einem Kurs waren 6 bis 9 Schüler, die aus den verschiedensten Motiven am Unterricht teilnahmen. Diese Motivationen reichten von „etwas anderes lernen als im obligatorischen Unterricht“ (solche Schüler, die die Abiturstufe besuchen wollten) bis „man muß ja einen fakultativen Kurs besuchen und dieser liegt für mich zeitlich günstig“. Es war also nicht so, daß alle Schüler mit großer Begeisterung bei der Sache waren. Insgesamt wurde der Kurs 4mal durchgeführt und für jeden Kurs wurden ca. 30 Unterrichtsstunden geplant.

Folgende Inhalte standen im Mittelpunkt:

- Charakterisierung von Vorgängen mit Zufallscharakter
- absolute und relative Häufigkeit und deren Eigenschaften
- Auswertung von Meßreihen mit Hilfe der relativen Häufigkeit, der relativen Summenhäufigkeit und Häufigkeitsverteilungen
- Wahrscheinlichkeit als stabiler Wert der relativen Häufigkeiten eines Ereignisses in einer langen Reihe von Wiederholungen
- klassische Wahrscheinlichkeit
- Rechnen mit Wahrscheinlichkeiten (Additionssatz, Multiplikationssatz)
- Auswahl einer Stichprobe aus einer Grundgesamtheit
- Arbeiten mit Zufallszahlen und Simulation.

Insbesondere wurden bei der Durchführung der Kurse und ihrer Auswertung 4 Fragestellungen in den Mittelpunkt gestellt:

1. Wie reagieren 15-jährige Schüler auf stochastische Fragestellungen und wie bearbeiten sie Vorgänge mit Zufallscharakter, wenn sie in den vorangegangenen Schuljahren keine „Ausrichtung“ auf eine stochastische Denkweise erhielten?

2. Wie bewältigen die Schüler die mathematische Modellierung von Vorgängen mit Zufallscharakter, insbesondere die Wiedergabe des Problems mit eigenen Worten, den Übergang zur mathematischen Formulierung und die Interpretation des mathematischen Ergebnisses? Erkennen die Schüler die Gleichwahrscheinlichkeit von Ereignissen, wann können der Additionssatz und der Multiplikationssatz angewendet werden?
3. Welche Anwendungsbezüge sprechen die Schüler besonders an?
4. Wie reagieren die Schüler bei der Arbeit mit Zufallszahlen und bei Simulationen?

Bevor ich auf eine Interpretation dieser Zielstellungen eingehe, sollen einige Aufgaben vorgestellt werden, die mit den Schülern bearbeitet wurden.

### 1. Beispiel: Münzwurf und Werfen von Kronenkorken

Mit Hilfe dieser Versuche erlebten die Schüler das Stabilwerden der relativen Häufigkeit. Sie erkannten, daß der sich nicht mehr ändernde Wert als Maß für das Eintreten eines Ereignisses genommen werden kann. Dieser stabile Wert wurde als Wahrscheinlichkeit des betrachteten Ereignisses bezeichnet. In der Auswertung der Versuche wurden aus den Eigenschaften der relativen Häufigkeit die Eigenschaften der Wahrscheinlichkeit abgeleitet. Weiterhin erfolgte die Einführung des klassischen Wahrscheinlichkeitsbegriffs. Im Vergleich zwischen den beiden Versuchen wurde herausgestellt, daß die klassische Berechnung der Wahrscheinlichkeit nur Anwendung finden kann, wenn eine Gleichwahrscheinlichkeit aller möglichen Ausgänge des Vorgangs vorliegt.

### 2. Beispiel: Rechnen mit Wahrscheinlichkeiten:

- In einem Behälter befinden sich 49 Kugeln, davon sind 43 rot und 6 blau. Wir entnehmen dem Behälter zufällig 6 Kugeln auf einmal. Wie groß ist die Wahrscheinlichkeit,
  - a) daß alle 6 entnommenen Kugeln blau sind,
  - b) daß unter den 6 entnommenen Kugeln 4 blaue und 2 rote sind?

Aus diesem Beispiel heraus wurden die Gewinnchancen beim Lotto (6 aus 49 und 5 aus 35) berechnet.

- Wie groß ist die Wahrscheinlichkeit, daß aus 32 gut gemischten Karten eine zufällig herausgegriffene Karte eine Herzkarte oder ein As ist?

An dieser Aufgabe wurde der Additionssatz herausgearbeitet, da man das Ereignis durch verschiedene Überlegungen erhalten kann.

**3. Beispiel:** Aussagen mit Hilfe des Wahrscheinlichkeitsbegriffs  
(Lebensdauer - Auswahl von Elementen)

Bei einem Versuch mit 100 Individuen derselben Art ergaben sich unter annähernd gleichen Lebensbedingungen (Modellannahme) folgende Lebensdauern:

Lebensdauer  $X$  in Monaten    absolute Häufigkeit  $H(x)$

17 bis unter 18	21
18 bis unter 19	36
19 bis unter 20	24
20 bis unter 21	12
21 bis unter 22	5
22 bis unter 23	1
23 bis unter 24	1

Die relativen Häufigkeiten und Summenhäufigkeiten wurden berechnet und in Diagrammen dargestellt. Bei der Auswertung wurden z.B. folgende Fragen diskutiert:

Wie groß ist die Wahrscheinlichkeit, im 21. Monat zu sterben?

Wie groß ist die Wahrscheinlichkeit, bis zum 21. Monat zu sterben bzw. älter als 21 Monate zu werden?

Auf der Grundlage des Beispiels wurden Probleme der Gütekontrolle von Schaltelementen, Glühlampen u.a. besprochen und auf die Bedeutung solcher Untersuchungen und deren mathematische Auswertung eingegangen.

**4. Beispiel:** Simulation einer Geschwisterverteilung in einer Familie mit Hilfe des Münzwurfs, indem folgendes Problem betrachtet wurde:

Ein Ehepaar möchte sich 4 Kinder anschaffen und wissen, wie groß die Wahrscheinlichkeit dafür ist, daß 3 Mädchen und ein Junge geboren werden.

**5. Beispiel:** Simulation einer Ameisenwanderung von 10 Ameisen durch Zufallszahlen vom



Probleme der Erzeugung von Zufallszahlen wurden diskutiert. Das 5. Beispiel wurde auch mit einem Kleincomputer bearbeitet.

Als zusammenfassende Ergebnisse der durchgeführten Kurse können herausgestellt werden:

Die Schüler zeigten trotz unterschiedlicher Motivation der Teilnahme an den Kursen i. allg. Interesse an der Lösung von stochastischen Aufgaben, insbesondere dann, wenn sie selbst experimentierten (z.B. Münzwurf, Kronkorkenwurf), messen (z.B. Meßwertreihen aufstellen), befragen (z.B. Erbkrankheiten, Geschwisteranzahlen) u.a. Tätigkeiten ausführen konnten. Die Berechnung von Gewinnchancen bei Glücksspielen und die Bearbeitung von Qualitätskontrollen und Vererbungsproblemen waren weitere Aufgabenstellungen, die mit großer Aufmerksamkeit von den Schülern angenommen wurden. Sie erkannten, daß mit Hilfe der Stochastik Aussagen über Vorgänge als Ganzes gemacht werden können. Dabei gab es auch „Rückschläge“. Zum Beispiel war die Einsicht vorhanden, daß bei der Ziehung der Lottozahlen Gleichwahrscheinlichkeit für alle Zahlen vorliegt. Als aber die Schüler einen Tip abgeben sollten, richteten sich einige Schüler nicht danach, sondern nach einer Veröffentlichung, in der die absoluten Anzahlen der Ziehung der einzelnen Zahlen über Jahre hinweg angeführt waren, z.B. mit dem Argument, daß wenig gezogene Zahlen mit einer größeren Wahrscheinlichkeit in der nächsten Ziehung kommen werden, da sie ja aufholen müssen.

Zusammenfassend kann bez. der 1. und 3. Fragestellung festgehalten werden:

15-jährige Schüler reagieren positiv auf stochastische Fragestellungen. Sie können sich in eine stochastische Denkweise hineinarbeiten, haben aber bei den konkreten Anwendungen gewisse Schwierigkeiten in der Anerkennung des „Zufälligen“. Es überwiegen dann subjektive Meinungen und Anschauungen. Die Anwendungsbezüge der Stochastik müssen so gewählt werden, daß die Schüler den Nutzen der Anwendung stochastischer Verfahren erkennen. Rein theoretische Betrachtungen, wie z.B. das empirische Gesetz der großen Zahlen, wurden nur teilweise angenommen.

Fragen der mathematischen Modellierung von Vorgängen mit Zufallscharakter (2. Fragestellung) müssen sehr differenziert beantwortet werden. Die Wiedergabe eines Problems mit eigenen Worten und das Herausarbeiten einer (sprachlichen) Struktur des Problems, d.h. die zu betrachteten Größen und ihre Zusammenhänge darzustellen, gelang i. allg. nur mit Hilfe des Lehrers. Schwierigkeiten traten auch immer wieder bei der Angabe von Ergebnismengen, der Verknüpfung von Ereignissen und beim Setzen von Bedingungen (z.B. bei Simulationen) auf. War das Problem erst einmal soweit bearbeitet worden, daß „nur noch gerechnet werden brauchte“, wurde das mathematische Ergebnis schnell gefunden. Die Deutung und Auswertung des mathematischen Ergebnisses bewältigten nicht alle Schüler, da sie es nicht auf den betrachteten Sachverhalt beziehen konnten.

Zum Beispiel: Die Wahrscheinlichkeit, daß bei Kindern in einer Familie Rheuma auftritt betrage  $\frac{2}{3}$ . Was bedeutet das, wenn nur 1 Kind geboren wurde oder wenn 3 Kinder geboren werden?

(Ich würde das als eine „kann“-„muß“-Diskussion, bezogen auf den Einzelfall bezeichnen).

Die Betrachtung der zweiten Fragestellung läßt folgende Schlußfolgerung zu:

Das Bearbeiten von stochastischen Fragestellungen muß kontinuierlich von Klassenstufe zu Klassenstufe wachsen, wobei nicht immer ein mathematischer Kalkül im Mittelpunkt stehen sollte. Es müssen auch pauschale Lösungen zugelassen werden. Die inhaltliche Klarheit des zu bearbeitenden Problems ist notwendig, was vor allem auch bei naturwissenschaftlichen Vorgängen zutrifft. Hierbei ist eine Koordination zu anderen Unterrichtsfächern unbedingt notwendig. Es muß ein konzentrischer Aufbau eines Lehrgangs Stochastik in der Schule über mehrerer Klassenstufen hinweg aufgebaut werden.

Bezüglich der 4. Fragestellung, Arbeiten mit Zufallszahlen und Simulation, kann eingeschätzt werden, daß die Schüler mit Freude und Spaß an die Lösung von Aufgaben herangingen. Zwar war für einige Schüler die Entstehung von Zufallszahlen durch einen Algorithmus uneinsichtig, die Arbeit mit ihnen wurde aber anerkannt, wobei hervorgehoben wurde, daß Manipulationen möglich sein könnten.

Insgesamt haben wir die Durchführung der Kurse mit Schülern 9. Klassen als positiv eingeschätzt. Wir konnten Erfahrungen sammeln im Unterrichten von Stochastik bezüglich des Inhalts und methodischer Vorgehensweisen. Wir sind uns aber auch darüber im klaren, daß dieser kleine Unterrichtsversuch nicht ohne weiteres auf den obligatorischen Unterricht übertragen werden kann. In ihm spielen Faktoren eine Rolle, die in den durchgeführten Kursen nicht auftraten, z.B. Frontalunterricht, Zensuren. Gewisse Schlußfolgerungen können wir aber ziehen für weitere geplante Untersuchungen auf verschiedenen Klassenstufen im obligatorischen Mathematikunterricht, die in den nächsten Jahren erfolgen sollen.

## Literatur

- [1] **Hellmann, R.** : *Zu Fragen der Behandlung stochastischer Sachverhalte im mathematischen und naturwissenschaftlichen Unterricht der allgemeinbildenden polytechnischen Oberschule.* Dissertation A, Universität Rostock 1987
- [2] **Kölbl, I.** : *Theoretisch-konzeptionelle Untersuchungen zu Fragen einer Koordination des Mathematikunterrichts mit Inhalten anderer Unterrichtsfächer, dargestellt am Beispiel der Methode der mathematischen Modellierung von außermathematischen Sachverhalten.* Dissertation B, Universität Rostock 1990
- [3] **Kölbl, I., and Leneke, B.** : *Stochastik im Schulunterricht* Preprint, Mathematik 3/90, Technische Universität „Otto von Guericke“ Magdeburg 1990
- [4] *Stochastik, Teil I und II (Material für einen Schulversuch).* (Verfaßt von: Dr. I. Hilsberg und Dr. E. Warmuth, Humboldt-Universität zu Berlin, Sektion Mathematik) Akademie der Pädagogischen Wissenschaften der DDR, Institut für mathematischen und naturwissenschaftlichen Unterricht, Abteilung Mathematik, Berlin 1987

eingegangen: Januar 1991

### Verfasser:

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

WOLF-DIETER RICHTER

## Eine geometrische Methode in der Stochastik

---

Ansätze zu einer geometrischen Behandlung von Aufgabenstellungen aus den Gebieten der Maß- und Integrationstheorie sowie der darauf aufbauenden Stochastik haben weit zurückliegende Ursprünge. Erinnert sei zunächst an das berühmte Prinzip von Cavalieri und Torricelli.

Ein Körper (eine Fläche) werde mit Hilfe von parallelen Ebenen (Geraden) in durchschnittsfremde Schichten (Streifen) zerlegt. Die nach unbegrenzter Verkleinerung der Schichtdicken (Streifenbreiten) entstehenden ebenen Figuren (Strecken) heißen nach B. Cavalieri [4] die Unteilbaren bzw. Indivisiblen des Körpers (der Fläche). In [4] wurde das aus der Anschauung gewonnene und zum Axiom erhobene Prinzip formuliert, wonach die Volumina (Flächeninhalte) zweier Körper (Flächen) zueinander in demselben Verhältnis stehen, wie die Gesamtheiten ihrer Unteilbaren. Dabei wird davon ausgegangen, daß die Richtung, in der ein Körper (eine Fläche) zerlegt wird, auf das Resultat der Untersuchungen ohne Einfluß ist.

Ein weiterführender Gedanke von E. Torricelli besteht darin, daß die betrachteten Körper (Flächen) nicht beide mit Hilfe von Netzen paralleler Ebenen (Geraden) zerlegt bzw. überdeckt werden, sondern daß ein solches Zerlegungsnetz aus einem (speziellen) System "krummer" Flächen (Linien) bestehen kann, siehe etwa [9]. Damit ist aber auch schon der Gedanke angeregt, zwei in gewisser Weise beliebige, aber vergleichbare Systeme krummer Flächen (Linien) als Überdeckende zuzulassen. Die im folgenden vereinfachend ebenfalls nur für den zwei- bzw. dreidimensionalen Fall formulierte Modifizierung des Prinzips von Cavalieri-Torricelli verwendet speziell zwei gleiche Systeme krummer Flächen (Linien) als Überdeckungen.

Es seien  $A$  eine Teilmenge des  $\mathbb{R}^k$ ,  $k \in \{2, 3\}$ ,

$$S_k(r) = \{(x_1, \dots, x_k) \in \mathbb{R}^k : x_1^2 + \dots + x_k^2 = r\}$$

die Einheitskreislinie bzw. die Einheitskugeloberfläche und  $F(A; r)$  der prozentuale Anteil des Inhaltes (d.h. der Bogenlänge bzw. des Flächeninhaltes) von  $A \cap S_k(r)$  in Bezug auf den Inhalt von  $S_k(r)$ ,  $r > 0$ . Stimmen für zwei Mengen  $A_1$  und  $A_2$  aus demselben  $\mathbb{R}^k$  die Werte  $F(A_1; r)$  und  $F(A_2; r)$  für alle  $r > 0$  überein, so haben  $A_1$  und  $A_2$  den gleichen Inhalt.

Es soll nun kurz die Idee des Beweises für diese Aussage erläutert werden. Wegen der Gleichungen

$$\frac{d}{dr}(\pi r^2) = 2\pi r \quad \text{und} \quad \frac{d}{dr}\left(\frac{4}{3}\pi r^3\right) = 4\pi r^2$$

läßt sich der Inhalt der Sphäre  $S_k(r)$  interpretieren als Ableitung des Inhalts der Kreisfläche bzw. Kugel

$$K_k(r) = \{(x_1, \dots, x_k) \in \mathbb{R}^k : x_1^2 + \dots + x_k^2 < r^2\}, \quad r > 0,$$

nach der Variablen  $r$ . Andererseits läßt sich der Inhalt von  $K_k(r)$  grob gesprochen als „Summe der Inhalte aller  $S_k(v)$  mit  $v \leq r$ “ interpretieren. Genauer kann man sich  $K_k(r)$  als Grenzwert der Vereinigung disjunkter Schichten

$$\{x \in \mathbb{R}^k : r_1^2 \leq x_1^2 + \dots + x_k^2 \leq r_2^2\}, \quad 0 \leq r_1 < r_2,$$

vorstellen, deren Dicken  $r_2 - r_1$  gegen Null streben. Durch geeignete Wahl von Paaren  $(r_1, r_2)$  aus  $[0, r] \times [0, r]$  kann man sich so eine Zerlegung der Menge  $K_k(r)$  erzeugt denken. Der Übergang zur Betrachtung der Mengen  $A_i \cap K_k(r)$ ,  $i \in \{1, 2\}$ ,  $r > 0$  birgt nur dann gewisse Schwierigkeiten in sich, wenn die Berandungen der Mengen  $A_i$  keinen weiteren einschränkenden Bedingungen genügen. Diese Schwierigkeiten werden im weiteren auf mathematischer Grundlage überwunden und bleiben deshalb hier zurückgestellt.

Obwohl sich das Cavalieri'sche Prinzip in vielen Fällen als sehr fruchtbar erwies, schloß seine Anwendung doch nicht gewisse Widersprüche aus [9]. An der Klärung und Überwindung aufgetretener Paradoxa beteiligten sich in hervorragender Weise die Vollender der Grundlegung der Differential- und Integralrechnung, I. Newton und G. W. Leibniz. Dank deren und anderer Leistungen lassen sich heute die von Cavalieri-Torricelli entwickelten Gedanken in einer auch heutigen Vorstellungen genügenden präziseren Sprache reproduzieren [8], [3]. Zum grundlegenden Gedankengut zählt in diesem Zusammenhang auch ein Satz von Fubini [8], welcher unter gewissen Bedingungen die Vertauschbarkeit von Integrationsreihenfolgen ermöglicht und aus welchem das Cavalieri'sche Prinzip nun als Spezialfall folgt. Eine wichtige

Konsequenz ergibt sich daraus im Hinblick auf die oben geschilderte Modifizierung des Prinzips von Cavalieri-Torricelli beispielsweise auch für das Standard-Normalverteilungsintegral

$$\Phi(A) = (2\pi)^{-k/2} \int_A \cdots \int \exp\{-(x_1^2 + \cdots + x_k^2)/2\} dx_1 \cdots dx_k.$$

Dieses läßt sich als mit

$$(2\pi)^{-k/2} \exp\{-(x_1^2 + \cdots + x_k^2)/2\}$$

gewichteter Inhalt der (meßbaren) Menge  $A$  ansehen, wobei die Wichtungsfunktion invariant bezüglich orthogonaler Transformationen (Drehungen um 0) ist. Deshalb ist auch  $\Phi(A)$  invariant bezüglich orthogonaler Transformationen von  $A \cap S_k(r)$ . Diese Gedanken münden in die folgende Indivisiblen-Darstellung für  $\Phi$  bei beliebiger endlicher Dimension  $k$ , welche in [10], [11] auf der Grundlage der entsprechenden maßtheoretischen und analytischen Hilfsmittel hergeleitet wurde.

