



fzg

INSTITUT FÜR DIGITALE
BILDVERARBEITUNG UND GRAPHIK

DIBAG - BERICHT No. 2

DESBOD - SYSTEMSPEZIFIKATION

M. RANZINGER, W. KAINZ, F. LEBERL

1. ZWISCHENBERICHT
PROJEKT DESBOD

DIGITALE ERFASSUNG, SPEICHERUNG UND
BEARBEITUNG ORTSBEZOGENER DATEN.

PROJEKTBEGINN:
1. MAI 1981

PROJEKTBEAUFTRAGTER :
FORSCHUNGSGESELLSCHAFT-
JOANNEUM

PROJEKTVERANTWÖRTLICHER:
F. LEBERL

GRAZ, 1. NOVEMBER 1981

KOPIEN DIESES BERICHTES KÖNNEN VOM INSTITUT FÜR DIGITALE
BILDVERARBEITUNG UND GRAPHIK AM FORSCHUNGSZENTRUM GRAZ,
WASTIANGASSE 6, 8010 GRAZ, TEL.: (0316)/82 5 31
BEZOGEN WERDEN.

PREIS: ÖS 130,-- OHNE VERSAND.

PROJEKT DESBOD

Digitale Erfassung, Speicherung und Bearbeitung Ortsbezogener
Daten

1. Zwischenbericht

S Y S T E M S P E Z I F I K A T I O N

Projektbeauftragter
Forschungsgesellschaft Joanneum

Projektverantwortlicher
F. Leberl

Projektbeginn
1. Mai 1981

Autoren dieses Berichtes
M. Ranzinger
W. Kainz
unter Mitarbeit von
F. Leberl

Abteilung für Digitale Bildauswertung
und Verarbeitung graphischer Daten
Forschungszentrum Graz

1. November 1981

ZUSAMMENFASSUNG

DESBOD ist ein Forschungs- und Entwicklungsprojekt zum Aufbau eines nahezu operationellen Systems der digitalen interaktiven Erfassung, Speicherung und Bearbeitung von ortsbezogenen Daten, also eines sogenannten Geoinformationssystems. Die Entwicklung ist durch die Anstrengungen zur Erfassung des Naturraumpotentials ausgelöst und begründet. Somit richtet sich DESBOD gerade auf die Bedürfnisse dieses Arbeitsgebietes.

Es ist eine Entwicklungsperiode von 4 Jahren vorgesehen. DESBOD wurde am 1. Mai 1981 in Angriff genommen. Dieser Bericht beschreibt den geplanten Umfang von DESBOD, ist das Ergebnis einer 6-monatigen Spezifikationsphase und eröffnet die Periode der Programmerstellung.

Es ist vorgesehen, analog vorgegebene ortsbezogene Daten in computerlesbare Form umzusetzen, wobei sowohl händische Verfahren wie die X, Y-Digitalisierung, als auch automatische Systeme verwendet werden können (Rasterscanner). Daraus wird eine Datenbank auf Vektor- und auf Rasterbasis erstellt, welche laufend zu halten ist. Weiters entsteht ein Auswertesystem, mit dem der Benutzer auf die Daten zugreift sowie ein Werkzeug zur wirkungsvollen Datenanalyse, z.B. mit Mitteln der Geostatistik, zur Verfügung hat.

Wegen der Vielzahl bestehender Vorarbeiten, welche in jüngster Zeit vor allem im nordamerikanischen Raum und in Japan entstanden sind, besteht die Erwartung, daß DESBOD trotz des erklärten Forschungscharakters zu einem so weit wie möglich und sinnvoll operationellen System wird.

ABSTRACT

DESBOD is a research and development project to create a near-operational system for interactive digitizing, storing and processing of spatial data, a so-called geo-information system. DESBOD is caused and motivated by Austrian efforts to map and study the potential of rural areas (suitability mapping). This is why DESBOD is particularly oriented towards the needs of those involved in these mapping efforts.

A 4-year period is planned for the development, the project started on 1 May 1981. This report describes the planned scope of the project as it stands after a 6-month specification period. It opens the process of computer programming.

DESBOD will provide the tools (a) to digitize analog geographical data, using either a manual X, Y digitizer or an automatic scanning process, (b) to establish a data base using both vector and raster data formats with the required editing capabilities, and (c) to access the data by a user so that he has a means for efficient spatial analysis, e.g. using the methods of geo-statistics.

