

## *Erwartungen (Hypothesen)*

Der empirische Teil dieser Arbeit folgt einem heuristischen Anliegen. Allgemein wird für die Anwendung der vier genannten Modelle auf die Wunschdaten erwartet, daß die Struktur unter den Wünschen in den resultierenden Bildern sichtbar wird.

Im günstigsten Fall sollten die Wünsche im Anschluß an die Modellierung im Sinn dieser aufzufindenden Struktur klassifiziert werden können. Mit Blick auf die Klärung des Begriffs Wunsch wäre damit zwar die jeweilige *differentia specifica* zur Trennung zwischen den Klassen von Wünschen noch nicht benannt. Wenn jedoch bereits eine Klassifikation vorliegt, steht man einer mit der Bestimmung der *differentia specifica* abgeschlossenen Begriffsbestimmung für den Wunsch sehr nahe. Erwartungen an die Resultate können nur im Licht der jeweiligen Modellierungen formuliert werden.

## Abbildungen im Euklidischen Raum

Die in der Literatur zu findenden intendierten Anwendungen der NMDS, wie beispielsweise die Landkartendaten bei Borg (1980), haben in der Regel die Eigenschaft, daß die Lösungsdimensionen gut zu interpretieren sind. Im Allgemeinen tritt diese günstige Situation aber leider nicht auf. Die bei Borg untersuchten Objekte, Städte in Deutschland, sind tatsächlich durch zwei Dimensionen hinreichend zu beschreiben, nämlich durch die Dimensionen Nord-Süd und Ost-West. Eine derartig klare Datenlage besteht in der Regel jedoch nicht.

Wenn für Daten geometrische Repräsentationen gesucht werden, geschieht dies unter Zuhilfenahme von Annahmen über die Beziehung von Datenaspekten zu spezifischen Konstellationen im Bildraum. So wird beispielsweise bei faktoranalytischen Modellen der Winkel von Vektoren oder die Projektionen von Vektoren aufeinander als Parameter für bivariate Kovariationen interpretiert. Degerman (1972) weist in diesem Zusammenhang jedoch darauf hin, daß ein geometrisches Repräsentationsmodell durch drei Quellen beeinflusst wird:

1. Eigenschaften der Daten, also der eigentliche Gegenstand einer Analyse,
2. Strukturen, die das gewählte analytische Modell vorgibt,
3. invariante Merkmale des gewählten Abbildungsraumes.

Der Einfluß der letzten beiden Einflußquellen ergibt sich allein aus dem gewählten Vorgehen und ist unabhängig von der zu untersuchenden empirischen Datenlage. Vor diesem Hintergrund ist die Prüfung unumgänglich, ob die gefundene Repräsentation auch wirklich auf Re-

striktionen durch die Daten zurückzuführen ist. Anderenfalls besteht die Gefahr, einem Methodenartefakt aufzusitzen. Eine hervorzuhebende unter diesen Gefahrenquellen stellt die Degeneration einer Lösung dar.

### *Das Degenerationsproblem*

Young (2001) fordert für die NMDS allgemein Nicht-Degeneration, daß nämlich der Abstand zwischen zwei Punkten dann und nur dann den Wert  $d = 0$  haben soll, wenn die Objekte, die durch die Punkte repräsentiert sind, identisch sind.

–Degeneration :  $d_j=0 \Leftrightarrow i = j.$

Von der Degeneration einer Lösung ist wiederum genau dann zu sprechen, wenn zwei verschiedene Objekte keinen Abstand besitzen oder für äquivalente Objekte ein Abstand entsteht.

Degeneration:  $\neg (d_j=0 \Leftrightarrow i = j) \Leftrightarrow ((d_j=0 \Leftrightarrow i \neq j) \vee (d_j \neq 0 \Leftrightarrow i = j))$

Das Problem der Degeneration kann durch die vorliegenden Algorithmen nicht intern gelöst werden, da sich der Umstand der Degeneration nicht notwendigerweise ungünstig auf die in den Algorithmen minimierten Kennwerte auswirkt. Die vorliegenden Optimierungsalgorithmen sind unter Umständen blind gegenüber Degeneration.