Es seien  $A$  eine Borel-Menge des  $\mathbb{R}^k$ ,  $k \geq 1$ ,  $\omega$  die gleichmäßige Wahrscheinlichkeitsverteilung auf der Einheitssphäre  $S_k(1)$ ,

$$F(A; v) = \omega(v^{-1}A \cap S_k(1)), \quad v > 0,$$

und  $\Gamma$  die Gammafunktion. Dann gilt

$$\Phi(A) = \frac{\omega_k}{(2\pi)^{k/2}} \int_0^\infty F(A; v) v^{k-1} e^{-v^2/2} dv, \quad (1)$$

wobei

$$\omega_k = 2\pi^{k/2} / \Gamma(k/2)$$

der Oberflächeninhalt der Sphäre  $S_k(1)$  ist.

Die folgenden drei Beispiele der Anwendung dieser Formel für spezielle Mengen  $A$  sollen andeuten, wie man (1) in der Mathematischen Statistik nutzen kann. Im Falle

$$A = A_1(r) = \{x \in \mathbb{R}^k : \|x\|^2 < r\}, \quad r > 0,$$

ist

$$F(A; v) = I_{[0, \sqrt{r}]}(v), \quad v > 0.$$

Damit erhält man nach Differentiation von

$$\Phi(A_1(r)) = \frac{\omega_k}{(2\pi)^{k/2}} \int_0^{\sqrt{r}} v^{k-1} e^{-v^2/2} dv$$

bezüglich  $r$  die Dichte der Chi-Quadrat-Verteilung mit  $k$  Freiheitsgraden:

$$\frac{d}{dr} \Phi(A_1(r)) = \frac{r^{k/2-1} e^{-r/2}}{2^{k/2} \Gamma(k/2)}, \quad r > 0.$$

Dieses Ergebnis widerspiegelt den Sachverhalt, daß

$$\Phi(A_1(r)) = P((X_1, \dots, X_n) \in A_1(r)) = P(X_1^2 + \dots + X_k^2 < r)$$

gilt, wobei  $X_1, \dots, X_n$  unabhängige und identisch standard-normalverteilte Zufallsgrößen sind.

Als zweites Beispiel betrachten wir den Doppelkegel

$$A = A_2(r) = \{x \in \mathbb{R}^{m+n} : \sum_{i=1}^m x_i^2 < r \sum_{i=1}^n x_{m+i}^2\}.$$

Führt man verallgemeinerte Kugelkoordinaten in geeignet gewählten  $m$ - bzw.  $n$ -dimensionalen linearen Teilräumen  $L_1$  und  $L_2$  des  $\mathbb{R}^{m+n}$  ein, sowie Polarkoordinaten in der vom Nullelement und den Resultaten der Orthogonalprojektion von  $x$  in  $L_1$  bzw.  $L_2$  aufgespannten zweidimensionalen Ebene, so erhält man

$$\omega_{m+n} F(A_2(r); v) = \omega_m \omega_n \int_0^{\arctan \sqrt{r}} (\cos \phi)^{n-1} (\sin \phi)^{m-1} d\phi$$

in gleicher Weise für alle  $v > 0$ . Nach Differentiation von

$$\Phi(A_2(r)) = F(A_2(r); 1)$$

bezüglich der Variablen  $r$  ergibt sich die Dichte der Fisher'schen F-Verteilung mit  $(m, n)$  Freiheitsgraden. Die Interpretation dieses Resultats ist analog der im vorangehenden Beispiel.

Der Fall des Kegels

$$A = A_3(r) = \left\{ x \in \mathbb{R}^k : \frac{\|\bar{x}_k \mathbf{1}_k\|}{\|\bar{x}_k \mathbf{1}_k - x\|/\sqrt{k-1}} > r, \bar{x}_k > 0 \right\},$$

$$\mathbf{1}_k = (1, \dots, 1) \in \mathbb{R}^k, \quad \bar{x}_k = (x_1 + \dots + x_k)/k,$$

schließlich läßt sich durch geeignete Drehung der Achse

$$\{\lambda \mathbf{1}_k : \lambda > 0\}$$

des Kegels  $A_3(r)$  auf den vorhergehenden Fall für  $m = 1$ ,  $n = k - 1$  zurückführen, mit dem einzigen Unterschied, daß nun kein Doppelkegel vorliegt sondern ein einfacher Kegel. Für die Interpretation dieses Sachverhaltes ist es nützlich zu bemerken, daß  $A_3(r)$  dem kritischen Bereich des Student-Tests zur Prüfung der Hypothese

$$H_0 := \mu \leq 0 \quad \text{gegen} \quad H_1 : \mu > 0$$

über den Erwartungswert  $\mu$  einer normalverteilten Zufallsgröße entspricht.

Es sei erwähnt, daß man die Funktion  $F(A; \cdot)$  beispielsweise auch für den kritischen Bereich  $A$  eines Tests zur Prüfung einer Hypothese über den Nichtlinearitätsparameter, etwa einer exponentiellen Wachstums- bzw. Regressionsfunktion, berechnen kann. Auf diesem Wege erhält man exakte Verteilungsaussagen auch für Situationen, für die man bisher überwiegend Simulationsergebnisse genutzt hat.

Die angeführten Beispiele illustrieren eindrucksvoll, daß die Darstellungsformel (1) für die Standard-Normalverteilung wesentlich Gebrauch macht von der Eigenschaft von  $\Phi$ , invariant bezüglich orthogonaler Transformationen zu sein. Die Funktion

$$v \longrightarrow v^{k-1} e^{-v^2/2}, \quad v > 0,$$

im Integranden auf der rechten Seite von (1) ging im ersten Beispiel auf sehr einfache Weise in die Rechnungen ein und in den folgenden zwei Beispielen überhaupt nicht. Dieser Umstand legt den Gedanken nahe, daß man für alle endlichdimensionalen sphärischen Verteilungen eine Indivisiblen-Darstellung wie in (1) ableiten kann. Dies ist auch völlig analog zur Vorgehensweise im Falle der Normalverteilung möglich und wurde im Detail in [13] durchgeführt.

In diesem Zusammenhang sei bemerkt, daß jedes beliebige Element aus der größeren Klasse der regulären elliptisch konturierten Verteilungen durch geeignete Zentrierung und Normierung auf eine entsprechende sphärische Verteilung zurückgeführt werden kann: Es habe  $Y_g$  die reguläre ( $\det \Sigma > 0$ ) elliptisch konturierte Dichte

$$\phi^*(y; \mu, \Sigma, g) = C(k, g)(\det \Sigma)^{-1/2} g((y - \mu)\Sigma^{-1}(y - \mu)), \quad y \in \mathbb{R}^k,$$

mit einer nichtnegativen Funktion  $g$  und einer geeigneten Normierungskonstanten  $C(k, g)$  so, daß

$$\frac{1}{C(k, g)} = \int_0^\infty r^{k-1} g(r^2) dr < \infty.$$

Dann hat

$$X_g = (Y_g - \mu)\Sigma^{-1/2}$$

die sphärische Dichte

$$\phi(x; g) = C(k, g)g(\|x\|^2), \quad x \in \mathbb{R}^k.$$

Die zugehörige Wahrscheinlichkeitsverteilung

$$\Phi(A; g) = \int_A \phi(x; g) dx, \quad A \in \mathcal{B}^k,$$

genügt der Darstellung

$$\Phi(A; g) = C(k, g)\omega_k \int_0^\infty F(A; v)v^{k-1}g(v^2)dv. \quad (2)$$

Analog der Vorgehensweise bei der geometrischen Herleitung der Chiquadratdichte erhält man nun

$$\frac{d}{dr} \Phi(A_1(r); g) = \frac{1}{2} C(k, g)\omega_k g(r)r^{k/2-1}, \quad r > 0,$$

als Dichte der Zufallsgröße  $\|X_g\|^2$ . Es erscheint gerechtfertigt, diese Funktion als  $g$ -Verallgemeinerung der Chiquadratdichte mit  $n$  Freiheitsgraden zu bezeichnen.

Zum Abschluß der Betrachtungen zu sphärischen Verteilungen wird noch einmal auf das Beispiel der Menge  $A_3(r)$  eingegangen. Für diesen Kegel hängt  $F(A_3(r); v)$  nicht von  $v, v > 0$ , ab. Aus (2) folgt damit, daß  $\Phi(A_3(r); g)$  nicht von  $g$  abhängt. Das ist gleichbedeutend damit, daß der oben betrachtete Student-Test die gleiche Fehlerwahrscheinlichkeit erster Art für alle zulässigen Funktionen  $g$  besitzt und in diesem Sinne also robust ist.

Die letzten beiden Resultate wurden in der erst unlängst erschienenen Monographie [7] über das sich gegenwärtig stürmisch entwickelnde Gebiet der symmetrischen mehrdimensionalen Verteilungen auf einem anderen Wege als dem hier eingeschlagenen bewiesen.

Neben der Möglichkeit der Verallgemeinerung der Darstellungsformel (1) auf die Klasse der sphärischen Verteilungen erlauben die eingangs angestellten geometrischen Betrachtungen zur Normalverteilung auch eine Übertragung auf eine typisch asymmetrische Situation. Speziell kann in Anlehnung an den Kegel  $A_3(r)$  aus der Definition des kritischen Gebietes beim Student-Test ein etwas abgewandelter Kegel  $A_3^E(r)$  als kritisches Gebiet für einen exakten Erwartungswerttest über zweiparametrig exponentiell verteilter Grundgesamtheit gewählt werden. Für diesen Kegel  $A_3^E(r)$  hat dann eine modifizierte Funktion  $F^E(A_3^E(r); v)$  die zu  $F(A_3(r); v)$  analoge Eigenschaft, von  $v$  unabhängig zu sein. In Übereinstimmung mit der Situation beim Student-Test läßt sich dann auch eine Statistik angeben, welche den Kegel  $A_3^E(r)$  in der üblichen Weise als das kritische Gebiet des Erwartungswerttests erzeugt:

$$A_3^E(r) = \{x \in \mathbb{R}^k : \sqrt{n} \frac{\bar{x}_n - \mu_0}{x_{(n)} - x_{(1)}} > r\}.$$

Hierbei sind  $x_{(n)}$  und  $x_{(1)}$  Realisierungen der maximalen bzw. minimalen Ordnungsstatistiken. Eigenschaften dieses Tests werden in [5] diskutiert.

Eine weitere unmittelbare Anwendungsmöglichkeit der Darstellungsformel (1) ergibt sich bei der numerisch-rechentechnischen Auswertung des Normalverteilungsintegrals

$$\Phi(A) = (2\pi)^{-k/2} \int_A \exp\{-\|x\|^2/2\} dx$$

für verschiedenartige Mengen  $A$ . In der Literatur finden sich zahlreiche unterschiedliche spezifische Methoden zur Lösung dieser Aufgabe für spezielle Klassen von Mengen  $A$ . In [2] wurde für den Fall  $k = 2$  und eine große Klasse von Mengen ein entsprechender, auf (1) basierender allgemeiner mathematischer Algorithmus sowie ein Computerprogramm zur Berechnung von  $\Phi(A)$  entwickelt. Mit dessen Hilfe wurden Quantiltabellen für gewisse Typen von  $A$  berechnet. In diesem Zusammenhang heiße eine Zahl  $\lambda = \lambda(A; q)$  das  $A$ -bezogene (oder gebietsbezogene) Quantil der Ordnung  $q$  der Standard-Normalverteilung  $\Phi$ , falls gilt

$$\Phi(\lambda A) = q.$$

Speziell ist hierbei an Mengen  $A$  gedacht, deren Komplemente positiven Abstand vom Koordinatenursprung haben.

An dieser Stelle sei darauf verwiesen, daß geometrische Betrachtungen in der statistischen Literatur seit langem einen breiten Raum einnehmen. Stellvertretend für viele Zeitschriftenpublikationen und Monographien seien hier nur [6] und [1] genannt. Trotz der vielen vorliegenden Resultate steht jedoch die Entwicklung einer durchgängigen Theorie noch aus, in welcher konsequent die algebraisch-geometrischen sowie die maßtheoretisch-stochastischen Aspekte der „klassischen“ Statistik, soweit sinnvoll, getrennt und damit hervorgehoben werden.

Abschließend soll nun die wahrscheinlichkeitstheoretische Frage nach dem asymptotischen ( $\lambda \rightarrow \infty$ ) Verhalten von  $\Phi(\lambda A)$  für meßbare Mengen  $A$  mit Einheitsabstand 1 vom Koordinatenursprung behandelt werden. Von sehr allgemeinem Interesse sind Aussagen über Bedingungen dafür, daß  $\Phi(\lambda A)$  nicht schneller als eine vorgegebene Funktion mit  $\lambda \rightarrow \infty$  gegen Null strebt. Kenntnisse dieser Art sind beispielsweise von grundlegender Bedeutung bei der Anwendung asymptotischer Methoden zur Quantilapproximation.

Es seien  $\tau \geq 0$ ,  $(\lambda_n)_{n \in \{1, 2, \dots\}}$  eine monoton fallende Nullfolge mit

$$\sup_{n \geq 1} \lambda_n / \lambda_{n+1} < \infty$$

und  $\mathcal{M} \subset [0, 1]$  eine meßbare Menge mit positivem Lebesguemaß. Genügt die Funktion  $F(A; v)$  für alle  $t \in \mathcal{M}$  der Bedingung

$$F(A; \sqrt{1 + t\lambda_n}) \geq C_1 (t\lambda_n)^\tau, \quad n = 1, 2, \dots,$$

dann gilt ([12])

$$\Phi(\lambda A) \geq C_2 \lambda^{k-2-2\tau} \exp\{-\lambda^2/2\}, \quad \lambda \geq \lambda_0 > 0.$$

Geometrische Eigenschaften des Gebietes großer Abweichungen  $\lambda A$  gehen also explizit über die mit der Funktion  $F(A; v)$  zusammenhängende Zahl  $\tau$  sowohl in die Bedingung als auch die Aussage ein. Zum Beispiel kann wieder Formel (1) bzw. die folgende, speziell für Situationen großer Abweichungen gedachte Modifizierung verwendet werden:

$$\Phi(\lambda A) = \frac{\omega_k}{(2\pi)^{k/2}} \lambda^{k-2} e^{-\lambda^2/2} \int_0^\infty F(A; \sqrt{1+2c/\lambda^2}) (1+2c/\lambda^2)^{k/2-1} e^{-c} dc.$$

## Literatur

- [1] Anderson, T.W. : *An Introduction to Multivariate Statistical Analysis. 2nd ed.* New York 1984
- [2] Boldt, U. : *Die Berechnung des zweidimensionalen Normalverteilungsmaßes ebener n-Ecke.* Diplomarbeit, Universität Rostock 1989
- [3] Brehmer, S., and Apelt, H. : *Analysis II.* Mathematik für Lehrer, Band 5. Berlin 1974
- [4] Cavalieri, B. : *Geometrica indivisibilis ... 1635 (Geometrie der unteilbaren Größen.)* (Russisch) Moskau-Leningrad 1940
- [5] Davids, J., and Richter, W.-D. : *Exakter Test für erwartete Lebensdauer bei unbekannter Mindestlebensdauer.* Z. Klin. Med. **46**, 783–785 (1991)
- [6] Durbin, J., and Kendall, M.G. *The geometry of estimation.* Biometrika **38**, 150–158 (1951)
- [7] Fang, K.-T., Kotz, S., and Ng, K.-W.: *Symmetric Multivariate and Related Distributions.* London 1990
- [8] Naas, J., and Schmid, H.-L. : *Mathematisches Wörterbuch.* Leipzig 1972
- [9] Nikiforowski, W.A., and Freiman, L.F. : *Wegbereiter der neuen Mathematik.* Moskau und Leipzig 1978
- [10] Richter, W.-D. : *Laplace-Gauß integrals, Gaussian measure asymptotic behaviour and probabilities of moderate deviations.* Z. Anal. Anwendungen **4** (3), 257–267 (1985)

- [11] **Richter, W.-D.** : *Zur Restgliedabschätzung im integralen Zentralen Grenzwertsatz der Wahrscheinlichkeitstheorie*. Math. Nachr. **135**, 103-117 (1988)
- [12] **Richter, W.-D.** : *Multidimensional domains of large deviations*. In: *Limit Theorems in Probability and Statistics*. Colloquia Mathematica Societas János Bolyai **57**, p. 443-458. Amsterdam 1990
- [13] **Schumacher, J.** : *Große Abweichungen elliptisch konturierter Verteilungen*. Diplomarbeit, Universität Rostock 1989

**eingegangen:** November 1990

**Verfasser:**

Dr. W.-D. Richter

Fachbereich Mathematik

Universität Rostock

Universitätsplatz 1

O-2500 Rostock

Germany

FRIEDRICH LIESE

## Aspekte der Statistik zufälliger Prozesse mit unabhängigen Zuwächsen

---

Ein reeller, auf einem gewissen Wahrscheinlichkeitsraum definierter zufälliger Prozeß  $X(t)$ ,  $0 \leq t \leq T$ , heißt *Prozeß mit unabhängigen Zuwächsen*, falls für beliebige  $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq T$  die Zufallsvariablen  $X(0), X(t_1) - X(0), \dots, X(t_n) - X(t_{n-1})$  unabhängig sind. Ohne wesentliche Einschränkungen nehmen wir im weiteren stets an, daß  $X(0) = 0$  gilt und  $X$  stochastisch stetig ist. Prozesse mit unabhängigen Zuwächsen besitzen eine übersichtliche wahrscheinlichkeitstheoretische Struktur. Sie sind darstellbar als Summe von zwei unabhängigen zufälligen Prozessen  $X_1(t), X_2(t)$ , wobei  $X_1(t)$  ein Gaußprozeß mit unabhängigen Zuwächsen ist, der stetige Realisierungen besitzt und im weiteren als allgemeiner Wienerprozeß bezeichnet wird.  $X_2(t)$  ist ein Limes von gemischten Poissonprozessen (siehe [12]). Im folgenden untersuchen wir als wichtige Spezialfälle den allgemeinen Wienerprozeß und den inhomogenen Poissonprozeß näher.