Numerous recent research projects, essentially in North-America and Japan, lead one to hope that DESBOD can result in a near-operational software system, in spite of the fact that it is a research project rather than the development of an industrial product.

INHALTSVERZEICHNIS

	Seite
EINLEITUNG	1
TEIL 1: Automatisierte Geographische Informationssysteme	
1.1 Definitionen	4
1.2 Datencodierung und Aufbereitung der Eingabe	6
1.3 Rasterverfahren	7
1.4 Vektorverfahren	12
1.5 Vergleich der Möglichkeiten für die Datenerfassung	17
1.6 Datenstrukturen in Geoinformationssystemen	22
1.6.1 Rasterstrukturen	22
1.6.2 Vektorstrukturen	27
1.6.3 "Hybride" Datenstrukturen	32
1.7 Datenverwaltungssystem	34
1.8 Datenauswertung und Analyse	36
1.9 Datenausgabe	39
1.9.1 Geräte zur Erzeugung von Hardcopies	39
1.9.2 Softcopysichtgeräte	43
1.9.3 Diskussion	45
TEIL 2: Überblick über das zu erstellende System DESBOD	
2.1 Einführung	46
2.2 Hardware	47
2.3 Gliederung von DESBOD	49
2.4 Datenerfassungssystem	54
2.4.1 Gesamttablauf	54
2.4.2 Erfassung der Themabeschreibungen	58
2.4.3 Erfassung von Grundrißelementen	61
2.4.4 Erfassung von Bildern	66

	Seite
2.5	Die DESBOD Datenbank 68
2.5.1	Begriffsdefinitionen 68
2.5.2	Allgemeine Überlegungen zum Datenbankentwurf 70
2.5.3	Struktur für Grundrißelemente 75
2.5.3.1	Topologische Relationen für Grundrißelemente in Vektorstruktur 78
2.5.3.2	Topologische Relationen für Grundrißelemente in Rasterstruktur 85
2.5.4	Struktur für Themalelemente 87
2.5.4.1	Relationen zwischen Themalelementen 90
2.5.5	Verbindung von Grundriß- und Themalelementen 92
2.5.6	Datenverwaltung 95
2.6	Analyse und Ausgabe von Daten 97
2.6.1	Übersetzer für die Kommandosprache 100
2.6.2	Auskunfts- und Kontrollsystem 100
2.6.3	Datenauffindungs- und Manipulationssystem 101
2.6.4	Datenanalyse 103
2.6.5	Ausgabe der Daten 109
ANHANG A: Darstellung von Zwischenergebnissen der Systemrealisierung	
A.1	Einleitung 112
A.2	Beschreibung der Funktionen 112
A.3	Beispiele 113
ANHANG B: Die Mathematik der Karten	
B.1	Einleitung 118
B.2	Topologie 119
B.3	Graphentheorie 124
B.4	Verbandstheorie 124
B.5	Schluß 124

ANHANG C: Polynomische Mehrparametertransformation 125

Literaturverzeichnis 127

E I N L E I T U N G

Dieser Bericht beschreibt den Inhalt und Umfang des Projektes DESBOD, wie es nach 6 Monaten Arbeit definiert ist. Dieses Projekt dient dazu, im Auftrage der Steiermärkischen Landesregierung und des Bundesministeriums für Wissenschaft und Forschung während einer Periode von 4 Jahren ein umfassendes System zu entwickeln, um ortsbezogene Daten insbesondere aus der Naturraumpotentialkartierung digital zu erfassen, zu speichern, maschinell zu bearbeiten, auszuwerten und darzustellen.

DESBOD ist daher ein Projekt zur Entwicklung eines umfassenden digitalen Geo- oder Landinformationssystems, ein Konzept, welches seit Beginn der siebziger Jahre in den Erdwissenschaften rasch an Bedeutung gewinnt. Es ist ja bekannt, daß die Kartierung jedes nationalen Territoriums von verschiedenen staatlichen, aber auch privaten Stellen in großer Vielfalt betrieben wird. Dabei ist das Ziel der meisten Kartierungen auf einen speziellen Zweck gerichtet, z.B. für Geologie, Hydrologie, Forstwesen, Landwirtschaft, Weinbau, Topographie, Katasterwesen usw.. Die Ergebnisse dieser sehr umfassenden und vielfältigen Arbeiten bestehen bis heute nicht in einem einheitlichen Format, es ist sehr mühsam, diese Ergebnisse miteinander zu verbinden, um in einer übergeordneten Analyse Aussagen von fachübergreifender Bedeutung zu gewinnen.