Ein Hilfsmittel zur Aufdeckung von Degeneration stellt das Shepard-Diagramm dar, ein Streudiagramm von Distanzen versus transformierte Dissimilaritäten. Idealerweise sollten die Punkte des Streudiagramms annähernd die Gestalt einer durchgehenden Linie aufweisen. Bilden sie eine Stufe oder eine L-Form ist dies ein wichtiger Hinweis auf Degeneration. Die zugehörige NMDS-Lösung kann dann praktisch nicht sinnvoll interpretiert werden. Wilkinson (1996, S. II-51) merkt dazu an: "Publish it and you will be excoriated by the clergy." Dieses allgemeine Problem gilt zwangsläufig auch für ein Unfolding im Rahmen der NMDS. Borg (1981, S. 225) klagt in diesem Zusammenhang, das „Modell erweist sich skaliertechnisch als ziemlich tückisch, d.h. insbesondere als anfällig für verschiedene Degenerationsformen“.

Bei HOMALS stellt sich eine Degeneration als mangelnde Diskrimination der Items dar. Diese schlägt sich auch auf die Modellanpassung nieder, die als Summe der Eigenwerte der Repräsentation festgelegt ist. Wenn die Items nicht diskriminieren, die optimalen Scores sich also in einer kleinen Umgebung versammeln, ist der von den Daten umschriebene Raum nur gering ausgedehnt. Ein Charakteristikum einer Raumausdehnung ist eine kleine Determinante und somit eine numerisch geringe Ausprägung der Eigenwerte.

Allerdings sind die Übergänge von relevanter Datenstruktur zur Degeneration fließend. So kann es durchaus informativ sein, wenn die Distanzen von Objektpunkten minimal sind, falls dadurch sinnvoll ein Cluster oder eine Äquivalenzklasse repräsentiert wird. Heiser (1989) verweist auf zwei Degenerationsformen, denen ein Informationsgehalt nicht vollständig abgesprochen werden kann. Im eindimensionalen Raum ergibt sich eine „2+3-Konfiguration“, wenn bei einer Untermenge von Personen  $P_1$  ein Präferenzobjekt  $A$  konsistent abgelehnt wird und dies auch für eine Untermenge von Personen  $P_2$  für das Objekt  $B$  der Fall ist (Abbildung 20).

Innerhalb  $P_1$  und  $P_2$  alternieren die Präferenzen bezüglich der Objekte  $\{C, \dots, Z\}$  stark. Das Verhältnis der Abstände  $d(A, P_1)$  zu  $d(B, P_2)$  steht in Funktion des Verhältnisses der Umfänge der Personenteilmengen.

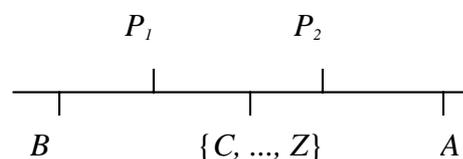


Abbildung 20: Eine „2+3-Konfiguration“ (Heiser 1989)

Existieren drei Personenuntermengen mit den genannten Eigenschaften der konsistenten Ablehnung jeweils eines Präferenzobjektes, läßt sich die Struktur unter Umständen besser in zwei Dimension darstellen und es ergibt sich eine „Dreieck-Stern-Konfiguration“ (Abbildung 21).

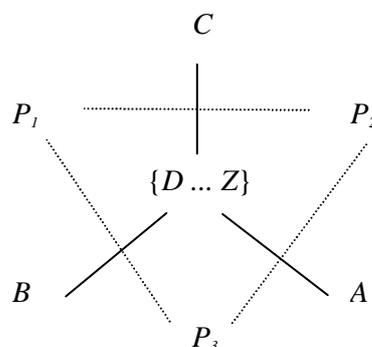


Abbildung 21: Eine „Dreieck-Stern-Konfiguration“ nach Heiser (1989)

Die „Dreieck-Stern“- wie auch die „2+3“-Konstellation lassen sich in dem Sinn interpretieren, daß sich die Menge der Präferenzobjekte nur partiell zur Diskrimination zwischen den Personen eignet. Diese Formen partieller Degeneration lassen sich auch in Räumen höherer Dimension finden. Analog zu ihrer Lage in den Eckpunkten eines Dreiecks in zwei Dimensionen würden sich die Personenpunkte bei einer dreidimensionalen Lösung entsprechend in den Eckpunkten eines Tetraeders auffinden lassen. Mindestens vier Objektpunkte würden an den Strahlenden eines vierstrahligen Sterns liegen, während sich alle übrigen Objektpunkte am

Schnittpunkt der Strahlen im Mittelpunkt der Lösung befänden. Sollte sich bei einer Unfol-  
dinglösung im Rahmen einer NMDS oder HOMALS jedoch die Konstellation ergeben, daß  
eine Punktwolke die Personen und eine zweite Punktwolke die Präferenzobjekte repräsentiert,  
so kann die Lösung nur als trivial bezeichnet werden, denn daß Personen und Objekte dis-  
junkte Klassen darstellen, ist stets bereits vor einer Modellierung bekannt, denn Personen und  
Objekte werden zweifellos als elementefremde Mengen verstanden.