1.  $a(t), b(t)$ ,  $0 \leq t \leq T$ , seien reelle stetige Funktionen mit  $a(0) = b(0) = 0$ , wobei  $b$  zusätzlich nicht fallend sei. Unter einem *allgemeinen Wienerprozeß* mit der Erwartungsfunktion  $a$  und der Varianzfunktion  $b$  versteht man einen Gaußprozeß mit den Eigenschaften  $W(0) = 0$ ,  $EW(t) = a(t)$ ,  $EW(s)W(t) = \text{Min}(b(s), b(t)) + a(s)a(t)$ . Auf Grund der speziellen Struktur der Kovarianzfunktion und der Stetigkeit von  $a$  und  $b$  läßt sich zeigen, daß die Realisierungen eines solchen Prozesses stets stetig und die Zuwächse unabhängig sind. Interessant ist, daß die Umkehrung dieser Aussage in folgender Weise gültig ist.

*Ist  $W$  ein Prozeß mit stetigen Realisierungen, dessen Zuwächse unabhängig sind, so ist er notwendigerweise ein allgemeiner Wienerprozeß (siehe [12]).*

Allgemeine Wienerprozesse werden zur Beschreibung der Abnutzungen mechanischer Bauteile benutzt. Hierbei geht man von der Vorstellung aus, daß die Abnutzung im Zeitintervall  $(t, t+h)$  von der Vergangenheit unabhängig ist und sich mit dem Zustand zur Zeit  $t$  additiv überlagert. Die Erwartungswertfunktion  $a(t)$  beschreibt die mittlere Tendenz der Abnutzung, während die Varianzfunktion  $b(t)$  die mittleren Schwankungen beschreibt, die wegen der Monotonie von  $b$  mit wachsendem  $t$  immer größer werden. Dies entspricht auch genau den Erfahrungen aus der Praxis.

2. Sei  $\Lambda(t)$  eine nichtfallende stetige Funktion. Ein zufälliger Prozeß  $N(t)$  heißt *inhomogener Poissonprozeß* mit der Intensitätsfunktion  $\Lambda$ , falls  $N(0) = 0$  gilt,  $N$  unabhängige Zuwächse besitzt und für alle  $0 \leq s \leq t$  die Zufallsvariablen  $N(t) - N(s)$  eine Poissonverteilung mit dem Parameter  $\Lambda(t) - \Lambda(s)$  besitzen. Deutet man  $N(t)$  als Anzahl der Kunden, die bis zur Zeit  $t$  in einer Bedienungseinrichtung eintreffen, so ist  $\Lambda(t) - \Lambda(s)$  die mittlere Anzahl der Kunden, die im Zeitintervall  $(s, t)$  eintreffen. Hat  $\Lambda$  eine Ableitung  $\lambda$ , so lassen sich durch periodische Funktionen  $\lambda$  etwa tageszeitlich schwankende Kundenströme beschreiben.

Wir bezeichnen durch  $P_{a,b}$  die im Raum der stetigen Funktionen auf  $[0, T]$  definierte Verteilung eines allgemeinen Wienerprozesses mit der Erwartungswertfunktion  $a$  und der Varianzfunktion  $b$ .  $a$  und  $b$  treten als unendlichdimensionale Parameter auf. In vielen Fällen hat man jedoch Zusatzinformationen, die einen Übergang zu einem parametrisierten Modell  $(a_\vartheta, b_\vartheta)$ ,  $\vartheta \in \Theta \subseteq R^k$ , gestatten. Bei der Konstruktion von Maximum-Likelihoodschätzungen muß zunächst die Frage geklärt werden, unter welchen Voraussetzungen für unterschiedliche  $\vartheta_1, \vartheta_2$  die zugehörigen Verteilungen  $P_{a_{\vartheta_1}, b_{\vartheta_1}}$ ,  $P_{a_{\vartheta_2}, b_{\vartheta_2}}$  gegenseitig absolut stetig sind und welche Struktur die Likelihoodfunktion hat. Aus einem allgemeinen Dichotomiesatz von Hajek und Feldman folgt, daß die Verteilungen der allgemeinen Wienerprozesse als Gaußprozesse stets entweder maßtheoretisch äquivalent (gegenseitig absolut stetig) oder orthogonal sind. Die Frage ist nur, wie man dies an den zugehörigen Funktionen  $a_i, b_i$  ablesen kann,  $i = 1, 2$ . Dieses Problem wurde bereits in [2] gelöst, vgl. auch [6], [10].

a) Existiert ein  $0 < t_0 \leq T$  mit  $b_1(t_0) \neq b_2(t_0)$ , so sind die Verteilungen  $P_{a_1, b_1}$  und  $P_{a_2, b_2}$  orthogonal.

b) Gilt  $b_1 = b_2$ , so sind  $P_{a_1, b_1}$  und  $P_{a_2, b_2}$  genau dann maßtheoretisch äquivalent, wenn eine meßbare Funktion  $\alpha(s)$  mit folgenden Eigenschaften existiert:

$$\int_0^T \alpha^2(s) db_1(s) < \infty, \quad (1)$$

$$\int_0^t \alpha(s) db_1(s) = a_1(t) - a_2(t), \quad 0 \leq t \leq T. \quad (2)$$

Sind die Voraussetzungen (1) und (2) im Falle b) erfüllt, so hat die Dichte  $f$  von  $P_{a_1, b_1}$  bezüglich  $P_{a_2, b_2}$  die Darstellung

$$f(W) = \exp \left\{ \int_0^T \alpha(s) dW(s) - \frac{1}{2} \int_0^T \alpha^2(s) db_1(s) \right\}, \quad (3)$$

wobei das stochastische Integral als  $L_2$ -Integral definiert werden kann, weil der Integrand nicht zufällig ist.

Das Problem, unter welchen Voraussetzungen die Verteilung von inhomogenen Poissonprozessen gegenseitig absolut stetig sind, läßt sich ebenfalls vollständig lösen. Sind  $P_{\Lambda_1}^T, P_{\Lambda_2}^T$  die Verteilungen zweier Poissonprozesse in  $[0, T)$ , so gilt  $P_{\Lambda_1}^T \ll P_{\Lambda_2}^T$  genau dann, wenn eine Funktion  $\lambda(t)$  derart existiert, daß

$$\Lambda_1(t) = \int_0^t \lambda(s) d\Lambda_2(s), \quad 0 \leq t \leq T, \quad (4)$$

$$\int_0^T (\sqrt{\lambda(s)} - 1)^2 d\Lambda_2(s) < \infty. \quad (5)$$

Für einen Beweis dieser Aussage verweisen wir auf [2], [5]. Aus der Ungleichung  $(\sqrt{\lambda} - 1)^2 \leq \lambda + 1$  folgt, daß die linke Seite der Ungleichung (5) höchstens  $\Lambda_1(T) + \Lambda_2(T)$  und damit für  $T < \infty$  stets endlich ist. In diesem Falle gestattet die Dichte  $f$  von  $P_{\Lambda_1}^T$  bezüglich  $P_{\Lambda_2}^T$  die Darstellung

$$\exp \left\{ \int_0^T \ln \lambda(s) dN(s) - \Lambda_1(T) + \Lambda_2(T) \right\}. \quad (6)$$

Ausgehend von der expliziten Darstellung der Dichten in (3) bzw. (6) kann man Eigenschaften der Maximum-Likelihoodschätzungen untersuchen. Wir betrachten zunächst das Problem der Konsistenz dieser Schätzungen für  $T \rightarrow \infty$  und verwenden die allgemeinen, in [3] entwickelten Konsistenzbedingungen. Zur Vereinfachung der Situation sei im weiteren stets vorausgesetzt, daß die Parametermenge ein endliches Teilintervall der reellen Achse ist. Die für die Konsistenz entscheidende Größe ist das Hellingerintegral  $H(P, Q)$  von zwei Verteilungen  $P, Q$ , das folgendermaßen definiert ist

$$H(P, Q) = \int \left[ \frac{dP}{d(P+Q)} \frac{dQ}{d(P+Q)} \right]^{1/2} d(P+Q).$$

Der Hellingerabstand ist dann definiert durch  $D(P, Q) := 1 - H(P, Q)$ . Es sei  $P_{\vartheta}^T, \vartheta \in [a, b]$ , eine dominierte Familie von Verteilungen zufälliger Prozesse. Wir nehmen an, daß der Likelihoodquotient stetig vom Parameter  $\vartheta$  abhängt. Zur Abkürzung setzen wir

$$\mathcal{H}(T, \vartheta, u_1, u_2) = H(P_{\vartheta+u_1/\sqrt{T}}^T, P_{\vartheta+u_2/\sqrt{T}}^T).$$

Entsprechend sei

$$\mathcal{D}(T, \vartheta, u_1, u_2) = D(P_{\vartheta+u_1/\sqrt{T}}^T, P_{\vartheta+u_2/\sqrt{T}}^T)$$

definiert. Wir setzen voraus, daß Konstanten  $c, m$  existieren mit

$$\mathcal{D}(T, \vartheta, u_1, u_2) \leq c(1 + R^m)(u_2 - u_1)^2 \quad (7)$$

für alle  $|u_i| \leq R$  und  $a \leq \vartheta \leq b$ . Weiterhin möge es Konstanten  $c_1, c_2$  und eine Funktion  $g$  auf  $[0, \infty)$  geben, die nicht fallend ist und der Bedingung

$$\lim_{x \rightarrow \infty} x^N e^{-g(x)} = 0 \quad (8)$$

für alle  $N > 0$  genügt, so daß die Hellingerintegrale in folgender Weise abschätzbar sind

$$\mathcal{H}(T, \vartheta, u_1, u_2) \leq c_1 e^{-c_2 g(|u_1 - u_2|)}. \quad (9)$$

Unter diesen und sogar noch unter etwas komplizierter zu formulierenden schwächeren Bedingungen wurde in [3] die Existenz von Konstanten  $b_0, B_0$  nachgewiesen, so daß folgende Exponentialabschätzung gilt

$$P_{\vartheta}^T(\sqrt{T}|\hat{\vartheta}_T - \vartheta| > t) \leq B_0 e^{-b_0 g(t)}, \quad (10)$$

wobei  $\hat{\vartheta}_T$  eine Maximum-Likelihoodschätzung ist. Für  $\varepsilon > 0$  gilt

$$P_{\vartheta}^T(|\hat{\vartheta}_T - \vartheta| > \varepsilon) \leq B_0 \exp\{-b_0 g(\varepsilon T^{1/2})\} \underset{T \rightarrow \infty}{\rightarrow} 0,$$

woraus die Konsistenz von  $\hat{\vartheta}_T$  folgt. Die Konvergenzgeschwindigkeit hängt von der Struktur der Funktion  $g$  und damit letztlich von den Hellingerintegralen auf der linken Seite von (9) ab.

Diese Hellingerintegrale wurden in [5] für Poissonprozesse mit beliebigen Phasenräumen und in [6] für eine sehr allgemeine Klasse von Gaußschen Feldern mit unabhängigen Zuwächsen berechnet. Analoge Ergebnisse für Prozesse auf  $[0, T]$  bzw. Abschätzungen der Hellingerintegrale sind auch in [4, 9, 11] zu finden.

Wir betrachten jetzt den Fall inhomogener Poissonprozesse genauer und nehmen dazu an, daß die Intensitätsfunktion  $\Lambda_{\vartheta}(x)$  stetig nach  $x$  differenzierbar ist, und setzen  $\lambda_{\vartheta}(x) = \Lambda'_{\vartheta}(x)$ . Dann gilt (siehe [5]) mit  $P_{\vartheta_i}^T = P_{\lambda_{\vartheta_i}}^T$  und der oben eingeführten Abkürzung

$$\mathcal{H}(T, \vartheta, u_1, u_2) = \exp\left\{-\int_0^T (\sqrt{\lambda_{\vartheta+u_1/\sqrt{T}}(x)} - \sqrt{\lambda_{\vartheta+u_2/\sqrt{T}}(x)})^2 dx\right\}. \quad (11)$$

Ist jetzt  $\lambda_{\vartheta}$  nach  $\vartheta$  differenzierbar, so existieren unter bestimmten Zusatzvoraussetzungen (siehe Lemma 4.2.2 in [4]) Konstanten  $d_1, d_2$  derart, daß sich das Integral auf der rechten Seite von (11) nach oben durch  $d_2(u_2 - u_1)^2$  und nach unten durch  $d_1(u_2 - u_1)^2$  abschätzen läßt. Damit sind die Voraussetzungen (7) und (9) mit  $g(t) = t^2$  erfüllt, woraus die Konsistenz der Maximumlikelihoodschätzung einschließlich einer Geschwindigkeitsabschätzung folgt.

Nach (6) gilt für den Logarithmus des Likelihoodquotienten von  $P_{\lambda_{\vartheta}}^T$  bezüglich  $P_{\lambda_{\vartheta_0}}^T$  die Darstellung

$$L_T(\vartheta) = \int_0^T \left( \ln \frac{\lambda_{\vartheta}}{\lambda_{\vartheta_0}} \right) dN - \int_0^T (\lambda_{\vartheta} - \lambda_{\vartheta_0}) dt.$$

Hierbei ist  $\vartheta_0$  ein fest gewählter Punkt aus der Parametermenge. Bei fester Realisierung  $N$  maximiert  $\hat{\vartheta}_T$  die Funktion  $L_T(\vartheta)$ . Ist schon bekannt, daß  $\hat{\vartheta}_T$  konsistent ist, d.h. stochastisch gegen den unbekannt Parameter  $\vartheta$  strebt, der zu der Grundgesamtheit gehört, aus der die Stichprobe entnommen wurde, und liegt  $\vartheta$  im Innern des Parameterbereichs, so realisiert  $\hat{\vartheta}_T$  ein lokales Maximum und genügt deshalb der Maximum-Likelihoodgleichung  $\dot{L}_T(\hat{\vartheta}_T) = 0$  oder

$$\int_0^T (\dot{\lambda}_{\hat{\vartheta}_T} / \lambda_{\hat{\vartheta}_T}) dN - \int_0^T \dot{\lambda}_{\hat{\vartheta}_T} dt = 0,$$

wobei  $\dot{\lambda}_{\vartheta}$  die Ableitung von  $\lambda_{\vartheta}$  nach  $\vartheta$  bedeutet. Die linke Seite dieser Gleichung wird nach Potenzen von  $(\hat{\vartheta}_T - \vartheta)$  entwickelt, die wegen der Konsistenz von  $\hat{\vartheta}_T$  klein sind. Dies ergibt mit einem gewissen Restglied  $R_T$

$$\int_0^T (\dot{\lambda}_{\vartheta} / \lambda_{\vartheta}) dN - \int_0^T \dot{\lambda}_{\vartheta} dt + (\hat{\vartheta}_T - \vartheta) \left( \int_0^T (\ddot{\lambda}_{\vartheta} \lambda_{\vartheta} - \dot{\lambda}_{\vartheta}^2) \lambda_{\vartheta}^{-2} dN - \int_0^T \ddot{\lambda}_{\vartheta} dt \right) + R_T = 0$$

oder

$$\frac{1}{\sqrt{T}} \int_0^T (\dot{\lambda}_{\vartheta} / \lambda_{\vartheta}) (dN - \dot{\lambda}_{\vartheta} dt) = \sqrt{T} (\hat{\vartheta}_T - \vartheta) \left( \frac{1}{T} \int_0^T (\ddot{\lambda}_{\vartheta} / \lambda_{\vartheta})^2 dN \right) + \tilde{R}_T$$

mit einem gewissen Restglied  $\tilde{R}_T$ . Unter gewissen Voraussetzungen (siehe [4]) läßt sich zeigen, daß  $\tilde{R}_T$  für  $T \rightarrow \infty$  stochastisch gegen Null konvergiert.

Weil der Prozeß  $N_t - \int_0^t \dot{\lambda}_{\vartheta} ds$  unabhängige Zuwächse hat, ist

$$U_T = \frac{1}{\sqrt{T}} \int_0^T (\dot{\lambda}_{\vartheta} / \lambda_{\vartheta}) (dN - \dot{\lambda}_{\vartheta} dt)$$

interpretierbar als Summe unabhängiger Zufallsvariablen mit Erwartungswert Null. Ist für die Summanden die Lindebergsche Bedingung erfüllt, so konvergiert die Verteilung von  $U_T$  für  $T \rightarrow \infty$  gegen eine Normalverteilung mit Erwartungswert Null und der Varianz

$$\lim_{T \rightarrow \infty} V(U_T) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{(\dot{\lambda}_{\vartheta})^2}{\lambda_{\vartheta}} dt = I(\vartheta). \quad (12)$$

Analog ist der Ausdruck

$$V_T = \frac{1}{T} \int_0^T (\dot{\lambda}_{\vartheta} / \lambda_{\vartheta})^2 dN$$

interpretierbar als arithmetisches Mittel aus unabhängigen Zufallsvariablen, für das ein Gesetz der großen Zahlen gilt. Wegen  $E(dN) = \lambda_{\vartheta} dt$  gilt dann

$$\lim_{T \rightarrow \infty} V_T = I(\vartheta)$$

mit  $I(\vartheta)$  aus (12). Insgesamt konvergiert also die Verteilung von  $T^{1/2}(\hat{\vartheta}_T - \vartheta)$  gegen eine Normalverteilung mit dem Erwartungswert Null und der Varianz  $I^{-1}(\vartheta)$ . Die genauen Voraussetzungen und die strenge Rechtfertigung der oben durchgeführten Schlüsse sind der Monographie [4] zu entnehmen.

Die Ergebnisse über inhomogene Poissonprozesse in [4] wurden in [8] auf Poissonprozesse mit beliebigem Phasenraum verallgemeinert.