### *Punktconstellationen nach Degerman*

Degerman (1972) liefert für den Euklidischen Raum interpretationsrelevante Grundstrukturen,  
deren Kenntnis eine Voraussetzung zur korrekten Einschätzung der Datenlage sein können.  
Dabei wird ausschließlich die Konstellation der Datenpunkte an sich betrachtet, ohne daß  
externale Attribute der Daten eine Rolle spielen. Carroll (1980) spricht in diesem Zusammen-  
hang von einer internalen Analyse. Es wird von drei elementaren räumlichen Grundstrukturen  
ausgegangen, diskrete Klassen C, Dimension D und Sphäroide S (siehe Abbildung 22).

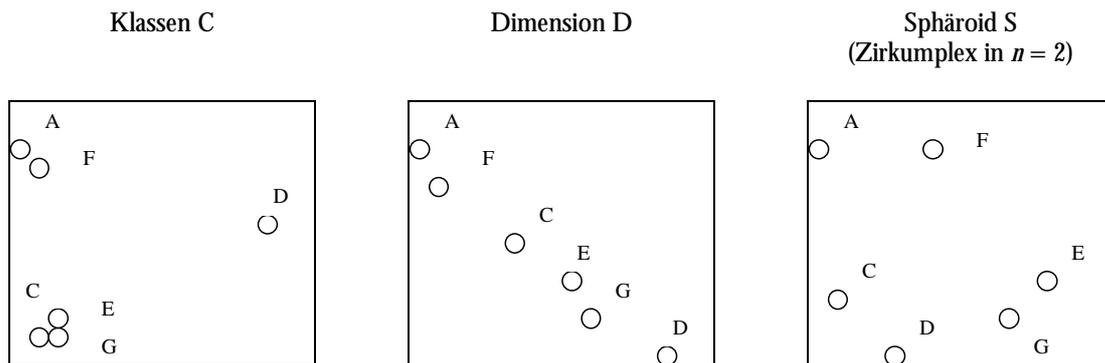


Abbildung 22: Elementare Punktconstellationen im Sinne Degermans (1972)

Auf diesen drei Grundstrukturen sind zwei Operationen definiert: orthogonale Komposition  
im Sinne eines Cartesischen Produkts („ $\times$ “) sowie eine Verschachtelung („ $>$ “).

Es ergeben sich zusätzlich zu den drei Basisstrukturen somit 15 abgeleitete Strukturen (Tabelle  
17 und Tabelle 18).

Die Kompositionen sind folgendermaßen zu verstehen: Wenn eine Struktur  $A$  in  $n$  Dimensio-  
nen  $r$  Punkte umfaßt und eine Struktur  $B$  in  $m$  Dimensionen  $s$  Punkte beinhaltet, gilt, daß die  
zusammengesetzte Struktur  $rs$  Punkte in  $(n + m)$  Dimensionen umfaßt. Damit stellt sich zu-  
nächst die Frage nach der Mindestgröße von  $r$  und  $s$ , um die Basisstrukturen identifizieren zu  
können. Auf der Basis einer solchen Abschätzung kann dementsprechend der Mindestumfang  
des Datenmaterials  $rs$  kalkuliert werden, der notwendigerweise bereitgestellt werden muß, um

eine der vorgeschlagenen zusammengesetzten Strukturen sichtbar zu machen, gegeben, die Daten weisen eine bestimmte Struktur auch auf.

Tabelle 17: Kartesisches Produkt räumlicher Basisstrukturen nach Degerman (1972)

×	Klassen	Dimension	Sphäroid
Klassen	„Paradigmenstruktur“ (ähnlich Facetten)		
Dimensionen	Quantitative Klassenstruktur	Hyperraum (Euklidischer Raum)	
Sphäroid	Klassen-Zirkumplex-Struktur	Hyperzylinder	Spherex

Tabelle 18: Verschachtelung räumlicher Basisstrukturen nach Degerman (1972)

>	Klassen	Dimension	Sphäroid
Klassen	Hierarchische Klassenstruktur	Merkmalprofil	Zirkumplexklassen
Dimensionen	Intensive Klassenstruktur	Guttman-Skala (Simplex)	Radex (konzentrische Sphäroide)
Sphäroid	Hypersphärische Klassen	Stereographische Abbildungsstruktur	Ringex