Wir untersuchen jetzt das Problem der Parameterschätzung in inhomogenen Wienerprozessen, wobei wir zunächst annehmen, daß nur die Erwartungswertfunktion  $a_\vartheta(t)$  vom unbekanntem Parameter  $\vartheta$  abhängt und  $b(t) = t$  gilt.  $a_\vartheta$  sei nach  $t$  differenzierbar und die Ableitung sei quadratisch integrierbar. Dann sind die zugehörigen Prozesse Lösungen der stochastischen Differentialgleichungen

$$dX_t = a'_\vartheta(t)dt + dW_t, \quad (13)$$

wobei  $W$  ein Wienerprozeß mit der Erwartungswertfunktion Null und der Varianzfunktion  $b(t) = t$  ist. Parameterschätzungen in der allgemeinen stochastischen Differentialgleichung

$$dX_t = a_\vartheta(t, X_t)dt + dW_t, \quad (14)$$

deren Lösungen Diffusionsprozesse sind, wurden intensiv in [1, 4, 11] untersucht. Insbesondere wurden Maximum-Likelihoodschätzungen studiert. Diese Ergebnisse sind unmittelbar auf die inhomogenen Wienerprozesse anwendbar, wenn die Varianzfunktion nicht vom Parameter abhängt. Ist dies nicht der Fall, so sind auch für  $T < \infty$  die zu zwei verschiedenen Parameterwerten gehörigen Verteilungen gegenseitig singulär. Einerseits läßt sich damit prinzipiell der unbekannt Parameter aus einer kontinuierlichen Beobachtung ermitteln. Andererseits ist aber die Likelihoodfunktion sehr irregulär und damit nicht analytisch durch ein glattes Funktional der Realisierungen beschreibbar. Deshalb läßt sich die Maximum-Likelihoodschätzung auch nicht als Lösung einer Gleichung beschreiben und ist schwer analytisch zugänglich. Daher wird in dieser Situation das Modell vergrößert. Man verwendet nur Beobachtungen an den diskreten Punkten  $(kt)/n$ ,  $k = 0, \dots, n$ . Da die Zuwächse der betrachteten Prozesse unabhängig sind, lautet, wenn man zu den Differenzen übergeht, die Likelihoodfunktion in diesem Falle

$$L_\vartheta(x_1, \dots, x_n) = \prod_{i=1}^n \varphi(\Delta_i a_\vartheta, \Delta_i b_\vartheta, x_i), \quad (15)$$

wobei  $\varphi(\mu, \sigma^2, x)$  die Dichte einer Normalverteilung mit Erwartungswert  $\mu$  und der Varianz  $\sigma^2$  ist und für eine beliebige Funktion  $f$

$$\Delta_i f = f\left(\frac{iT}{n}\right) - f\left(\frac{(i-1)T}{n}\right)$$

gesetzt wurde. Um die Darstellung zu vereinfachen, setzen wir im weiteren voraus, daß nur die Varianzfunktion vom Parameter abhängt und die Erwartungswertfunktion Null ist. Entsprechend den allgemeinen Kriterien für die Konsistenz in (7) und (9) sind jetzt die Hellingerintegrale von  $P_{\vartheta_1}^n = N(0, \Delta_1 b_{\vartheta_1}) \times \cdots \times N(0, \Delta_n b_{\vartheta_1})$  und  $P_{\vartheta_2}^n = N(0, \Delta_1 b_{\vartheta_2}) \times \cdots \times N(0, \Delta_n b_{\vartheta_2})$  zu berechnen, wobei  $N(\mu, \sigma^2)$  die Normalverteilung mit dem Erwartungswert  $\mu$  und der Varianz  $\sigma^2$  ist. Eine einfache Rechnung zeigt

$$H(N(0, \sigma_1^2), N(0, \sigma_2^2)) = \left( \frac{\sigma_1 \sigma_2}{\frac{1}{2}(\sigma_1^2 + \sigma_2^2)} \right)^{1/2}.$$

Da das Hellingerintegral von Produktmaßen gleich dem Produkt der einzelnen Hellingerintegrale ist, ergibt sich mit  $d_i(\vartheta_j) = \Delta_i b_{\vartheta_j}$

$$H^2(P_{\vartheta_1}^n, P_{\vartheta_2}^n) = \exp \left\{ \sum_{i=1}^n \frac{1}{2} (\ln d_i(\vartheta_1) + \ln d_i(\vartheta_2)) - \ln \left( \frac{1}{2} (d_i(\vartheta_1) + d_i(\vartheta_2)) \right) \right\}.$$

Ersetzt man jetzt  $\vartheta_i$  durch  $\vartheta + u_i/\sqrt{T}$ , so lassen sich unter gewissen Regularitätsbedingungen an  $b_{\vartheta}(t)$  die Voraussetzungen (7) und (9) und damit die Konsistenz der Maximum-Likelihoodschätzungen nachweisen (siehe [13]). Analog den Betrachtungen für den Poissonprozeß ergibt sich durch Taylorentwicklung der Likelihoodfunktion in (15), daß die Verteilung von  $n^{1/2}(\hat{\vartheta}_n - \vartheta)$  mit  $\hat{\vartheta}_n$  als Maximum-Likelihoodschätzung für  $n \rightarrow \infty$  gegen eine Normalverteilung mit dem Erwartungswert Null und der Varianz

$$\sigma^2 = \left[ \frac{1}{2T} \int_0^T \left( \frac{\partial^2}{\partial \vartheta \partial t} b_{\vartheta}(t) \right)^2 \left( \frac{\partial}{\partial t} b_{\vartheta}(t) \right)^{-2} dt \right]^{-1}$$

konvergiert (siehe [7]).

## Literatur

- [1] Arato, M.: *Über die Schätzung von Parametern von Prozessen, die linearen stochastischen Gleichungen genügen*. Stud. Sci. Math. Hung. **5**, 273–296 (1970)
- [2] Gichman, I.I., and Skorochod, A.V.: *On densities of probability measures in function spaces* (in Russian). Uspekhi Mat. Nauk **21**, 83–152 (1966)
- [3] Ibragimov, I.A., and Khas'minskii, R.Z.: *Statistical Estimation. Asymptotic Theory*. New York 1981

- [4] **Kutoyants, Yu.A.:** *Parameter Estimation for Stochastic Processes*. Berlin 1984
- [5] **Liese, F.:** *Eine informationstheoretische Bedingung für die Äquivalenz unbegrenzt teilbarer Punktprozesse*. Math. Nachr. **70**, 183–196 (1976)
- [6] **Liese, F.:** *Hellinger integrals of Gaussian processes with independent increments*. Stochastics **6**, 81–96 (1982)
- [7] **Liese, F.:** *Konsistenz von Maximum-Likelihoodschätzungen in inhomogenen Wienerprozessen*. Unveröff. Manuskript, Univ. Rostock 1990
- [8] **Lorz, U.:** *Beiträge zur Statistik unbegrenzt teilbarer Felder mit unabhängigen Zuwächsen*. Dissertation, Universität Rostock 1988
- [9] **Memin, J., and Shiryaev, A.N.:** *Distance de Hellinger-Kakutani des lois correspondant a accroissements independants*. Z. Wahrsch. Verw. Gebiete **70**, 67–89 (1985)
- [10] **Newman, C.M.:** *The orthogonality of independent increment processes*. In: Strook, D.W., and Varadhan, S.R.S (Eds.): *Topics in Probability Theory*. Seminar 1971–72, Courant Institute of Math. Sci. New York. pp. 93–111, New York 1973
- [11] **Prakasa Rao, B.L.S.:** *Estimation of the drift for diffusion process*. Statistics **16,2**, 263–275 (1985)
- [12] **Skorochod, A.V.:** *Random Processes with Independent Increments* (in Russian). 2. ed., Moscow 1986
- [13] **Wienke, A.:** *Maximum-Likelihoodschätzungen in transformierten Wienerprozessen*. Jahresarbeit, Fachbereich Mathematik, Universität Rostock 1990

eingegangen: Februar 1991

**Verfasser:**

Prof. Dr. F. Liese

Universität Rostock

Fachbereich Mathematik

Universitätsplatz 1

0-2500 Rostock

Germany

GERHARD OSIUS

# Aspekte statistischer Modellbildung und Auswertung

---

## 1. Einleitung

Es soll hier ein Einblick in diejenigen Arbeitsschwerpunkte am Institut für Statistik gegeben werden, an denen der Autor selbst oder seine Doktoranden beteiligt waren. Dabei handelt es sich sowohl um theoretische Untersuchungen, die überwiegend aus Fragestellungen zur Modellbildung, Planung und Auswertung von Studien erwachsen sind, als auch um konkrete Datenanalysen. Dementsprechend bestehen die folgenden Ausführungen aus einem theoretischen und einem praktischen Teil.

Im theoretischen Teil wird zunächst die Klasse der *Generalisierten Linearen Modelle* kurz vorgestellt, die den Rahmen der weiteren Ausführungen bilden. Danach betrachten wir das asymptotische Verhalten der Chiquadrat-Anpassungsstatistiken in parametrischen Multinomialmodellen bei wachsendem Freiheitsgrad, wobei die erwarteten Anzahlen pro Zelle auch klein sein dürfen. Dann gehen wir auf das Bootstrap-Verfahren ein, dessen asymptotische Untersuchungen gewisse Verschärfungen des Begriffs der Verteilungskonvergenz erfordern, die anschließend diskutiert werden.

Im praktischen Teil erläutern wir die Problematik einer statistischen Absicherung eines möglichen Einflusses von Umweltbelastungen auf die Gesundheit des Menschen am Beispiel der Säuglingssterblichkeit. Zum Abschluß diskutieren wir etwas ausführlicher die statistische Analyse der Übereinstimmung zweier Diagnose-Verfahren anhand der konkreten Auswertung einer Studie zur Treffsicherheit der Krebsvorsorgeuntersuchung.

## 2. Generalisierte Lineare Modelle

Nelder und Wedderborn [14] haben die *Generalisierten Linearen Modelle* (kurz GLM) als eine neue Klasse von Modellen eingeführt, in deren Rahmen eine gemeinsame Behandlung von

zunächst so unterschiedlich erscheinenden Modellen möglich ist wie z.B. die klassischen linearen Modelle für Normalverteilungen, die log-linearen Modelle für Poissonverteilungen und die Probit- und Logit-Modelle für Binomialverteilungen. Seitdem werden diese Modelle in wachsendem Umfang weiter untersucht und bei statistischen Datenanalysen angewandt. Wegen ihrer großen Bedeutung werden sowohl die Theorie als auch die Anwendung der Generalisierten Linearen Modelle regelmäßig im Rahmen von Lehrveranstaltungen für den Schwerpunkt Statistik im Studiengang Mathematik angeboten. Hierbei ist im Laufe der Jahre einerseits ein Vorlesungsskript entstanden und andererseits wurde ein Programm (GLM2) zur praktischen Datenanalyse in der Programmiersprache PASCAL für Personal Computer entwickelt, welches für die speziellen Bedürfnisse in Lehre und Forschung konzipiert ist.

Die Generalisierten Linearen Modelle sollen hier nur insoweit vorgestellt werden, wie es zum Verständnis des Folgenden notwendig erscheint. Eine ausführliche Darstellung findet man im Standardwerk von Mc Cullagh und Nelder [10].

Den Ausgangspunkt der Betrachtungen bilden  $J$  unabhängige reelle Zufallsvariablen  $Y_j$  (*Zielvariablen*) für Beobachtungen  $j = 1, \dots, J$ , deren Erwartungswert  $\mu_j := E(Y_j)$  unter Verwendung zusätzlicher beobachteter Covariablenvektoren  $\mathbf{x}_j = (x_{j1}, \dots, x_{jS}) \in \mathbb{R}^S$  wie folgt modelliert wird

$$\mu_j = G(\mathbf{x}_j^T \boldsymbol{\theta}) = G(\theta_1 x_{j1} + \dots + \theta_S x_{jS}) \quad \text{für alle } j. \quad (1)$$

Hierbei ist  $G$  eine fest vorgegebene reelle Funktion und  $\boldsymbol{\theta} \in \mathbb{R}^S$  ein unbekannter Parameter. Unter Verwendung der inversen Funktion  $g = G^{-1}$  (*Linkfunktion*) läßt sich das Modell äquivalent beschreiben durch

$$\eta_j := g(\mu_j) = \mathbf{x}_j^T \boldsymbol{\theta} \quad (\text{linearer Prediktor}) \quad \text{für alle } j, \quad (2)$$

d.h., der Link-transformierte Erwartungswert ist eine *lineare* Funktion des Parameters  $\boldsymbol{\theta}$ . Die Spalten der  $J \times S$  Covariablenmatrix  $\mathbf{X} = (x_{js})$  erzeugen einen linearen Teilraum  $\mathcal{M} = \{\mathbf{X}\boldsymbol{\theta} \mid \boldsymbol{\theta} \in \mathbb{R}^S\}$  des  $\mathbb{R}^J$ . Unter Verwendung des Vektors  $\boldsymbol{\eta} = (\eta_j)$  der linearen Prediktoren läßt sich das Modell (2) auch vektoriell schreiben als

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\theta} \quad \text{bzw.} \quad \boldsymbol{\eta} \in \mathcal{M}. \quad (3)$$

Entsprechend läßt sich (1) darstellen als

$$\boldsymbol{\mu} = \mathbf{G}(\mathbf{X}\boldsymbol{\theta}) \quad \text{bzw.} \quad \boldsymbol{\mu} \in \mathcal{G} := \mathbf{G}[\mathcal{M}], \quad (4)$$

wobei  $\mathbf{G}$  auf  $\mathbb{R}^J$  komponentenweise durch  $G$  definiert ist und der Modellraum  $\mathcal{G}$  für  $\boldsymbol{\mu}$  im allgemeinen *kein linearer* Teilraum ist.

Bei den klassischen linearen Modellen mit normalverteilten  $Y_j$  ist  $G$  und somit auch der Link  $g$  die identische Transformation. Für Poisson-verteilte  $Y_j$  und  $G = \exp$  als Exponentialfunktion bzw.  $g = \log$  als Logarithmus erhält man die log-linearen Modelle.

Eine weitere Teilklasse bilden die *Quantal-Response-Modelle*, bei denen  $Y_j = R_j/N_j$ ; die beobachtete Response-Rate unter  $N_j$  unabhängigen Wiederholungen ist, d.h.,  $R_j$  ist  $B(N_j, \mu_j)$ -verteilt. Unter „Response“ ist hierbei allgemein das Eintreten eines interessierenden Ereignisses zu verstehen (z.B. Tod, Krankheit). Für  $G$  verwendet man Verteilungsfunktionen spezieller stetiger Verteilungen und bezeichnet die resultierenden Modelle nach der zugehörigen Link-Funktion, z.B. Probit-Modelle für  $g$  als Probit-Transformation

$$\text{Probit}(y) := \Phi^{-1}(y) \quad \text{mit } \Phi \text{ als } N(0,1) - \text{Verteilungsfunktion}$$

und Logit-Modelle für  $g$  als Logit-Transformation

$$\text{Logit}(y) := \log(y/(1-y)).$$

Obwohl die Modelle der Form (4) bereits sehr umfangreich sind, erweisen sich weitere Verallgemeinerungen auf nicht notwendig unabhängige Zielvariablen  $Y_j$  als sinnvoll, um z.B. auch die im folgenden behandelten Multinomialmodelle mit einzubeziehen. Hierbei geht man davon aus, daß die Kovarianzmatrix des Zielvektors  $\mathbf{Y}$  von der Form  $\text{Cov}(\mathbf{Y}) = \sigma^2 \mathbf{V}(\boldsymbol{\mu})$  ist, wobei  $\sigma^2 > 0$  ein zusätzlicher reeller Dispersionsparameter und  $\mathbf{V}(\boldsymbol{\mu})$  eine  $J \times J$ -Matrix ist, deren Komponenten *bekannt*e Funktionen  $V_j(\boldsymbol{\mu})$  des Erwartungswertes  $\boldsymbol{\mu}$  sind. Weiterhin ist es in diesem Zusammenhang zweckmäßig, das Modell (1) dahingehend zu erweitern, daß die Funktion  $G$  auch noch vom Beobachtungsindex  $j$  abhängen kann:

$$\mu_j = G_j(\mathbf{x}_j^T \boldsymbol{\theta}) = G_j(\theta_1 x_{j1} + \dots + \theta_S x_{jS}) \quad \text{für alle } j. \quad (5)$$

Die äquivalente Modellformulierung (4) bleibt weiter gültig für  $G = (G_j)$ .

Die Hauptaufgaben einer statistischen Analyse der Beobachtungsdaten  $(Y_j, x_j)$  unter Verwendung obiger Modelle sind dann

- Schätzung des Parameters  $\boldsymbol{\theta}$ , z.B. Maximum-Likelihood Schätzung  $\hat{\boldsymbol{\theta}}$
- Test von (linearen) Hypothesen über  $\boldsymbol{\theta}$ , z.B.  $H_0: \theta_1 = 0$
- Beurteilung der Modellanpassung durch Vergleich des beobachteten Vektors  $\mathbf{Y}$  mit dem geschätzten Erwartungsvektor  $\hat{\boldsymbol{\mu}} = G(\mathbf{X}\hat{\boldsymbol{\theta}})$ .

Für große Anzahlen unabhängiger Beobachtungen (d.h.  $J \rightarrow \infty$ ) kann man unter relativ allgemeinen Bedingungen die Konsistenz und die asymptotische Normalität des Maximum-

Likelihood Schätzers  $\hat{\theta}$  zeigen (Fahrmeier & Kaufmann [7]). Hieraus ergeben sich in bekannter Weise asymptotisch gültige Tests für Hypothesen über dem Parameter  $\theta$ .

Die Modellanpassung kann – neben einer Analyse der Residuen ( $Y_j - \hat{\mu}_j$ ) – in speziellen Situationen auch durch einen Chiquadrat-Anpassungstest überprüft werden, z.B. bei Poissonverteilten Zielvariablen oder in Quantal-Response-Modellen (vgl. z.B. Habermann [9], Bishop, Fienberg, Holland [4]). Die Anwendbarkeit dieses asymptotischen Chiquadrat-Tests ist jedoch nicht immer gesichert, und deshalb wurde in Osius [15] ein auf Normalverteilung basierender Anpassungstest speziell für schwache Besetzungen in Quantal-Response-Modellen vorgeschlagen. Dieser Test wurde von Rojek [23] auf Multinomial-Modelle verallgemeinert und wird im nächsten Abschnitt vorgestellt.

### 3. Anpassungstest für Multinomial-Modelle

Wir betrachten jetzt eine  $J \times K$ -Tafel von beobachteten Anzahlen  $Y_{jk}$ , wobei  $j = 1, \dots, J$  ein Gruppenindex mit zugehörigem Covariablenvektor  $\mathbf{x}_j \in \mathbb{R}^S$  und  $k = 1, \dots, K$  eine disjunkte Klassifikation eines interessierenden Zielereignisses codiert, z.B. einer Lebensdauer  $T$  (d.h. Status  $k$  bedeutet  $t_{k-1} \leq T < t_k$  für vorgegebene Abschnitte  $t_k$ ) oder eines Gesundheitszustandes (z.B. 1=gesund, ..., K=schwer-krank).

	1	...	k	...	K	$\Sigma$
1	$Y_{11}$	...		...	$Y_{1K}$	$N_1$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
j	$Y_{j1}$	...	$Y_{jk}$	...	$Y_{jK}$	$N_j$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
J	$Y_{J1}$	...		...	$Y_{JK}$	$N_J$

Die Zeilen  $\mathbf{Y}_j = Y_{j1}, \dots, Y_{jK}$  seien voneinander unabhängig und jeweils multinomialverteilt mit  $K$  Klassen:  $\mathbf{Y}_j \sim \text{Multi}_K(N_j, \boldsymbol{\pi}_j)$ . Als parameterisches Modell für die Wahrscheinlichkeit  $\pi_{jk}$  von Status  $k$  in Gruppe  $j$  betrachten wir das Modell

$$\pi_{jk} = G_k(\mathbf{x}_j^T \boldsymbol{\theta}), \quad (6)$$

wobei  $G_k$  bekannte Funktionen und  $\boldsymbol{\theta} \in \mathbb{R}^S$  der unbekannte Parameter sind. Dies ist ein Generalisiertes Lineares Modell im Sinne von Abschnitt 2 mit den beobachteten Raten  $\bar{Y}_{jk} = Y_{jk}/N_j$  als Zielvariable, deren Erwartungswert  $E(\bar{Y}_{jk}) = \pi_{jk}$  ist.