Da aber die Anwendung von NMDS und HOMALS genau die Frage klären soll, welche Struktur für vorliegende Daten am angemessensten ist, wirft eine zu kleine Anzahl von Datenpunkten das Problem auf, komplexere Strukturen unter Umständen nicht aufdecken zu können. Darüber hinaus entfällt der Vorteil der Anschaulichkeit bei der Kombination von mehrdimensionalen Grundstrukturen, wenn die Dimensionszahl größer als drei ( $d > 3$ ) gerät. Läßt sich allerdings eine der komplexen Konstellationen identifizieren, so ermöglicht dies eine Dekomposition in Grundstrukturen im Sinne der beiden Degermanschen Operationen Produkt „×“ und Verschachtelung „>“. Unter Berücksichtigung der jeweils anzunehmenden Operationen stellt diese Reduktion auf Basisstrukturen eine erhebliche Erleichterung für die Interpretation der gefundenen räumlicher Konstellationen dar. In diesem Sinne ist die erste Erwartung an die Lösungen zu verstehen:

***Hypothese 1:***

Für die beiden in Euklidische Räume abbildenden Verfahren NMDS und HOMALS wird erwartet, daß für die Daten mindestens eine der zusammengesetzten oder einfachen Degermanschen Punktkonstellationen in geeigneter Dimension gefunden werden.

## *Facetten*

Bekannter als der quasi an der Gestalt der Punkterepräsentation orientierte Interpretationsansatz Degermans ist jedoch die auf Guttman (1959) zurückgehende Facettentheorie. Facetten sind „Merkmal[e] zur Unterscheidung der Elemente eines Gegenstandsbereiches in verschiedene Typen“ (Borg, 1996, S. 229). Analog zu den unabhängigen Variablen in einem experimentellen Design dienen die Facetten einer Stratifizierung des betrachteten Datenmaterials. Im Unterschied zu Degerman kommen hier also Vermutungen über relevante Merkmale des Gegenstandsbereiches ins Spiel. Benutzt man mehrere Facetten, so erzielt man eine multiple Typisierung (bei  $m$  Facetten „a  $m$ -way simultaneous classification“ Guttman, 1992, S. 595). Daraus ergeben sich sowohl Vorteile bei der Konstruktion von Itemmaterial für Fragebogen und Tests wie auch für heuristische Untersuchungen.

Der offensichtlichste Effekt der Facettierung liegt in der Verbesserung der Übersichtlichkeit im Untersuchungsbereich. Darüber hinaus tragen die a priori gesetzten Facetten den Charakter von Hypothesen. So wie sich Facetten bei der Konstruktion eines Fragebogens möglicherweise als nicht diskriminativ herausstellen können, so besteht auch die Möglichkeit, daß sie sich unter Umständen als wertlos für die Klassifikation im Rahmen einer Heuristik erweisen. Guttman meint zur Rolle, die Facetten im wissenschaftlichen Prozeß einnehmen können: „Developing a fruitful facet design for content is an evolutionary process ... . There is always room for correction and extension, based on cumulative research.“ (1992, S. 595).

Für den ersten Schritt bei der Formulierung eines Abbildungssatzes von Facetten müssen diese zunächst nur dem Kriterium genügen „Sinn zu machen“. Dieser erste Schritt fußt auf plausiblen Aussagen darüber, daß ein untersuchtes Objekt einer Facette zugehört. Der Wahrheitsgehalt dieser Aussagen bleibt dabei vorläufig offen. Erst durch die Anwendung von Modellen, die in Euklidische Räume abbilden, erhalten diese Aussagen empirischen Gehalt, d.h. sie können falsifiziert werden. Mit der Formulierung von Facetten liegt eine Hypothesen zur Topologie des Abbildes vor. Die Zuordnung der Objekte in Facetten wird nur genau dann als gelungen betrachtet, wenn sich die Punkte, welche die Objekte repräsentieren, in isotonen Regionen der Repräsentation lokalisieren lassen. Unter isotonen Regionen versteht man in diesem Zusammenhang Partitionen des Repräsentationsraumes, die ausschließlich Objektpunkte genau einer Facette beinhalten. Borg (1997) spricht von einer isotonen Region auch als einem „Struktupel der Punkte“ und differenziert in Abhängigkeit von den Bildpartitionen in polare, modulierende und axiale Facetten. Die Eigenschaften dieser basalen Facettentypen finden sich in Abbildung 23.





$$G = (V, X)$$

$$x = (u, v) \text{ mit } x \in X \text{ und } u, v \in V.$$

Äquivalent dazu läßt sich ein Graph auch durch eine auf der Eckenmenge  $V$  definierten Nachbarschafts-Abbildung  $\text{Adj}$  von der Ecken- in die Potenzmenge von  $V$  beschreiben (Golubovic, 1980).

$$\text{Adj}: V \rightarrow \wp(V)$$

$$\text{mit } (v, w) \in X \Leftrightarrow w \in \text{Adj}(v)$$

Die durch das Verfahren HICLAS generierten Graphen unterscheiden sich vom Hassediagramm der FPA sowohl durch die entsprechende Eckenmenge als auch in den Eigenschaften der jeweiligen Abbildung  $\text{Adj}$ . Wegen der verschiedene Bedeutungen müssen beide in unterschiedliche Weise interpretiert werden.

### *HICLAS-Graphen*

Bei HICLAS-Lösungen existieren zwei Modi von Ecken, ein Modus bezieht sich hier auf Wünsche, der andere auf die wünschenden Personen. Die Interpretation basiert auf den drei Modellrestriktionen von HICLAS: Verbundenheit, Äquivalenz und Hierarchie. Die Äquivalenzbedingung hat Auswirkungen auf die Anzahl der Ecken des Graphen. Maximal möglich sind bei  $n$  untersuchten Personen und  $m$  Wünschen  $n + m$  Ecken. Erweisen sich jedoch einige Objekte als äquivalent, so wird die Anzahl der Ecken reduziert und die einander äquivalenten Objekte werden zu Bezeichnern ein und derselben Ecke. Im trivialen Fall existieren nur zwei Ecken in HICLAS-Graphen, jeweils eine Ecke für die Äquivalenzklasse aller Personen und für die aller Wünsche. Ein solcher Graph führt zu keinem Informationsgewinn sondern muß als degeneriert bezeichnet werden. Die Verbundenheit schlägt sich in den Übergängen zwischen den beiden simultanen Teilen des HICLAS-Graphen nieder, dem personenbezogenen und dem wunschbezogenen.

Im Mittelpunkt der Interpretation steht jedoch die Hierarchie sowohl bei den Personen als auch bei den Wünschen. Die maximal mögliche Anzahl von Hierarchieebenen entspricht dem Schein-Rang der Lösung. Auf der Wunschseite des Graphen bedeutet das Vorliegen hierarchischer Beziehungen, daß ein Wunsch auf der niedrigeren Ebene den Wunsch oberhalb in der Hierarchie impliziert. Fehlt das Element der höchsten Hierarchieebene, so zerfällt der entsprechende Modus in durch die jeweiligen Assoziationsbereiche charakterisierte, mehr oder weniger überlappende Cluster. Dabei ist sowohl der Überlappingsgrad der Cluster als auch der Umfang von Besetzungen der Knoten innerhalb der Cluster von Belang.

### *Hypothese 3:*

Es wird erwartet, daß sich auf der Basis einer Verbundenheitsstruktur von Ecken für Personen und Wünsche mindestens eine Hierarchiebeziehungen zwischen Ecken des Wunsch- oder des Personenmodus in HICLAS-Graphen auffinden läßt.

### *Überlappungen im Hassediagramm*

Die Ecken im Hassediagrammen einer FPA-Lösung sind Muster von Itemreaktionen und somit Charakteristika von Personen. Für jedes Item J lassen sich jedoch Teilgraphen, sogenannte Gebiete  $Geb(J)$  identifizieren. Ein Gebiet eines Graphen ist allgemein eine Unter-  
menge der Eckenmenge  $V$ , deren Elemente bestimmte Eigenschaften aufweisen. Die Ecken des Gebietes sind durch einen Kantenzug mit einer maximalen Anzahl von Kanten verbunden. Für Gebiete  $Geb(J)$  im Hassediagramm lautet die relevante Eigenschaft der Ecken, daß genau diejenigen Ecken verbunden sind, in deren Muster das betreffende Item J gelöst ist. Bei einer gegebenen Anzahl  $m$  von Items in der Analyse gilt  $j = 1$  für jedes Muster  $\langle a...j...m \rangle$  in  $Geb(J)$ . Ein Gebiet  $Geb(J)$  umfaßt mindestens  $m$  Ecken. Dies folgt aus einer Eigenschaft des Pseudogeradenarrangements als Dualgraph zum Hassediagramm. Dort wird die Item J repräsentierenden Pseudogerade von allen andern  $(m - 1)$  Pseudogeraden genau einmal geschnitten. Wenn die Schnittpunkte der anderen  $(m - 1)$  Pseudogeraden untereinander in dem Halbraum bezüglich J liegen, der durch die Orientierung  $j = 0$  bezeichnet ist, wird der andere Halbraum, mit  $j = 1$  bezeichnet, in  $(m - 1) + 1$  isotone Regionen geteilt. Für diesen Fall resultieren genau  $m$  Ecken in  $Geb(J)$  des korrespondierenden Hassediagramms (ein Beispiel für vier Items findet sich in Abbildung 24).