Unter Verwendung des Maximum-Likelihood-Schätzers  $\hat{\theta}$  erhält man Schätzungen

$$\hat{\mu}_{jk} = N_j G_k(\mathbf{x}_j^T \hat{\theta}) \quad (7)$$

für die erwarteten Anzahlen  $\mu_{jk} = E(Y_{jk})$ . Zur globalen Beurteilung der Modellanpassung berechnet man meist eine der beiden Statistiken

$$X^2 = \sum_j \sum_k (Y_{jk} - \hat{\mu}_{jk})^2 / \hat{\mu}_{jk} \quad (\text{Pearson - Statistik})$$

$$G^2 = 2 \sum_j \sum_k Y_{jk} \cdot \log(Y_{jk} / \hat{\mu}_{jk}) \quad (\text{Likelihood - Ratio - Statistik}).$$

Beide Statistiken gehören zu einer von Cressi und Read[5] vorgeschlagenen Familie von *Power-Divergenzen* der Form

$$SD_\lambda = \sum_j \sum_k a_\lambda(Y_{jk}, \hat{\mu}_{jk}) \quad (\text{Sum of Deviations})$$

mit Abstandsfunktionen  $a_\lambda$ , die für jedes  $\lambda \in \mathbb{R}$  definiert sind als

$$a_\lambda(y, \mu) = \frac{2y}{\lambda(\lambda+1)} \left[ \left( \frac{y}{\mu} \right)^\lambda - 1 \right] - \frac{2}{\lambda+1} [y - \mu] \geq 0.$$

Die zunächst undefinierten Fälle  $\mu = 0$ ,  $\lambda = -1$  sind durch stetige Fortsetzungen  $\lambda \rightarrow 0$ ,  $\lambda \rightarrow -1$  erklärt. Für  $\lambda = 1$  bzw.  $\lambda = 0$  ergibt sich  $SD_1 = X^2$  bzw.  $SD_0 = G^2$ .

Ein formaler, auf  $SD_\lambda$  beruhender Anpassungstest benutzt die asymptotische Verteilung von  $SD_\lambda$ , die allerdings von der betrachteten Asymptotik abhängt. Die klassische „Fixed-Cells“ Asymptotik geht davon aus, daß  $J$  und  $K$  fest bleiben, während alle Umfänge  $N_j \rightarrow \infty$  im wesentlichen proportional anwachsen. Unter zusätzlichen Regularitätsbedingungen ergibt sich, daß alle  $SD_\lambda$  asymptotisch äquivalent sind und eine zentrale  $\chi^2$ -Verteilung mit Freiheitsgrad  $J - S$  besitzen. Bei einer praktischen Anwendung des entsprechenden  $\chi^2$ -Anpassungstests sollten die erwarteten Anzahlen  $\hat{\mu}_{jk}$  „groß genug“ sein, damit die Asymptotik  $N_j \rightarrow \infty$  gerechtfertigt erscheint.

Sind dagegen die erwarteten Anzahlen zumindest teilweise klein, ist aber der Gesamtumfang  $N_+ = \sum_j N_j$  groß, so bietet sich die „Increasing-Cells“ Asymptotic an, bei der mit  $N_+ \rightarrow \infty$  auch die Anzahl der Gruppen  $J \rightarrow \infty$  (bei festem  $K$ ) anwächst, ohne daß hierbei die Umfänge  $N_j$  anwachsen müssen (sie dürfen sogar alle gleich 1 bleiben!). Unter gewissen Bedingungen an die Umfänge  $N_j$  und Regularitätsbedingungen ergibt sich, daß die  $SD_\lambda$  nicht mehr asymptotisch äquivalent sind, aber jedes  $SD_\lambda$  asymptotisch Normal( $\mu_\lambda, \sigma_\lambda^2$ )-verteilt ist. Hieraus ergibt sich ein auf der Normalverteilung von  $SD_\lambda$  basierender Anpassungstest. Einzelheiten dazu findet man bei Rojek [23] und Osius und Rojek [21].

## 4. Bootstrap-Verfahren für Generalisierte Lineare Modelle

Obwohl die asymptotische Verteilung des Maximum-Likelihood-Schätzers  $\hat{\theta}$  und daraus abgeleiteter Teststatistiken  $T(\hat{\theta})$  in Generalisierten Linearen Modellen bekannt ist, erlaubt dies noch keine Aussagen über die Zulässigkeit der daraus resultierenden Tests oder Konfidenzintervalle einer *konkreten Analyse*, weil oft Unklarheit darüber herrscht, ob die Anwendung der *asymptotischen* Methoden im vorliegenden *finiten* Fall schon gerechtfertigt ist.

Betrachten wir den Schätzer  $\hat{\theta} = \hat{\theta}(\mathbf{Y})$  als eine Funktion der Zielvariablen  $\mathbf{Y}$ , so interessiert uns die Verteilung des Schätzers um den wahren Parameter  $\theta_0$ , also die Verteilung von

$$\hat{\theta}(\mathbf{Y}) - \theta_0. \quad (8)$$

Der klassische Ansatz schätzt die Verteilung von (8) durch eine multivariate Normalverteilung  $N_S(0, \hat{\Sigma})$  mit einer geschätzten Covarianzmatrix  $\hat{\Sigma}$ . Eine alternative Möglichkeit, die Verteilung von (8) zu schätzen, bietet die von Efron [6] eingeführte Idee der Bootstrap-Verfahren. Bootstrap-Methoden wurden bereits von Freedmann [8] auf das klassische lineare Modell angewandt und dann auf verschiedene Weise auf Generalisierte Lineare Modelle verallgemeinert (Rothe [24], Moulton und Zeger [13]).

Die Idee des Bootstrap besteht hierbei darin, die Verteilung von (8) zu schätzen durch die Verteilung eines anderen Zufallvektors

$$\hat{\theta}(\mathbf{Y}^*) - \hat{\theta}_0, \quad (9)$$

wobei  $\hat{\theta}_0 := \hat{\theta}(\mathbf{Y})$  der aus  $\mathbf{Y}$  geschätzte Parameter und  $\mathbf{Y}^*$  eine *Bootstrap-Beobachtung* ist, deren Verteilung eine Schätzung auf die Verteilung der ursprünglichen Beobachtung  $\mathbf{Y}$  ist. Das *parametrische* Bootstrap (Rothe [24]) verwendet eine Bootstrap-Beobachtung  $\mathbf{Y}^*$ , deren Verteilung man aus der von  $\mathbf{Y}$  erhält, wenn man dort den wahren Parameter  $\theta_0$  durch seine Schätzung  $\hat{\theta}_0$  ersetzt. Dagegen wird beim *nicht-parametrischen* Bootstrap (Moulton und Zeger [13]) ein  $\mathbf{Y}^*$  benutzt, dessen Komponenten  $Y_j^*$  voneinander unabhängig jeweils *zufällig* aus der empirischen Verteilung der ursprünglichen Beobachtungen  $Y_1, \dots, Y_J$  gezogen werden (mit dem zugehörigen Covariablenvektor). Da aber die Schätzung  $\hat{\theta}(\mathbf{Y}^*)$  im allgemeinen nicht als explizite Funktion von  $\mathbf{Y}^*$  gegeben ist (sondern iterativ berechnet wird), kann auch die Verteilung von (9) nicht explizit angegeben werden. Man schätzt daher diese Verteilung durch ihre empirische Verteilung, die sich aus der Simulation mit hinreichend vielen unabhängigen Bootstrap-Wiederholungen von  $\mathbf{Y}^*$  ergibt. Hierbei ist einerseits für jede solche Wiederholung ein rechenintensives Iterationsverfahren zur Bestimmung der

Schätzung  $\hat{\theta}(Y^*)$  erforderlich, und andererseits ist die Existenz dieser Schätzung nicht für jede Wiederholung gesichert. Deshalb wird von Mosbach [11] vorgeschlagen, statt  $\hat{\theta}(Y^*)$  einen anderen Schätzer  $\theta^*(Y^*)$  zu verwenden, der in  $Y^*$  linear ist. Dies garantiert sowohl eine schnelle Berechnung als auch die Existenz dieses Schätzers für alle  $Y^*$ . Hierbei wird für  $\hat{\theta}(Y^*)$  derjenige Parameter verwendet, der sich beim Iterationsverfahren für  $\hat{\theta}(Y^*)$  ergibt, wenn man dies ausgehend vom Startwert  $\hat{\theta}_0$  nach dem *ersten* Schritt abbricht und linearisiert, d.h., man führt lediglich eine gewichtete Minimale-Quadrat-Schätzung durch, deren Gewichte sich aus der Modellstruktur für den geschätzten Parameter  $\hat{\theta}_0$  ableiten lassen. Dies liefert das sogenannte *lineare Bootstrap-Verfahren*, bei dem die Verteilung von (8) durch die von

$$\theta^*(Y^*) - \hat{\theta}_0 \quad (10)$$

geschätzt wird. Hierbei kann die Verteilung von  $Y^*$  sowohl parametrisch als auch unparametrisch spezifiziert sein. Die grundlegenden asymptotischen Eigenschaften des linearen Bootstrap findet man in Mosbach [11, 12].

## 5. Asymptotik: Starke Konvergenzbegriffe

Asymptotische Tests oder Konfidenzbereiche werden meist aus Grenzwertsätzen hergeleitet, in denen die Verteilungskonvergenz einer Folge von reellen Zufallsvariablen  $X_n$  gegen eine Zufallsvariable  $X$  für einen Grenzprozeß  $n \rightarrow \infty$  nachgewiesen wird. Manchmal ist es wichtig zu wissen, ob hierbei auch die Momentenfolge  $E\{X_n^r\}$  gegen das Moment  $E\{X^r\}$  konvergiert. Da dies im allgemeinen nicht der Fall ist, haben Bickel und Freedman [2] im Zusammenhang mit asymptotischen Eigenschaften von Bootstrap-Methoden folgende Verschärfung der Verteilungskonvergenz für Zufallsvariablen mit Werten in Banach-Räumen eingeführt. Gilt zusätzlich zur Verteilungskonvergenz  $X_n \xrightarrow{\mathcal{L}} X$  noch die Konvergenz des  $r$ -ten absoluten Moments

$$E\{\|X_n\|^r\} \rightarrow E\{\|X\|^r\}, \quad (11)$$

so heißt  $X_n$   *$d_r$ -konvergent* gegen  $X$ , abgekürzt  $X_n \xrightarrow{d_r} X$ . Der noch stärkere Begriff der  *$e_r$ -Konvergenz* ergibt sich, wenn man (11) verschärft zu

$$E\{\exp \|rX_n\|\} \rightarrow E\{\exp \|rX\|\}. \quad (12)$$

Speziell für reellwertige Zufallsvariable gilt z.B.

$$X_n \xrightarrow{d_2} X \Leftrightarrow X_n \xrightarrow{L} X, E\{X_n\} \rightarrow E\{X\}, \text{Var}\{X_n\} \rightarrow \text{Var}\{X\} \quad (13)$$

$$X_n \xrightarrow{e_r} X \Leftrightarrow \text{Die Moment-erzeugende Funktion von } X_n \text{ konvergiert in } [-r, r] \text{ punktweise gegen die von } X. \quad (14)$$

Für die d- und e-Konvergenz gelten entsprechende Regeln wie für die Verteilungskonvergenz sowie Analoga zum Zentralen Grenzwertsatz. Diese Resultate sind mit einigen statistischen Anwendungen in Osius [17] zusammengestellt.

## 6. Analyse der Säuglingssterblichkeit

Im Zusammenhang mit Umweltbelastungen wie z.B. der Radioaktivität taucht immer wieder die Frage auf, ob diese Belastungen einen statistisch nachweisbaren Einfluß auf die Gesundheit des Menschen und speziell der Neugeborenen haben. Dabei sind in erster Linie die angeborenen Mißbildungen und die Säuglingssterblichkeit von Interesse, wobei wir uns hier auf letztere konzentrieren, weil diese in der amtlichen Bevölkerungsstatistik systematisch erfaßt wird.

Bezeichnet  $T$  die Lebenszeit eines Neugeborenen, so sind neben den Sterberaten der Form  $P\{T \leq t_1\}$  für einen gewissen Zeitabschnitt  $t_1$ , wie z.B.  $t_1 = 0$  (Totgeburt),  $t_1 = 7$  Tage (Frühsterblichkeit) und  $t_1 = 1$  Jahr (Säuglingssterblichkeit), oft auch die bedingten Wahrscheinlichkeiten  $P\{T \leq t_1 | T > t_0\}$  von Interesse, wobei man für  $t_0 = 0$  speziell auf einen Lebendgeborenen bedingt.

Ist nun  $\pi$  irgendeine dieser Sterberaten, z.B.  $\pi = P\{T \leq 7 \text{ Tage} | T > 0\}$ , so besteht das statistische Problem darin, Modelle zu konstruieren, mit deren Hilfe man den Einfluß der in Frage kommenden Belastungsvariablen  $u$  auf  $\pi$  durch einen statistischen Test an Hand konkreter Daten überprüfen kann. Hierfür bieten sich z.B. logistische Quantal-Response-Modelle der Form an

$$\eta := \text{logit}(\pi) = \theta_1 x_2 + \dots + \theta_S x_S + \beta u,$$

wobei  $u$  die interessierende Belastungsvariable ist und  $x_1, \dots, x_S$  weitere Covariable sind, die ebenfalls einen (bereits nachgewiesenen oder vermuteten) statistischen Einfluß auf die Sterblichkeit  $\pi$  haben (z.B. Geschlecht, Geburtsgewicht, Region etc.). Das interessierende Testproblem hat dann die Hypothesen

$$H_0 : \beta = 0 \quad (\text{Nullhypothese : kein Einfluß von } u \text{ auf } \pi)$$

$$H : \beta > 0 \quad (\text{Alternative : positiver Einfluß von } u \text{ auf } \pi)$$

Liegen nun beobachtete Sterberaten  $Y_j$  für  $j = 1, \dots, J$  verschiedene Gruppen vor (die durch Übereinstimmung in allen Covariablen definiert sind), so kann der Test im Rahmen der Quantal-Response-Modelle durchgeführt werden, wobei wir hier  $E(Y_j)$  suggestiv mit  $\pi_j$  (statt wie bisher mit  $\mu_j$ ) bezeichnen. Der Parameter  $\beta$  läßt sich übrigens für das Logit-Modell (im Gegensatz zum entsprechenden Pobit-Modell) leicht unter Verwendung der Chance  $\pi/(1-\pi)$  interpretieren.

Unterscheiden sich zwei Populationsgruppen  $j$  und  $k$  nur in einer Belastung  $u$ , aber nicht in den restlichen Covariablen  $x_1, \dots, x_S$ , so ist der Logarithmus des Chancen-Verhältnisses (engl.: *odds ratio*) im Logit-Modell gegeben durch

$$\log \left( \frac{\pi_j}{1-\pi_j} / \frac{\pi_k}{1-\pi_k} \right) = \text{logit } \pi_j - \text{logit } \pi_k = \beta(u_j - u_k).$$

Für kleine Wahrscheinlichkeiten  $\pi$  ist  $\pi \approx \pi/(1-\pi)$  und das Chancen-Verhältnis entspricht daher für kleine  $\pi_j, \pi_k$  ungefähr dem *relativen Risiko*  $\pi_j/\pi_k$ .

Das Problem beim Testen der Nullhypothese besteht nun darin, daß der Test nur eine geringe Schärfe hat, wenn der Einfluß von  $u$  auf  $\pi$  ebenfalls gering ist, d.h., wenn  $\beta$  klein ist. Insofern kommt der Analyse der Testschärfe sowohl bei der Einschätzung bereits durchgeführter Studien zur Säuglingssterblichkeit, als auch bei der Planung zukünftiger Studien entscheidende Bedeutung zu, wobei für die Planung auch die Verfügbarkeit der „gewünschten“ Covariablen  $x_1, \dots, x_S$  unter Datenschutzaspekten berücksichtigt werden muß. Zu diesem Problemkreis liegen Untersuchungen von Westerhoff [26] vor, die noch weitergeführt werden.

## 7. Analyse der Treffsicherheit der Krebsvorsorgeuntersuchung

Der Ausgangspunkt der folgenden Untersuchungen war eine Anfrage des Zytologischen Labors der Frauenklinik St.-Jürgen- Straße in Bremen (Leitung: Prof. Dr. D. Langnickel) über statistische Aussagen zur Treffsicherheit der *zytologischen* Diagnose bei Abstrichen von Ekto- und Endocervix des Uterus im Rahmen der gesetzlichen Krebsfrüherkennung. Hierfür standen für ein Teilkollektiv zusätzlich eine *pathologische* Diagnose einer Biopsie zur Verfügung. Tabelle 1 zeigt die Klassifikation von insgesamt 1957 Fällen nach der zytologischen und pathologischen Diagnose, die nach Papanicolaou wie folgt klassifiziert sind:

- Pap I** : normales Zellbild
- Pap II** : leicht entzündliche oder regenerative Veränderungen
- Pap III** : schwere entzündliche und/oder degenerative Veränderungen  
(Carcinom nicht auszuschließen)
- Pap III D L** : leichte Dysplasie
- Pap III D M** : mittelgradige Dysplasie
- Pap IV** : vereinzelt verdächtige Zelle
- Pap IV a s** : schwere Dysplasien
- Pap IV a Cis** : Carcinoma in situ
- Pap IV b** : Carcinoma in situ mit beginnender Infiltration
- Pap V** : auf Carcinom verdächtige Zellen
- Pap V P** : Plattenepithel-Carcinom
- Pap V A** : Adeno-Carcinom
- Pap V C** : Corpus-Carcinom

Die statistische Analyse soll einerseits den Zusammenhang beider Diagnosen *quantifizieren* und andererseits *überprüfen*, ob sich dieser Zusammenhang im Laufe der Jahre verändert hat oder nicht. Hierbei gehen wir aufgrund der Datenerhebung davon aus, daß die beobachteten Anzahlen  $Y_{jk}$  für die Kombination  $(j, k)$  der zytologischen und pathologischen Diagnose  $j$  und  $k$  für  $j, k = 1, \dots, J$  voneinander unabhängig und jeweils Poisson-verteilt sind. Es sei  $\mu_{jk} := E\{Y_{jk}\} > 0$  die erwartete Anzahl für die Diagnosekombination  $(j, k)$  und  $\eta_{jk} := \log \mu_{jk}$  deren Logarithmus.