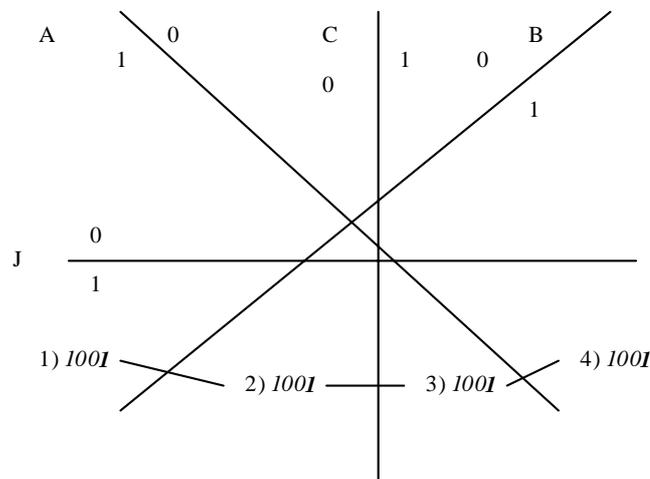


Abbildung 24: Wenn sich die drei Items A, B, C eines Pseudogeradenarrangements von  $m = 4$  Items im Halbraum für  $j = 0$  schneiden, so wird der  $(j = 1)$ -Halbraum in  $m = 4$  isotone Regionen geteilt. Das Gebiet  $Geb(J)$  im dualen Hassediagramm besitzt dementsprechend vier Ecken.

Liegen Schnittpunkte der anderen  $(m - 1)$  Pseudogeraden auch im J-bezüglichen Halbraum mit  $j = 1$ , so besitzt das Gebiet  $\text{Geb}(J)$  mehr als  $m$  Elemente, weil der Halbraum mit  $j = 1$  dann in mehr als  $(m - 1) + 1$  isotone Regionen geteilt wird.

Eine Überlappung  $\text{Üb}(\text{Geb}(A), \dots, \text{Geb}(J), \text{Geb}(K), \dots)$  von Gebieten ist eine Untermenge der Eckenmenge  $V$  des Hassediagramms und als solche genau dann nichtleer, wenn es mindestens ein Muster gibt, in denen die jeweiligen Items alle gelöst sind und somit gilt:  
 $\text{Üb}(\text{Geb}(A), \dots, \text{Geb}(J), \text{Geb}(K), \dots) \neq \emptyset \Leftrightarrow \exists \langle a \dots jk \dots m \rangle$  mit  $(a = 1 \wedge \dots \wedge j = 1 \wedge k = 1 \wedge \dots)$ .

Es ergibt sich ebenfalls aus den Eigenschaften des Pseudogeradenarrangements, daß die Gebiete zweier Items  $\text{Geb}(J)$  und  $\text{Geb}(L)$  sich paarweise stets bezüglich mindestens einer Ecke überlappen. Die sich schneidenden Pseudogeraden erzeugen mindestens vier isotone Regionen. Eine Region liegt nicht im Gebiet von den Items J oder L, eine Region beinhaltet mindestens eine Ecke der Überlappung  $\text{Üb}(\text{Geb}(J), \text{Geb}(L))$ , die beiden anderen Regionen sind durch mindestens ein Muster entweder des Gebiets  $\text{Geb}(J)$  oder des Gebietes  $\text{Geb}(L)$  gekennzeichnet. Demzufolge gibt es bezüglich des Gebietes  $\text{Geb}(J)$  mindestens eine Ecke, die nicht in der Überlappungsmenge  $\text{Üb}(\text{Geb}(J), \text{Geb}(L))$  liegt und vice versa eine Ecke bezüglich  $\text{Geb}(L)$ , die ebenfalls nicht Element von  $\text{Üb}(\text{Geb}(J), \text{Geb}(L))$  ist. Aus letzterem folgt, daß im Rahmen einer FPA-Lösung niemals ein Itemgebiet ein Teilgebiet eines anderen Itemgebietes sein kann, da stets eine Ecke von  $\text{Geb}(J)$  außerhalb von  $\text{Geb}(L)$  liegt.

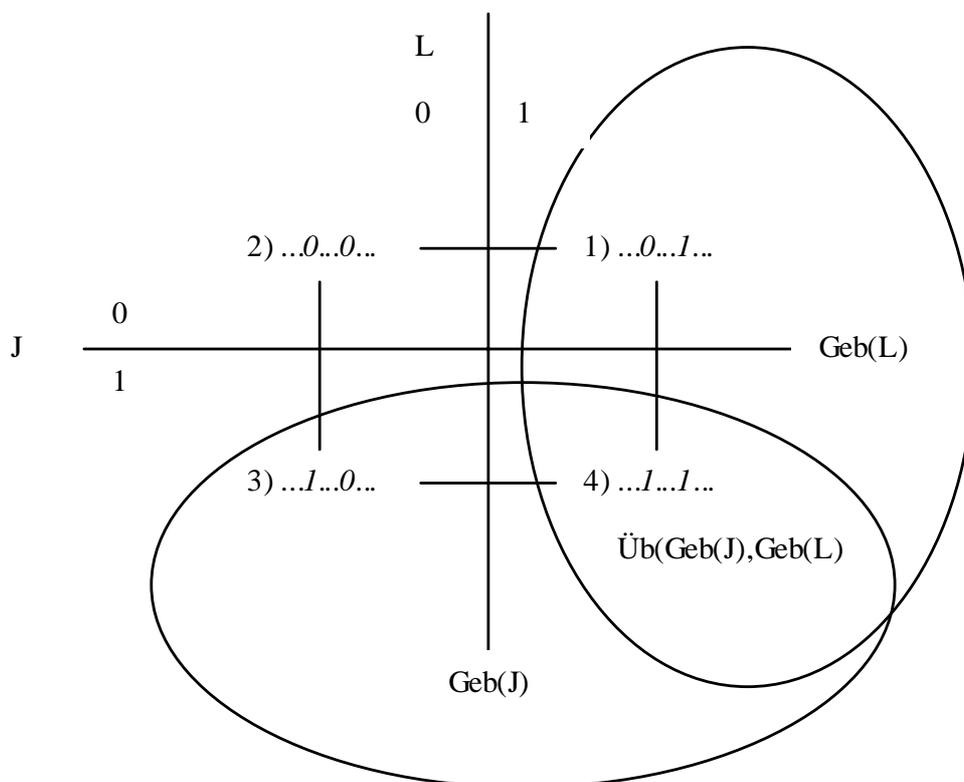


Abbildung 25: Zwei sich schneidende Pseudogeraden für Items J und L erzeugen vier Teilräume, in denen sich jeweils mindestens eine isotone Region mit genau einer der vier möglichen Lösungsmuster befindet. Das hier mit 4) bezeichnete Muster ist stets Element von  $\text{Üb}(\text{Geb}(J), \text{Geb}(L))$

Empirische bedeutsam ist somit nicht die Existenz, sondern das Ausmaß der Überlappung von Itemgebieten. Von starker paarweiser Überlappung soll gesprochen werden, wenn die Anzahl der Elemente der Überlappung zweier Items der Anzahl maximal möglicher Ecken annähernd entspricht. Die maximal mögliche Anzahl von Ecken in der Überlappung von Itemgebieten wird dabei vom Minimum der Anzahl von Ecken in den überlappenden Gebieten begrenzt. In diesem Sinn schöpft eine starke Überlappung das kleinere der beiden Itemgebiete maximal oder zumindest annähernd maximal aus.

starke Überlappung :

$$\text{Üb}_{\text{stark}}(\text{Geb}(J), \text{Geb}(L)) : \quad |\text{Üb}(\text{Geb}(J), \text{Geb}(L))| \rightarrow \min(|\text{Geb}(J)| - 1, |\text{Geb}(L)| - 1)$$

Von einer schwachen Überlappung ist zu sprechen, wenn die Anzahl der Elemente von  $\text{Üb}(\text{Geb}(J), \text{Geb}(L))$  nahe dem Minimum paarweiser Überlappung liegt. Nach den oben beschriebenen Modelleigenschaften der FPA muß mindestens eine Ecke in jeder paarweisen Überlappung liegen. Aufgrund dessen ist eine Ecken der minimale Umfang einer paarweisen Überlappung. Tritt eine Überlappung mit einem Umfang annähernd diesem Minimum auf, so soll sie schwache Überlappung heißen.

schwache Überlappung :

$$\text{Üb}_{\text{schwach}}(\text{Geb}(J), \text{Geb}(L)) : \quad |\text{Üb}(\text{Geb}(J), \text{Geb}(L))| \rightarrow 1$$