Nun ist die erwartete Tafel  $\mu = (\mu_{jk})$  eindeutig bestimmt durch (Sinkhorn [25])

- die Gesamtsumme  $\mu_{++} := \sum_j \sum_k \mu_{jk}$
- die Zeilensummen  $\mu_{j+} := \sum_k \mu_{jk}$  für  $j > 1$ ,
- die Spaltensummen  $\mu_{+k} := \sum_j \mu_{jk}$  für  $k > 1$ ,
- die Kreuzverhältnisse  $\theta_{j_1 k_1 j_2 k_2} := (\mu_{j_1 k_1} \mu_{j_2 k_2}) / (\mu_{j_1 k_2} \mu_{j_2 k_1})$  für  $j_1 < j_2, k_1 < k_2$ ,

wobei nur die Kreuzverhältnisse Information über den *Zusammenhang* beider Diagnose-Verfahren enthalten. Betrachten wir nun für zwei Diagnose-Stufen  $j < k$  die  $2 \times 2$ -Teiltafel mit nur diesen beiden Zeilen und Spalten, so läßt sich der Zusammenhang beider Diagnose-Verfahren für diese beiden Stufen durch das *Kreuzverhältnis* (engl.: *cross product ratio*, kurz: CPR)

	j	k
j	$\mu_{jj}$	$\mu_{jk}$
k	$\mu_{kj}$	$\mu_{kk}$

$$\theta_{jjkk} = \frac{\mu_{jj}\mu_{kk}}{\mu_{jk}\mu_{kj}} \quad (15)$$

bzw. durch dessen Logarithmus (den log-CPR)

$$\Psi_{jk} := \log \theta_{jjkk} = \eta_{jj} + \eta_{kk} - \eta_{jk} - \eta_{kj} \quad (16)$$

beschreiben. Für  $\theta_{jjkk} = 1$  bzw.  $\Psi_{jk} = 0$  liegt kein Zusammenhang vor, und je größer  $\theta_{jjkk}$  bzw.  $\Psi_{jk}$  ist, desto stärker ist die Übereinstimmung beider Diagnose-Verfahren für diese beiden Stufen  $j, k$ .

Weiter betrachten wir für  $1 < j < k \leq J$  den CPR derjenigen  $2 \times 2$ -Teiltafel mit den Zeilen 1, j und den Spalten 1, k

	1	k
1	$\mu_{11}$	$\mu_{1k}$
j	$\mu_{j1}$	$\mu_{jk}$

$$\theta_{11jk} = \frac{\mu_{11}\mu_{jk}}{\mu_{1k}\mu_{j1}} \quad (17)$$

bzw. dessen Logarithmus

$$\Psi_{jk}^- := \log \theta_{11jk} = \eta_{11} + \eta_{jk} - \eta_{1k} - \eta_{j1}. \quad (18)$$

Man kann nun zeigen, daß sich die Tafel  $\eta = (\eta_{jk})$  eindeutig darstellen läßt in der Form

$$\eta_{jk} = \lambda_0 + \lambda_j^Z + \lambda_k^S - \frac{1}{2}\Psi_{jk} + \frac{1}{2}\Psi_{jk}^-, \quad (19)$$

wobei die Parameter folgenden Nebenbedingungen genügen

$$\begin{aligned} \lambda_1^Z &= 0, & \lambda_1^S &= 0, \\ \Psi_{jk} &= \Psi_{kj}, & \Psi_{jj} &= 0 & \text{für alle } j, k \\ \Psi_{jk}^- &= -\Psi_{kj}^-, & \Psi_{jj}^- &= 0, & \Psi_{1k}^- = 0 & \text{für alle } j, k. \end{aligned} \quad (20)$$

Hierbei ist  $\lambda_0$  ein Gesamteffekt,  $\lambda_j^Z$  ein Zeilen- und  $\lambda_k^S$  ein Spalteneffekt. Aus den Nebenbedingungen ergibt sich, daß die Parameter  $\Psi_{jk}$  den *symmetrischen* und die Parameter  $\Psi_{jk}^-$  den *asymmetrischen* Anteil des Zusammenhangs modellieren. Aufgrund der Interpretation von  $\Psi_{jk}$  als log-CPR kann man spezielle Modelle für  $\Psi$  betrachten, die man teilweise bereits in Agresti [1] findet:

- (I) Keine Übereinstimmung  
 $\Psi_{jk} = 0$  für alle  $j, k$ .
- (II) Übereinstimmung proportional zum „Abstand“ der Klassifikationen:  
 $\Psi_{jk} = \rho \cdot (k - j)$  für alle  $j < k$ .
- (III) Übereinstimmung abhängig vom „Abstand“ der Klassifikationen:  
 $\Psi_{jk} = \rho_{k-j}$  für alle  $j < k$ .
- (IV) Übereinstimmung additiv-abhängig von beiden Klassifikationen:  
 $\Psi_{jk} = \tau_j + \nu_k$  für alle  $j < k$ .
- (V) Übereinstimmung additiv-abhängig von Klassifikationen und „Abstand“:  
 $\Psi_{jk} = \tau_j + \nu_k + \rho_{k-j}$  für alle  $j < k$ .

Die Parameter  $\tau_j, \nu_k$  und  $\rho_i$  sind in diesen Modellen gegebenenfalls erst durch zusätzliche (lineare) Bedingungen eindeutig bestimmt.

Jedes dieser Modelle für  $\Psi$  liefert nach (19) ein log-lineares Modell für  $\mu$ . Die Verwendung dieser Modelle (I-V) für die beobachtete Tafel ergab nur für die Modelle (IV) und (V) eine zufriedenstellende Anpassung, wobei ein Untermodell von (V) mit  $\rho_i = 0$  für  $i \neq 6$  bereits eine gute Anpassung lieferte, die deutlich besser war als das Modell (IV), bei dem auch noch  $\rho_6 = 0$  ist.

In allen Modellen ist die ML-Schätzung  $\hat{\Psi}_{jk}$  asymptotisch  $N(\Psi_{jk}, \sigma_{jk})$ -verteilt, und hieraus lassen sich Tests und Konfidenzbereiche für  $\Psi_{jk}$  konstruieren. Die Schätzungen  $\hat{\Psi}_{jk}$  und ihre geschätzten asymptotischen Standardabweichungen  $\hat{\sigma}_{jk}$  sind für das obige Teilmodell von (V) in Tabelle 2 aufgeführt. Damit ist die Frage nach der Quantifizierung der Übereinstimmung beider Diagnose-Verfahren durch die Schätzung von  $\Psi$  gelöst.

Will man einen zeitlichen Trend analysieren, so betrachtet man die Anzahlen  $Y_{jkt}$  für verschiedene Zeitabschnitte  $t = 1, \dots, T$ , z.B. im vorliegenden Fall mit  $T = 2$  Gruppen für je 6 Jahre, und erhält die zu (19) analoge Darstellung für jeden Zeitabschnitt  $t$

$$\eta_{jkt} = \lambda_{0t} + \lambda_{jt}^Z + \lambda_{kt}^S - \frac{1}{2}\Psi_{jkt} + \frac{1}{2}\Psi_{jkt}^-, \quad (21)$$

wobei wieder die (20) entsprechenden Nebenbedingungen für jedes  $t$  gelten.

Speziell für  $T = 2$  Zeitabschnitte können wir den Unterschied

$$\Delta\Psi_{jk} := \Psi_{jk2} - \Psi_{jk1} \quad (22)$$

des Zusammenhangs für die Diagnose-Paare  $j < k$  betrachten und die Hypothesen testen

$$H_0^{jk} : \Delta\Psi_{jk} = 0 \quad (\text{Zusammenhang unverändert in beiden Zeitgruppen}),$$

$$H^{jk} : \Delta\Psi_{jk} > 0 \quad (\text{Zusammenhang für } T = 2 \text{ positiv-stärker als für } T = 1).$$

Die Schätzung  $\Delta\hat{\Psi}_{jk}$  ist wieder asymptotisch normalverteilt, und hieraus lassen sich Konfidenzintervalle für  $\Delta\Psi_{jk}$  und ein Test für  $H_0^{jk}$  ableiten.

Wir haben für zwei Zeitgruppen (T=2) für das folgende Modell eine gute Modellanpassung an die beobachteten Daten aus Tabelle 1 erhalten.

$$\Psi_{jk1} = \tau_j + r_k + \phi_{k-j}, \quad \text{wobei nur } \phi_6 \neq 0 \text{ ist (siehe oben),}$$

$$\Psi_{jk2} = \Psi_{jk1} + \Delta\Psi_{jk}, \quad \text{mit beliebigem } \Delta\Psi_{jk},$$

$$\Psi_{jk1}^- = \Psi_{jk2}^- \quad \text{d.h. die Asymmetrie ist über die Zeit konstant.}$$

Die Schätzungen  $\Delta\hat{\Psi}_{jk}$  zusammen mit ihren geschätzten Standardabweichungen sind in Tabelle 3 aufgeführt. Die Paare  $j < k$ , bei denen  $H_0^{jk}$  zum 5%-Niveau abgelehnt wird (d.h. eine Verbesserung des Zusammenhangs über die Zeit statistisch abgesichert ist) sind dort gekennzeichnet. Damit ist auch die zeitliche Änderung des Zusammenhangs in beiden Zeitabschnitten analysiert.

## Anhang: Tabellen

Zytologie	Pathologie								Total
	IIIDL	IIIDM	IVas	IVaCis	IVb	VPl	VAd	VCo	
IIIDL	14	19	7	4	0	0	0	0	44
	39	23	7	5	0	0	2	0	76
IIIDM	9	35	24	43	1	2	1	1	116
	26	77	61	36	0	2	1	0	203
	0	2	5	18	3	154	3	5	190
IVas	2	12	22	22	2	4	0	0	64
	2	15	86	54	1	3	1	0	162
IVaCis	5	7	19	161	13	24	2	1	232
	3	7	37	132	3	17	0	0	199
IVb	0	0	3	20	10	43	4	5	85
	0	0	3	17	4	2	0	0	26
VPl	0	1	3	62	11	276	9	2	364
	0	2	5	18	3	154	3	5	190
VAd	0	0	0	3	0	5	42	16	66
	0	0	0	1	0	4	36	12	53
VCo	1	0	1	0	0	0	1	37	40
	0	1	0	0	0	1	0	35	37
Total	31	74	79	315	37	354	59	63	1011
	70	125	199	263	11	183	43	52	946

Tabelle 1: Beobachtete Anzahlen der Klassifikation nach zytologischer und pathologischer Diagnose und zwei Zeitabschnitten: 1972-77 (obere Zahl), 1978-83 (untere Zahl). Weitere Erläuterungen findet man im Text.

Diagnose k	Diagnose j						
	IIIDL	IIIDM	IV a s	IV a Cis	IV b	V Pl	V Ad
IIIDM	1,40						
	0,28						
IV a s	4,31	1,72					
	0,43	0,25					
IV a Cis	5,97	3,37	1,95				
	0,43	0,28	0,20				
IV b	9,23	6,63	5,22	1,94			
	0,65	0,55	0,53	0,39			
V Pl	10,93	8,33	6,91	3,64	2,28		
	0,55	0,44	0,40	0,20	0,41		
V Ad	9,25	13,12	1,70	8,42	7,07	5,72	
	1,52	0,94	0,92	0,84	0,92	0,45	
V Co	18,47	9,40	4,46	11,18	9,82	8,48	5,30
	1,32	1,36	1,25	1,19	1,25	0,96	1,03

**Tabelle 2:** Schätzungen  $\hat{\Psi}_{j,k}$  (oberer Wert) mit geschätzter asymptotischer Standardabweichung (unterer Wert) aller Diagnosepaare  $j < k$  für das Untermodell von (V) (mit  $\rho_i = 0$  außer bei  $i = 6$ ), angepaßt an die Daten des gesamten Zeitraums 1972-83.

Diagnose k	Diagnose j						
	IIIDL	IIIDM	IV a s	IV a Cis	IV b	V Pl	V Ad
IIIDM	0,62 0,61						
IV a s	1,67 * 0,83	0,93 * 0,53					
IV a Cis	1,63 * 0,82	0,89 * 0,51	-0,26 0,44				
IV b	1,40 1,28	0,66 1,11	-0,49 1,08	0,31 0,84			
V Pl	1,92 * 1,06	1,19 0,84	0,04 0,80	0,83 0,42	3,90 * 1,20		
V Ad	-1,18 2,53	3,38 * 1,58	2,23 1,56	3,03 * 1,50	6,10 * 1,85	-1,10 0,89	
V Co	3,41 * 1,81	-2,62 2,62	1,53 1,69	2,32 1,66	5,39 * 1,96	-0,80 1,42	1,29 1,50

**Tabelle 3:** Schätzungen  $\Delta\hat{\Psi}_{jk}$  (oberer Wert) mit geschätzter asymptotischer Standardabweichung (unterer Wert) aller Diagnosepaare  $j < k$  für das im Text beschriebene Modell. Bei den mit \* markierten Paaren ist  $\Delta\Psi_{jk} > 0$  zum Niveau  $\alpha = 5\%$  abgesichert.

## Literatur

- [1] Agresti, A. : *Analysis of Ordered Categorical Data*. New York 1984
- [2] Bickel, P.J., and Freedman, D.A. : *Some asymptotic theory for the bootstrap*. Ann. Statist. **9**, 1196-1217 (1981)
- [3] Billingsley, P. : *Convergence of Probability Measures*. New York 1968
- [4] Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. : *Discrete Multivariate Analysis: Theory and Practice*. Cambridge 1975
- [5] Cressi, N., and Read, T.R.C. : *Multinomial goodness-of-fit tests*. J. Royal Statist. Soc. Ser. B **46**, 3, 440-464 (1984)
- [6] Efron, B. : *Bootstrap methods: Another look at the jackknife*. Ann. Statist. **7**, 1-26 (1979)
- [7] Fahrmeir, L., and Kaufmann, H. : *Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models*. Ann. Statist. **13**, 342-368 (1985)
- [8] Freedman, D.A. : *Bootstrapping regression models*. Ann. Statist. **9**, 1218-1228 (1981)
- [9] Haberman, S.J. : *The Analysis of Frequency Data*. Chicago 1974
- [10] McCullagh, P., and Nelder, J.A. : *Generalized Linear Models (2nd Edition)*. London 1989
- [11] Mosbach, O. : *Bootstrap-Verfahren im Allgemeinen Linearen Modell*. Diplomarbeit, Universität Bremen 1988
- [12] Mosbach, O. : *One-step bootstrapping generalized linear models*. Erscheint im Tagungsbericht der Int. Conference on Bootstrapping and Related Techniques, Trier 1990
- [13] Moulton, L.H., and Zeger, S.L. : *Bootstrapping generalized linear models*. Johns Hopkins University: Department of Biostatistics, Paper No. 572, Baltimore 1985
- [14] Nelder, J.A., and Wedderborn, W.M. : *Generalized linear models*. J. Royal Statist. Soc. **135**, part 3, 370, (1972)
- [15] Osius, G. : *Goodness-of-fit tests for binary data with (possible) small expectations but large degrees of freedom*. Statist. Decisions, Supplement Issue No. 2, 213-224 (1985)
- [16] Osius, G. : *Statistische Lineare Modelle: I, II*. Vorlesungsmanuskript, Fachbereich Mathematik/Informatik, Universität Bremen 1988/89
- [17] Osius, G. : *Some results on convergence of moments and convergence in distributions with applications in statistics*. Mathematik Arbeitspapiere, Preprint Nr. 33, Universität Bremen 1989
- [18] Osius, G. : *Statistische Probleme bei epidemiologischen Studien zur Strahlendosis und Säuglingssterblichkeit*. In: W. Köhnlein, H. Traut und M. Fischer (Ed.): *Die Wirkung niedriger Strahlendosen*. S. 221-223, Berlin 1989

- [19] Osius, G. : *Log-lineare Modelle für den Zusammenhang der Klassifikation nach zwei verschiedenen Diagnoseverfahren*. Vortragsausarbeitung, 36. Biometrisches Kolloquium (Marburg)
- [20] Osius, G., und Drescher, K.: *GLM: Programm zur Datenanalyse mit Generalisierten Linearen Modellen auf Personal Computern*. (Version 2 für ATARI ST und IBM-kompatible unter MS-DOS). Fachbereich Mathematik/Informatik, Universität Bremen 1990
- [21] Osius, G., and Rojek, D. : *Normal goodness-of-fit tests for parametric multinomial models with large degrees of freedom*. Mathematik Arbeitspapiere, Preprint Nr. 36, Universität Bremen 1989
- [22] Read, R.C., and Cressi, N.A.C. : *Goodness-of-fit Statistics for Discrete Multivariate Data*. New York 1988
- [23] Rojek, D. : *Asymptotic für Anpassungstests in Produkt-Multinomialmodellen bei wachsendem Freiheitsgrad*. Dissertation, Universität Bremen 1989
- [24] Rothe, G. : *Bootstrap in Generalisierten Linearen Modellen*. Mannheim: ZUMA-Arbeitsbericht Nr. 86/11, 1986
- [25] Sinkhorn, R. : *Diagonal equivalence to matrices with prescribed row and column sums*. Amer. Math. Month. 74, 402-405 (1967)
- [26] Westerhoff, B. : *Statistische Aspekte bei epidemiologischen Dosis-Wirkungs-Studien*. Dissertation, Universität Bremen 1990

**eingegangen:** Januar 1991

**Verfasser:**

Prof. Dr. G. Osius  
Institut für Statistik Fachbereich 3  
Universität Bremen  
Bibliotheksstraße  
W-2800 Bremen 33  
Germany

MANFRED TASCHKE

## Konstruktion schneller Algorithmen für diskrete Fouriertransformationen

---

Die Entwicklung effektiver Algorithmen für häufig wiederkehrende Grundprobleme ist ein wesentliches Anliegen der Numerischen Mathematik. Von Interesse sind dabei numerisch stabile Algorithmen mit geringer arithmetischer Komplexität (d.h. Anzahl benötigter Multiplikationen und Additionen), einfacher Implementierung, geringem Speicherplatzbedarf und niedriger Anzahl von Datenumordnungen. Ferner sollten die Algorithmen für eine Parallelverarbeitung geeignet sein.

J. W. Cooley und J. W. Tukey [1] haben 1965 den wohl bekanntesten schnellen Algorithmus für die diskrete Fouriertransformation (DFT( $N$ )) der Länge  $N = 2^n$  wiederentdeckt, bei dem die arithmetische Komplexität von  $O(N^2)$  auf  $O(N \log_2 N)$  Operationen gesenkt wird. Erstmals wurde diese Idee bereits 1805 von C. F. Gauß bei astronomischen Berechnungen benutzt. Diese Erkenntnis ist von fundamentaler Bedeutung für die Numerische Mathematik, weil durch das Auffinden schneller Algorithmen eine effektive, numerisch stabile Berechnung vielfach erst möglich wurde. In den letzten 25 Jahren wurden sehr viele schnelle DFT-Algorithmen gefunden (siehe [6], [9]). Als Grundprinzipien zur Konstruktion schneller Algorithmen haben sich dabei

- die Verwendung von Rekursionen,
- die Teile-und-Herrsche-Strategie und
- das Prinzip der Parallelisierung

bewährt. Nicht selten sind die Wurzeln schneller Algorithmen algebraischer Natur, was im folgenden anhand der Indextransformationen und der Polynomarithmetik dargestellt werden soll. Heute besitzt die DFT vielfältige Anwendungen in der Numerischen Mathematik, der Mathematischen Physik, der Kodierungstheorie und der digitalen Signalverarbeitung. Die DFT wird bei der schnellen Faltung, der schnellen Multiplikation von Polynomen, der schnellen Invertierung strukturierter Matrizen, der mehrdimensionalen Boxspline-Interpolation und der Lösung von Operatorgleichungen erfolgreich angewandt.

## 1. Gaußscher DFT-Algorithmus

Historisch entstand die DFT aus der trigonometrischen Interpolation. Die Grundidee eines schnellen DFT-Algorithmus ist bereits in der erst im Nachlaß [3] veröffentlichten Arbeit „Theoria interpolationis methodo nova tractata“ von C. F. Gauß enthalten. C. F. Gauß benutzte 1805 die trigonometrische Interpolation in der Astronomie, um die Bahn des Planetoiden Juno zu berechnen. Seine Methode sei hier in der heute üblichen komplexen Terminologie erläutert.

Bei der *komplexen trigonometrischen Interpolation* besteht die Aufgabe, ein Polynom

$$P(z) := \sum_{j=0}^{N-1} p_j z^j$$

derart zu bestimmen, daß für vorgegebene Werte  $x_j \in \mathbb{C}$

$$P(w_N^{-j}) = x_j \quad (j = 0, \dots, N-1) \quad (1)$$

gilt. Dabei ist  $w_N := \exp(-2\pi i/N)$ . Offenbar gilt

$$1 + w_N^n + w_N^{2n} + \dots + w_N^{(N-1)n} = \begin{cases} N, & \text{falls } n \equiv 0 \pmod{N}, \\ 0, & \text{falls } n \not\equiv 0 \pmod{N}. \end{cases} \quad (2)$$

Die *Fouriermatrix* ist dann definiert durch  $\mathbf{F} := (w_N^{jk})_{j,k=0}^{N-1}$ . Wegen (2) gilt  $\mathbf{F}\overline{\mathbf{F}} = N\mathbf{I}$  und somit  $\mathbf{F}^{-1} = \frac{1}{N}\overline{\mathbf{F}}$ .

Die *diskrete Fouriertransformation* der Länge  $N$  (DFT( $N$ )) ist diejenige Abbildung, die jedem Spaltenvektor  $\mathbf{x} = (x_k)_{k=0}^{N-1} \in \mathbb{C}^N$  einen Spaltenvektor  $\hat{\mathbf{x}} = (\hat{x}_j)_{j=0}^{N-1} \in \mathbb{C}^N$  mit  $\hat{\mathbf{x}} = \mathbf{F}\mathbf{x}$  zuordnet, d.h.

$$\hat{x}_k = \sum_{j=0}^{N-1} x_j w_N^{jk} \quad (k = 0, \dots, N-1). \quad (3)$$

Die *inverse DFT*( $N$ ) lautet dann  $\mathbf{x} = \frac{1}{N}\overline{\mathbf{F}}\hat{\mathbf{x}}$ , d.h.

$$x_j = \frac{1}{N} \sum_{k=0}^{N-1} \hat{x}_k w_N^{-jk} \quad (j = 0, \dots, N-1).$$

Zur Berechnung der DFT( $N$ ) würde man  $N^2$  komplexe Multiplikationen und  $N(N-1)$  komplexe Additionen benötigen, vorausgesetzt, daß die  $w_N$ -Potenzen vorberechnet worden sind. Das Ziel besteht nun darin, die DFT( $N$ ) mit wesentlich weniger Operationen zu berechnen. Die Interpolationsbedingung (1) lautet folglich  $\overline{\mathbf{F}}\mathbf{p} = \mathbf{x}$  mit  $\mathbf{p} = (p_j)_{j=0}^{N-1}$ , d.h.

$$\mathbf{p} = \frac{1}{N} \mathbf{F} \mathbf{x} = \frac{1}{N} \hat{\mathbf{x}}.$$

Das komplexe trigonometrische Interpolationsproblem besitzt somit die eindeutig bestimmte Lösung

$$p_k = \frac{1}{N} \sum_{j=0}^{N-1} x_j w_N^{jk} = \frac{1}{N} \hat{x}_k \quad (k = 0, \dots, N-1). \quad (4)$$

Verallgemeinern wir (1) und suchen wir für festes  $a \in \mathbb{C}$ , ( $a \neq 0$ ) ein Polynom  $P(z)$  mit

$$P(a^{-1} w_N^{-j}) = x_j \quad (j = 0, \dots, N-1), \quad (5)$$

dann gilt völlig analog

$$p_k = \frac{1}{N} a^k \sum_{j=0}^{N-1} x_j w_N^{jk} \quad (k = 0, \dots, N-1). \quad (6)$$

C. F. Gauß interessierte sich für eine effektive Berechnungsvorschrift der Koeffizienten (4) des trigonometrischen Interpolationspolynoms  $P(z)$  in den Fällen  $N = 12, 24, 36$ . Dabei wendete er die Teile-und-Herrsche-Strategie an. Im folgenden leiten wir den Gaußschen DFT-Algorithmus auf zwei verschiedene Weisen her. Der erste Zugang entspricht der von C. F. Gauß benutzten Methode in komplexer Schreibweise. Der zweite Weg ist kürzer und betont die Bedeutung der Indextransformationen bei schnellen DFT-Algorithmen.

**1. Herleitung.** Es sei  $N = N_1 N_2$ . Zuerst werden  $N_1$  Interpolationsprobleme für jeweils  $N_2$  gegebene Daten betrachtet. Für jedes  $j_1 = 0, \dots, N_1 - 1$  wird ein Polynom

$$Q_{j_1}(z) = \sum_{k_2=0}^{N_2-1} q_{j_1+k_2 N_1} z^{k_2}$$

mit der Interpolationseigenschaft

$$Q_{j_1}(w_N^{-j_1} w_{N_2}^{-k_2}) = x_{j_1+k_2 N_1} \quad (k_2 = 0, \dots, N_2 - 1)$$

bestimmt. Die Lösungen lauten nach (6)

$$q_{j_1+k_2 N_1} = \frac{1}{N_2} w_N^{j_1 k_2} \sum_{j_2=0}^{N_2-1} x_{j_1+j_2 N_1} w_{N_2}^{j_2 k_2} \quad (k_2 = 0, \dots, N_2 - 1). \quad (7)$$

Im zweiten Schritt werden  $N_2$  Teilprobleme für  $N_1$  gegebene Daten gelöst. Für jedes  $k_2 = 0, \dots, N_2 - 1$  wird ein Polynom

$$R_{k_2}(z) = \sum_{k_1=0}^{N_1-1} r_{k_2+k_1N_2} z^{k_1}$$

mit

$$R_{k_2}(w_{N_1}^{-j_1}) = q_{j_1+k_2N_1} \quad (j_1 = 0, \dots, N_1 - 1)$$

gesucht. Wegen (4) und (7) ergibt sich dann

$$r_{k_2+k_1N_2} = \frac{1}{N_1} \sum_{j_1=0}^{N_1-1} q_{j_1+k_2N_1} w_{N_1}^{j_1 k_1} = \frac{1}{N} \sum_{j_1=0}^{N_1-1} \left( \sum_{j_2=0}^{N_2-1} (x_{j_1+j_2N_1} w_N^{j_1 k_2}) w_{N_2}^{j_2 k_2} \right) w_{N_1}^{j_1 k_1}. \quad (8)$$

Mittels  $j = j_1 + j_2 N_1$  ( $j_t = 0, \dots, N_t - 1$ ;  $t = 1, 2$ ) folgt

$$r_{k_2+k_1N_2} = \frac{1}{N} \sum_{j=0}^{N-1} x_j w_N^{j(N_2 k_1 + k_2)} = p_{N_2 k_1 + k_2}. \quad (9)$$

**2. Herleitung.** Durch die beiden *Indextransformationen*

$$j = j_1 + j_2 N_1, \quad k = N_2 k_1 + k_2 \quad (j, k = 0, \dots, N - 1)$$

mit  $j_t, k_t = 0, \dots, N_t - 1$  ( $t = 1, 2$ ) folgt aus (3)

$$\begin{aligned} \hat{x}_{N_2 k_1 + k_2} &= \sum_{j_1=0}^{N_1-1} \sum_{j_2=0}^{N_2-1} x_{j_1+j_2N_1} w_N^{(j_1+j_2N_1)(N_2 k_1 + k_2)} \\ &= \sum_{j_1=0}^{N_1-1} \left( \sum_{j_2=0}^{N_2-1} (x_{j_1+j_2N_1} w_N^{j_1 k_2}) w_{N_2}^{j_2 k_2} \right) w_{N_1}^{j_1 k_1}. \end{aligned} \quad (10)$$

Geht man dagegen von den Indextransformationen  $j = N_2 j_1 + j_2$ ,  $k = k_1 + k_2 N_1$  aus, so erhält man anstelle von (10) die Vorschrift

$$\hat{x}_{k_1+k_2N_1} = \sum_{j_1=0}^{N_1-1} \left( \sum_{j_2=0}^{N_2-1} (x_{N_2 j_1 + j_2} w_N^{j_1 k_2}) w_{N_2}^{j_2 k_2} \right) w_{N_1}^{j_1 k_1}. \quad (11)$$

Offensichtlich sind die Formeln (9) und (10) gleichbedeutend wegen (4). Somit kann im Fall  $N = N_1 N_2$  eine DFT( $N$ ) durch  $N_1$  DFT( $N_2$ ) und  $N_2$  DFT( $N_1$ ) berechnet werden.

Dazu müssen die Daten  $x_{j_1+j_2N_1}$  noch mit den sogenannten *Drehfaktoren*  $w_N^{j_1k_2}$  multipliziert werden. Wir erhalten aus (10) folgenden

### DFT-Algorithmus von C. F. Gauß

Gegeben:  $x_j$  ( $j = 0, \dots, N-1$ ),  $N = N_1N_2$

Gesucht:  $\hat{x}_k$  ( $k = 0, \dots, N-1$ )

1. Berechne für  $j_1 = 0, \dots, N_1-1$

$$y_{j_1+k_2N_1} := \sum_{j_2=0}^{N_2-1} (x_{j_1+j_2N_1} w_N^{j_1k_2}) w_{N_2}^{j_2k_2}.$$

2. Berechne für  $k_2 = 0, \dots, N_2-1$

$$\hat{x}_{N_2k_1+k_2} = \sum_{j_1=0}^{N_1-1} y_{j_1+k_2N_1} w_{N_1}^{j_1k_1}.$$

Es muß betont werden, daß C. F. Gauß in [3] diesen Algorithmus mit *reellen* trigonometrischen Funktionen formulierte, was das Verständnis erschwert. Aus (11) ergibt sich ein ähnlicher Algorithmus. Wir wollen beide Algorithmen kurz illustrieren.

**Beispiel 1.** Es sei  $N > 2$  eine gerade Zahl. Mit  $N_1 = N/2$  und  $N_2 = 2$  ergibt sich aus (10) für  $k = 0, \dots, N/2-1$

$$\begin{aligned} \hat{x}_{2k} &= \sum_{j=0}^{N/2-1} (x_j + x_{j+N/2}) w_{N/2}^{jk}, \\ \hat{x}_{2k+1} &= \sum_{j=0}^{N/2-1} ((x_j - x_{j+N/2}) w_N^j) w_{N/2}^{jk}. \end{aligned} \quad (12)$$

Somit läßt sich die DFT( $N$ ) durch  $N$  Additionen,  $N/2$  Multiplikationen und zwei DFT( $N/2$ ) berechnen.

**Beispiel 2.** Wieder sei  $N > 2$  gerade. Für  $N_1 = N/2$  und  $N_2 = 2$  erhält man für  $k = 0, \dots, N/2-1$  aus (11)

$$\begin{aligned} \hat{x}_k &= \sum_{j=0}^{N/2-1} x_{2j} w_{N/2}^{jk} + w_N^k \sum_{j=0}^{N/2-1} x_{2j+1} w_{N/2}^{jk}, \\ \hat{x}_{k+N/2} &= \sum_{j=0}^{N/2-1} x_{2j} w_{N/2}^{jk} - w_N^k \sum_{j=0}^{N/2-1} x_{2j+1} w_{N/2}^{jk}. \end{aligned} \quad (13)$$

**Beispiel 3.** Es sei  $N > 4$  durch 4 teilbar. Für  $N_1 = N/4$  und  $N_2 = 4$  ergibt sich aus (10) für  $k = 0, \dots, N/4 - 1$

$$\begin{aligned}
 \hat{x}_{4k} &= \sum_{j=0}^{N/4-1} ((x_j + x_{j+N/2}) + (x_{j+N/4} + x_{j+3N/4})) w_{N/4}^{jk}, \\
 \hat{x}_{4k+1} &= \sum_{j=0}^{N/4-1} ((x_j - x_{j+N/2}) - i(x_{j+N/4} - x_{j+3N/4})) w_N^j w_{N/4}^{jk}, \\
 \hat{x}_{4k+2} &= \sum_{j=0}^{N/4-1} ((x_j + x_{j+N/2}) - (x_{j+N/4} + x_{j+3N/4})) w_N^{2j} w_{N/4}^{jk}, \\
 \hat{x}_{4k+3} &= \sum_{j=0}^{N/4-1} ((x_j - x_{j+N/2}) + i(x_{j+N/4} - x_{j+3N/4})) w_N^{3j} w_{N/4}^{jk}.
 \end{aligned} \tag{14}$$

**Beispiel 4.** Es sei  $N > 4$  durch 4 teilbar. Durch Kombination von (12) und (14) erhält man für  $k = 0, \dots, N/2 - 1$  und  $l = 0, \dots, N/4 - 1$

$$\begin{aligned}
 \hat{x}_{2k} &= \sum_{j=0}^{N/2-1} (x_j + x_{j+N/2}) w_{N/2}^{jk}, \\
 \hat{x}_{4l+1} &= \sum_{j=0}^{N/4-1} ((x_j - x_{j+N/2}) - i(x_{j+N/4} - x_{j+3N/4})) w_N^j w_{N/4}^{jl}, \\
 \hat{x}_{4l+3} &= \sum_{j=0}^{N/4-1} ((x_j - x_{j+N/2}) + i(x_{j+N/4} - x_{j+3N/4})) w_N^{3j} w_{N/4}^{jl}.
 \end{aligned}$$

Hier läßt sich die DFT( $N$ ) durch  $3N/2$  Additionen,  $N/2$  Multiplikationen, eine DFT( $N/2$ ) und zwei DFT( $N/4$ ) berechnen.

Abschließend sei noch bemerkt, daß sich der Gaußsche DFT-Algorithmus auf den Fall  $N = N_1 N_2 \dots N_r$  ( $r \geq 2$ ) erweitern läßt. Entsprechende Indextransformationen sind beispielsweise

$$j = j_1 + j_2 N_2 + \dots + j_r N_1 \dots N_{r-1}, \quad k = k_r + k_{r-1} N_r + \dots + k_1 N_2 \dots N_r$$

mit  $j_t, k_t = 0, \dots, N_t - 1$  ( $t = 1, \dots, r$ ).

## 2. Algorithmus von C. Runge

Nach C. F. Gauß sind effektive Methoden zur diskreten Fourieranalyse erst von C. Runge [7] entwickelt worden. C. Runge betrachtete 1903 das folgende Problem der *reellen trigonometrischen Interpolation*: Gesucht ist ein trigonometrisches Polynom

$$T(t) := \frac{1}{N} \hat{a}_0 + \frac{2}{N} \sum_{j=0}^N (\hat{a}_j \cos jt + \hat{b}_j \sin jt), \quad (t \in \mathbf{R})$$

mit der Interpolationseigenschaft

$$T\left(\frac{2j\pi}{2N}\right) = x_j \quad (j = 0, \dots, 2N-1),$$

wobei  $x_j \in \mathbf{R}$  gegebene Daten sind. Diese Aufgabe ist eindeutig lösbar und es gilt

$$\hat{a}_j = \sum_{k=0}^N a_k \cos \frac{jk\pi}{N} \quad (j = 0, \dots, N), \quad (15)$$

$$\hat{b}_j = \sum_{k=1}^{N-1} b_k \sin \frac{jk\pi}{N} \quad (j = 1, \dots, N-1), \quad \hat{b}_N = 0. \quad (16)$$

Dabei ist  $a_0 := x_0$ ,  $a_k := x_k + x_{2N-k}$  ( $k = 1, \dots, N-1$ ),  $a_N := x_N$  und  $b_k := x_k - x_{2N-k}$  ( $k = 1, \dots, N-1$ ) gesetzt worden. Diese Transformationen (15) und (16) heißen *diskrete Fourier-Kosinus-Transformation* bzw. *diskrete Fourier-Sinus-Transformation*. Es besteht ein enger Zusammenhang zwischen diesen Transformationen und einer DFT(2N), worauf an dieser Stelle nicht näher eingegangen werden soll.

C. Runge gab für  $N = 12, 24, 36$  effektive Berechnungsschemata für (15) und (16) an. Auch er verwendete die Teile-und-Herrsche-Strategie. Es sei  $N > 2$  gerade. Die Ausgabewerte von (15) mit geraden bzw. ungeraden Indizes werden getrennt betrachtet. Dazu bildet man

$$c_k := a_k + a_{N-k} \quad (k = 0, \dots, N/2 - 1), \quad c_{N/2} := a_{N/2}$$

$$d_k := a_k - a_{N-k} \quad (k = 0, \dots, N/2 - 1).$$

Wegen

$$\cos \frac{(N-k)2j\pi}{N} = \cos \frac{2jk\pi}{N}, \quad \cos \frac{(N-k)(2j+1)\pi}{N} = -\cos \frac{k(2j+1)\pi}{N}$$

folgt dann aus (15)

$$\hat{a}_{2j} = \sum_{k=0}^{N/2} c_k \cos \frac{2jk\pi}{N} \quad (j = 0, \dots, N/2),$$

$$\hat{a}_{2j+1} = \sum_{k=0}^{N/2-1} d_k \cos \frac{(2j+1)k\pi}{N} \quad (j = 0, \dots, N/2-1).$$

Analog setzt man

$$e_k := b_k - b_{N-k} \quad (k = 1, \dots, N/2-1),$$

$$f_k := b_k + b_{N-k} \quad (k = 1, \dots, N/2-1), \quad f_{N/2} = b_{N/2}.$$

Dann gilt für (16)

$$\hat{b}_{2j} = \sum_{k=1}^{N/2-1} e_k \sin \frac{2jk\pi}{N} \quad (j = 1, \dots, N/2-1),$$

$$\hat{b}_{2j+1} = \sum_{k=1}^{N/2} f_k \sin \frac{(2j+1)k\pi}{N} \quad (j = 0, \dots, N/2-1).$$

Dieses Reduktionsverfahren ist effektiv für relativ kleines  $N$ . Auch diese Methode läßt sich bereits bei C. F. Gauß nachweisen. In der 1811 entstandenen Arbeit „Allgemeine Störungen der Pallas durch Jupiter. Erste Rechnung“ [4] hat C. F. Gauß diskrete Fourier-Analysen bis  $N = 24$  durchgeführt.