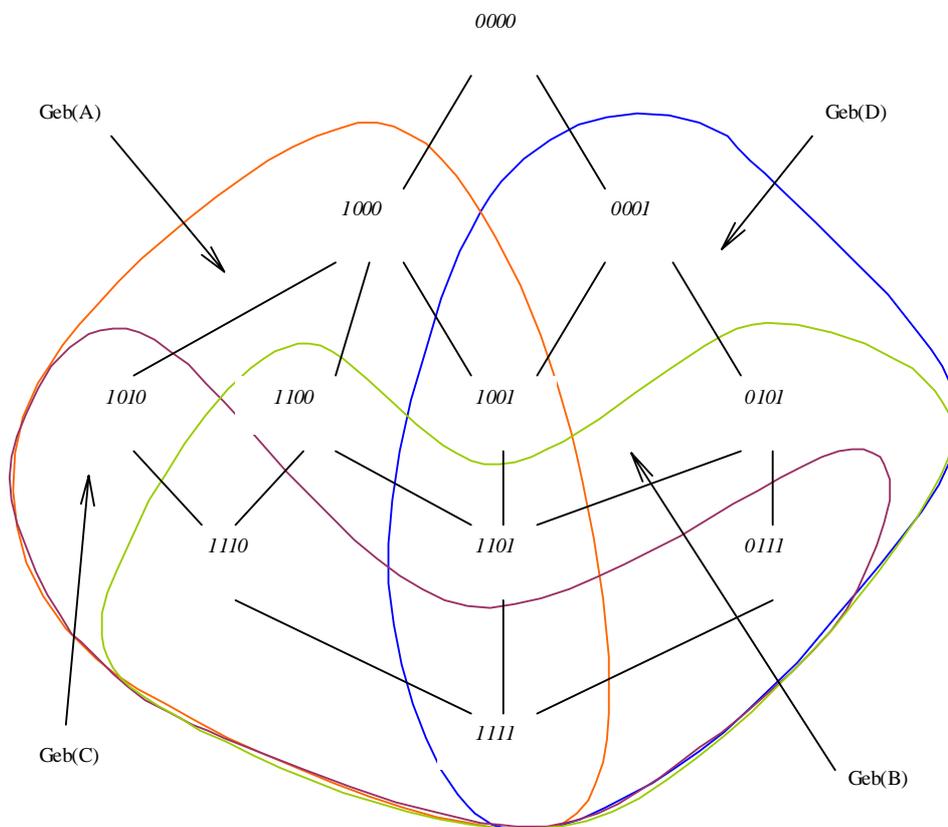


Abbildung 26: Ein Hassediagramm (aus Abbildung 19) und die Gebiete aller vier Items. Die Überlappungen  $\text{Üb}(\text{Geb}(C), \text{Geb}(B))$  und  $\text{Üb}(\text{Geb}(C), \text{Geb}(A))$  sind stark.

In Abbildung 26 findet sich ein Hassediagramm, bei dem die Gebiete der einzelnen Items gekennzeichnet sind. Die jeweiligen Umfänge der Gebiete und paarweisen Überlappungen sind in Tabelle 19 dargestellt. Hier erweisen sich die beiden Überlappungen  $\text{Üb}(\text{Geb}(A),\text{Geb}(C))$  und  $\text{Üb}(\text{Geb}(B),\text{Geb}(C))$  als stark. Die hier schwächste Überlappung ergibt sich hinsichtlich  $\text{Üb}(\text{Geb}(B),\text{Geb}(D))$ .

Tabelle 19: Umfang paarweiser Überlappungen in Abhängigkeit der Umfänge der Itemgebiete bezüglich Abbildung 26

Itemgebiet	Geb(J)	In Überlappung mit ...			
		Geb(A)	Geb(B)	Geb(C)	Geb(D)
Geb(A)	7	-			
Geb(B)	6	4	-		
Geb(C)	4	3	3	-	
Geb(D)	6	3	4	2	-

Lassen sich für einige Items jeweils starke paarweise Überlappungen untereinander feststellen, so kann das als Indikation für Konsistenz der durch diese Items konstituierten Teilmenge gewertet werden.

Finden sich darüber hinaus andere Items, die jeweils schwache Überlappungen zu den Items der konsistenten Teilmenge aufweisen, so stellt dies einen Ausdruck von Diskrimination zwischen der konsistenten Teilmenge und den anderen Items dar.

Treten beide Umstände auf, so können diese zur Identifikation von Itemclustern genutzt werden.

***Hypothese 4:***

Es wird erwartet, daß sich aus den FPA-Lösungen auf Überlappung basierende Aussagen über gemeinsam auftretende Nennungen oder den gegenseitigen Ausschluß von Wünschen ableiten lassen.