### 3. DFT-Algorithmus von J. W. Cooley und J. W. Tuckey

Das Verdienst von J. W. Cooley und J. W. Tuckey besteht darin, daß sie 1965 einen schnellen, einfach programmierbaren Algorithmus für die DFT( $N$ ) mit großer Radix-2-Länge  $N = 2^n$  angegeben haben. Dieser Algorithmus beruht auf einer rekursiven Anwendung des Gaußschen DFT-Algorithmus, der J. W. Cooley und J. W. Tuckey aber nicht bekannt war. Folglich ist der Gaußsche DFT-Algorithmus allgemeiner als der Cooley-Tuckey-Algorithmus, der sich aus Beispiel 1 unmittelbar ablesen läßt.

Bezeichnet  $M_N$  die Anzahl der benötigten komplexen Multiplikationen und  $A_N$  die Anzahl der komplexen Additionen bei einem schnellen DFT( $N$ )-Algorithmus, so gilt wegen (12) für die sog. *Methode der Frequenzminderung* (decimation-in-frequency)

$$M_N = 2M_{N/2} + N/2, \quad M_2 = 1,$$

$$A_N = 2A_{N/2} + N, \quad A_2 = 2.$$

Dabei wird angenommen, daß die  $w_N$ -Potenzen vorberechnet worden sind. Wegen  $N = 2^n$  setzt man  $\mu_n := M_N$ , so daß sich die lineare Differenzgleichung

$$\mu_n = 2\mu_{n-1} + 2^{n-1} \quad (n = 2, 3, \dots), \quad \mu_1 = 1$$

ergibt, deren Lösung  $\mu_n = n2^{n-1} = \frac{1}{2}N \log_2 N$  lautet. Analog zeigt man  $A_N = N \log_2 N$ . Im Fall  $N = 2^n$  besitzt der Cooley-Tuckey-Algorithmus die arithmetische Komplexität  $O(N \log_2 N)$ .

Aus Beispiel 2 ergibt sich für  $N = 2^n$  nach der sog. *Methode der Zeitverminderung* (decimation-in-time) ein weiterer schneller Radix-2-Algorithmus.

Aus Beispiel 3 folgt für  $N = 4^n$  nach der Methode der Frequenzverminderung ein Radix-4-Algorithmus mit einer geringeren Komplexität als die entsprechenden, vorher beschriebenen Radix-2-Algorithmen. Der Radix-4-Algorithmus reduziert die Anzahl der Multiplikationen um etwa 25% gegenüber dem ersten Radix-2-Algorithmus.

Wird Beispiel 4 für  $N = 2^n$  rekursiv angewandt, so gelangt man zu einem Split-radix-Algorithmus [2], der weniger Operationen als die zuvor erwähnten Algorithmen benötigt. Ein „optimaler“ DFT( $N$ )-Algorithmus ist selbst im Fall  $N = 2^n$  unbekannt.

Zum Vergleich geben wir die Anzahl der nichttrivialen reellen Multiplikationen ( $M_N^r$ ) und reellen Additionen ( $A_N^r$ ) für verschiedene schnelle DFT( $N$ )-Algorithmen an. Dabei wird angenommen, daß eine nichttriviale komplexe Multiplikation durch 3 reelle Multiplikationen und 3 reelle Additionen realisiert wird.

Länge	Radix-2-Alg.		Radix-4-Alg.		Split-radix-Alg.	
	$M_N^r$	$A_N^r$	$M_N^r$	$A_N^r$	$M_N^r$	$A_N^r$
16	24	152	20	148	20	148
32	88	408	-	-	68	388
64	264	1032	208	976	196	964
128	712	2504	-	-	516	2308
256	1800	5896	1392	5488	1284	5380
512	4360	13566	-	-	3076	12292
1024	10248	30728	7856	28336	7172	27652

Die vielfältigen Anwendungen der DFT und die rasante Entwicklung der Rechentechnik lösten seit 1965 eine Fülle von Untersuchungen zu schnellen Algorithmen aus. Die bekanntesten Verfahren sind – neben den oben erwähnten Algorithmen – der Primfaktoralgorithmus, der Rader-Algorithmus und der Winograd-Algorithmus.

Ein einheitlicher Zugang zu vielen Algorithmen ist auch über die Polynomarithmetik möglich, worauf bereits S. Winograd [10] und H. J. Nussbaumer [6] hinwiesen. Als Beispiel betrachten wir erneut den Fall  $N = 2^n$ . Dem gegebenen Datenvektor  $(x_j)_{j=0}^{N-1} \in \mathbb{C}^N$  wird das Polynom

$$P(z) := \sum_{j=0}^{N-1} x_j z^j$$

zugeordnet. Für  $z = w_N^k$  ( $k = 0, \dots, N-1$ ) erhalten wir die gesuchten Werte  $\hat{x}_k = P(w_N^k)$  ( $k = 0, \dots, N-1$ ). Mittels Horner-Schema lassen sich derartige Polynomauswertungen bekanntlich mit  $O(N^2)$  Operationen realisieren. Wir wollen aber mit wesentlich weniger Operationen auskommen.

Der folgende Radix-2-Algorithmus basiert auf der rekursiven Anwendung der Teile-und-Herrsche-Strategie. Er entspricht Beispiel 1. Es sei

$$k = (k_{n-1}, \dots, k_0)_2 := 2^{n-1}k_{n-1} + \dots + 2k_1 + k_0 \quad (k_j \in \{0, 1\})$$

die binäre Darstellung von  $k \in \{0, \dots, N-1\}$ . Wir spalten  $P(z)$  wie folgt auf:

$$P(z) = \sum_{j=0}^{N/2-1} x_j z^j + z^{N/2} \sum_{j=0}^{N/2-1} x_{N/2+j} z^j.$$

Da  $z^{N/2}$  an den  $N$  Knoten  $w_N^k$  ( $k = 0, \dots, N-1$ ) nur die beiden Werte  $\exp(-\pi i k) = (-1)^{k_0}$  annimmt, lautet nun der

**Schritt 1.** Berechne für  $i_0 = 0, 1$  und  $j = 0, \dots, N/2 - 1$

$$x_{j+i_0N/2}^{(1)} := x_j + (-1)^{i_0} x_{N/2+j}.$$

Bildet man die beiden Polynome

$$P_{i_0}^{(1)}(z) := \sum_{j=0}^{N/2-1} x_{j+i_0N/2}^{(1)} z^j \quad (i_0 = 0, 1),$$

so gilt

$$\hat{x}_k = P_{i_0}^{(1)}(w_N^k) = P(w_N^k)$$

für alle  $k = 0, \dots, N - 1$  mit  $k_0 = i_0$ . Analog wie mit  $P(z)$  verfahren wir jetzt mit  $P_{i_0}^{(1)}(z)$ :

$$P_{i_0}^{(1)}(z) = \sum_{j=0}^{N/4-1} x_{j+i_0N/2}^{(1)} z^j + z^{N/4} \sum_{j=0}^{N/4-1} x_{j+N/4+i_0N/2}^{(1)} z^j.$$

Da  $z^{N/4}$  an den Knoten  $w_N^k$  ( $k = 0, \dots, N - 1$ ) nur die 4 Werte  $\exp(-\pi i k/2) = (-1)^{k_1} (-i)^{k_0}$  ( $k = 0, \dots, N - 1$ ) annimmt, lautet nun

**Schritt 2.** Berechne für  $i_0, i_1 = 0, 1$  und  $j = 0, \dots, N/4 - 1$

$$x_{j+(i_1, i_0)_2 N/4}^{(2)} := x_{j+i_0N/2}^{(1)} + (-1)^{i_1} (-i)^{i_0} x_{N/4+j+i_0N/2}^{(1)}.$$

Die Fortsetzung dieses Verfahrens liefert schließlich im

**Schritt n.** Berechne für  $i_0, \dots, i_{n-1} = 0, 1$

$$x_{(i_{n-1}, \dots, i_0)_2}^{(n)} := x_{2(i_{n-2}, \dots, i_0)_2}^{(n-2)} + (-1)^{i_{n-1}} w_N^{(i_{n-2}, \dots, i_0)_2} x_{1+2(i_{n-2}, \dots, i_0)_2}^{(n-1)}.$$

Dann lautet die

**Ausgabe.**  $\hat{x}_k = x_k^{(n)}$  ( $k = 0, \dots, N - 1$ ).

Die Polynomdarstellung bietet auch den Vorteil, daß die Umordnungen der Zwischenwerte genau beschrieben werden.

## 4. Primfaktoralalgorithmus von I. J. Good

Im folgenden sei  $N = N_1 \dots N_r$  ein Produkt von paarweise teilerfremden Zahlen  $N_1, \dots, N_r$  ( $r > 1$ ). Bei der Berechnung einer DFT( $N$ ) wird oft das „Prinzip der Dimensionserhöhung“ angewandt. Dabei wird die DFT( $N$ ) in eine  $r$ -dimensionale DFT( $N_1 \times \dots \times N_r$ ) vom Format  $N_1 \dots N_r$

$$\hat{x}_{k_1 \dots k_r} = \sum_{j_1=0}^{N_1-1} \dots \sum_{j_r=0}^{N_r-1} x_{j_1 \dots j_r} w_{N_1}^{j_1 k_1} \dots w_{N_r}^{j_r k_r}$$

mit  $k_t = 0, \dots, N_t - 1$  ( $t = 1, \dots, r$ ) überführt. Eine effektive Berechnung der DFT( $N_1 \times \dots \times N_r$ ) kann schrittweise entsprechend der Klammersetzung

$$\sum_{j_1=0}^{N_1-1} \left( \dots \left( \sum_{j_r=0}^{N_r-1} x_{j_1 \dots j_r} w_{N_r}^{j_r k_r} \right) \dots \right) w_{N_1}^{j_1 k_1}$$

mittels schneller Algorithmen für DFT( $N_r$ ), ..., DFT( $N_1$ ) erfolgen. Das Berechnungsschema für die DFT( $N$ ), welches erstmals 1958 von I. J. Good [5] vorgeschlagen wurde, sieht wie folgt aus:

1. Durch eine Indextransformation der Eingabedaten  $x_j$  ( $j = 0, \dots, N - 1$ ) wird die DFT( $N$ ) in eine  $r$ -dimensionale DFT( $N_1 \times \dots \times N_r$ ) mit den Eingabedaten  $x_{j_1 \dots j_r}$  ( $j_t = 0, \dots, N_t - 1$ ;  $t = 1, \dots, r$ ) überführt.
2. Mittels schneller Algorithmen für die DFT( $N_t$ ) ( $t = 1, \dots, r$ ) wird die  $r$ -dimensionale DFT( $N_1 \times \dots \times N_r$ ) parallel berechnet.
3. Durch die Indextransformation der Ausgabedaten  $\hat{x}_{k_1 \dots k_r}$  ( $k_t = 0, \dots, N_t - 1$ ;  $t = 1, \dots, r$ ) der  $r$ -dimensionalen DFT( $N_1 \times \dots \times N_r$ ) erhält man die gesuchten Ausgabedaten  $\hat{x}_k$  ( $k = 0, \dots, N - 1$ ) der DFT( $N$ ).

Der gebräuchlichste Primfaktoralalgorithmus hat folgendes Aussehen:

Es seien  $M_t = N/N_t$  ( $t = 1, \dots, r$ ). Ferner seien  $M'_t$  ( $1 \leq M'_t < N_t$ ) durch  $M_t M'_t \equiv 1 \pmod{N_t}$  ( $t = 1, \dots, r$ ) definiert. Für die Eingabedaten wählt man die Indextransformation

$$j \equiv \sum_{t=1}^r M_t j_t \pmod{N} \quad (j = 0, \dots, N - 1; j_t = 0, \dots, N_t - 1; t = 1, \dots, r) \quad (17)$$

und für die Ausgabedaten

$$k \equiv \sum_{t=1}^r M_t M'_t k_t \pmod{N} \quad (k = 0, \dots, N - 1; k_t = 0, \dots, N_t - 1; t = 1, \dots, r). \quad (18)$$

Es sei bemerkt, daß  $k \equiv k_t \pmod{N_t}$  ( $t = 1, \dots, r$ ) gilt und daß (18) der *Chinesische Restsatz* ist, der bereits um das Jahr 100 von Sun Tsu erwähnt und erstmals von L. Euler 1734 bewiesen wurde.

Aus (17) und (18) ergibt sich für  $j, k = 0, \dots, N-1$

$$jk \equiv \sum_{t=1}^r M_t j_t k_t \pmod{N}$$

und somit

$$w_N^{jk} = w_{N_1}^{j_1 k_1} \dots w_{N_r}^{j_r k_r}.$$

Nun setzt man

$$x_{j_1 \dots j_r} = x_{M_1 j_1 + \dots + M_r j_r} \quad (j_t = 0, \dots, N_t - 1; t = 1, \dots, r)$$

und

$$\hat{x}_{k_1 \dots k_r} = \hat{x}_{M_1' k_1 + \dots + M_r' k_r} \quad (k_t = 0, \dots, N_t - 1; t = 1, \dots, r).$$

Bei dem Primfaktoralalgorithmus treten wegen der vorausgesetzten Teilerfremdheit von  $N_1, \dots, N_r$  keine Drehfaktoren wie bei den Radix-2-Algorithmen auf.

Abschließend sei bemerkt, daß man einen beliebigen Isomorphismus von der additiven Gruppe  $\mathbb{Z}/N\mathbb{Z}$  auf die direkte Summe  $\mathbb{Z}/N_1\mathbb{Z} \oplus \dots \oplus \mathbb{Z}/N_r\mathbb{Z}$  als Indextransformation wählen kann. Die andere Indextransformation ist dann eindeutig bestimmt und ebenfalls ein Isomorphismus. Somit ist die Konstruktion von Indextransformationen ein schönes Anwendungsbeispiel für den Isomorphiesatz endlicher abelscher Gruppen. Hinsichtlich der Eigenschaften und der Konstruktion von Indextransformationen sei auf [8] verwiesen.

## Literatur

- [1] Cooley, J.W., and Tukey, J.W.: *An algorithm for the machine calculation of complex Fourier series*. Math. Comp. **19**, 297–301 (1965)
- [2] Duhamel, P.: *Implementation of 'split-radix' FFT algorithm for complex, real, and real-symmetric data*. IEEE Trans. Acoust. Speech Signal Process. **34**, 285–295 (1986)
- [3] Gauß, C.F.: *Theoria interpolationis methodo nova tractata*. In: Carl Friedrich Gauß Werke, Bd. III, S. 265–327, Göttingen 1876

- [4] Gauß, C.F. : *Allgemeine Störungen der Pallas durch Jupiter. Erste Rechnung.* In: Carl Friedrich Gauß Werke, Bd. VII, S. 489–528, Göttingen 1906
- [5] Good, I.J. : *The interaction algorithm and practical Fourier analysis.* J. Roy. Statist. Soc. Ser. B **20**, 361–372 (1958); **22**, 372–375 (1960)
- [6] Nussbaumer, H.J. : *Fast Fourier Transform and Convolution Algorithms.* Berlin 1982
- [7] Runge, C. : *Über die Zerlegung empirisch gegebener periodischer Funktionen in Sinuswellen.* Z. Math. Phys. **48**, 443–456 (1903); **52**, 117–123 (1905)
- [8] Steidl, G., and Tasche, M.: *Index transforms for multidimensional DFT's and convolutions.* Numer. Math. **56**, 513–528 (1989)
- [9] Tolimieri, R., An, M., and Lu, C.: *Algorithms for Discrete Fourier Transform and Convolution.* New York 1989
- [10] Winograd, S. : *Arithmetic Complexity of Computations.* Philadelphia 1980

**eingegangen:** März 1991

**Verfasser:**

Prof. Dr. M. Tasche  
Fachbereich Mathematik  
Universität Rostock  
Universitätsplatz 1  
O-2500 Rostock  
Germany

## Hinweise für Verfasser

---

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

Insbesondere beachte man:

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

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

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

[8] **Gnedenko, B.W.:** *Über die Arbeiten von C.F. Gauß zur Wahrscheinlichkeitsrechnung*. In: Reichard, H. (Ed.): C.F. Gauß, Gedenkband anlässlich des 100. Todestages. S. 193–204, Leipzig 1967

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

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

Weiterhin besteht die Möglichkeit, mit dem **Satzsystem L<sup>A</sup>T<sub>E</sub>X** erstellte Manuskripte auf unter **MS-DOS** formatierten Disketten (**3.5"**, **0.72MB**, **1.44MB**; **5.25"**, **0.36MB**, **0.72MB**, **1.2MB**) einzureichen.

## SCHRIFTENREIHEN DER UNIVERSITÄT ROSTOCK

- |                                                                    |                |
|--------------------------------------------------------------------|----------------|
| - Archiv der Freunde der Naturgeschichte in Mecklenburg            | ISSN 0518-3189 |
| - Rostocker Agrarwissenschaftliche Beiträge                        | ISSN 0138-3299 |
| - Rostocker Betriebswirtschaftliche Manuskripte                    | ISSN 0232-3066 |
| - Rostocker Mathematisches Kolloquium                              | ISSN 0138-3248 |
| - Rostocker Physikalische Manuskripte                              | ISSN 0138-3140 |
| - Rostocker Wissenschaftshistorische Manuskripte                   | ISSN 0138-3191 |
| - Lateinamerika                                                    | ISSN 0458-7944 |
| - Erziehungswissenschaftliche Beiträge                             | ISSN 0138-2373 |
| - Beiträge zur Geschichte der Universität Rostock                  | ISSN 0232-539X |
| - Rostocker Beiträge zur Hoch- und Fachschul-<br>pädagogik         | ISSN 0233-0539 |
| - Rostocker Informatik-Berichte                                    | ISSN 0233-0784 |
| - Studien zur Geschichte der deutsch-polnischen<br>Beziehungen     | ISSN 0233-0687 |
| - Rostocker Forschungen zur Sprach- und Literatur-<br>wissenschaft | ISSN 0233-0644 |
| - Agrargeschichte                                                  | ISSN 0863-2170 |
| - Migrationsforschung                                              | ISSN 0863-1735 |
| - Manuskripte zur Rostocker Universitätsgeschichte                 | ISSN 0863-1727 |
| - Rostocker Universitätsreden                                      |                |

### Bezugsmöglichkeiten

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

Ferner sind die Hefte im Rahmen des Schriftentausches über die Universität Rostock, Universitätsbibliothek, Tauschstelle, Universitätsplatz 5, Rostock, 0-2500, zu beziehen.