So stellen wir für die derzeitige, prädigitale Lage im erdwissenschaftlichen Kartierungswesen die folgenden drei großen Probleme fest:

- Unverträglichkeit der Arbeitsergebnisse aus verschiedenen geowissenschaftlichen Fachgebieten;
- fehlende Laufendhaltung jener Kartierungen, wo veränderliche Erscheinungen erfaßt werden, oder fehlende Verfeinerung, wo nur gröbere Ergebnisse vorliegen;
- Mangel an wirksamen Verfahren zur Analyse der Datenvielfalt aus unterschiedlichen aber zusammenwirkenden Fachbereichen.

Die Disziplin, welche in DESBOD durch die computermäßige Umsetzung und Speicherung dieser vielfältigen Daten auferlegt wird, ermöglicht die Einführung rigoroser, eindeutig definierter Datenformate, welche nunmehr sehr wohl die Verknüpfung der verschiedenartigen Daten erlauben und begünstigen. Überdies wird durch die (Teil-) Automatisierung des Kartiervorganges Kapazität frei, um das zweite große Problem in Angriff zu nehmen, nämlich die Laufendhaltung und Verfeinerung. Als Beispiel für die Leistungssteigerung sei nur das Problem der Vervielfältigung von Kartenunikaten in beschränkter Auflagezahl genannt, welches in herkömmlichen Umgebungen sehr arbeitsintensiv ist, jedoch im Digitalfall weitgehend automatisiert werden kann.

Schließlich bietet die Computerverarbeitung von ortsbezogenen Daten vielfältige Möglichkeiten zur wirkungsvollen Datenauswertung durch Verwendung automatischer Abfragen und Darstellungen sowie Anwendungen von Verfahren der Geostatistik.

Es ist somit einleuchtend, daß gerade in umfassenden komplexen Problemstellungen die Konzepte der automatischen Geoinformationssysteme auf reges Interesse stoßen. Die Verwaltung des Naturraumes beinhaltet solche Problemstellungen, wo eine Vielzahl geowissenschaftlicher Daten berücksichtigt werden muß. Jedoch sind auch die kleinen Schritte der geowissenschaftlichen Kartierarbeit wegen der geräte- und verfahrenstechnischen Möglichkeiten des Computerwesens vom Grunde auf neu zu überdenken. Weltweit zeugt eine umfangreiche Literatur, daß in industrialisierten Ländern die Möglichkeit geprüft wird und die Realisierung entsprechender Arbeiten eingesetzt hat, um die Bearbeitung von Daten über die Erdoberfläche zu automatisieren. Die im Projekt DESBOD zu leistende Arbeit muß auf diese bestehenden Erfahrungen eingehen.

Der vorliegende Bericht gliedert sich in zwei Teile. Der erste Teil erläutert grundsätzlich Geoinformationssysteme (GIS) und Möglichkeiten des Computerwesens in Anwendungen auf ortsbezogene Daten der Geowissenschaften. Obwohl die gestellte Aufgabe durch die Naturraumpotentialkartierung motiviert und begründet ist,

geht die Darstellung im ersten Teil der Arbeit darüber hinaus und beschreibt die Grundlagen der graphischen Informationsverarbeitung im Rahmen von automatischen GIS.

Der 2. Teil des Berichtes spezifiziert das im Rahmen des Projektes zu erstellende Geoinformationssystem, welches mit DESBOD bezeichnet wird. Der Zweck dieser Spezifikation ist eine umfassende, jedoch kompakte Beschreibung der Möglichkeiten des zu entwickelnden DESBOD.

In einem Anhang werden einige gesonderte Konzepte und Erläuterungen dargestellt, die den Rahmen des Hauptberichtes sprengen würden.

Wie sich aus dem Text und den Literaturangaben entnehmen läßt, ist der vorliegende Bericht das Ergebnis intensiven Literaturstudiums, Briefwechsels mit internationalen Zentren der GIS-Technologie und Prüfung verfügbarer Programmdokumentationen von Systemen, welche vor allem in den USA bestehen. Es werden heute offenstehende Möglichkeiten der Computertechnik (Hardware) und Programmierung (Software) berücksichtigt. Es besteht daher aus der Sicht der DESBOD-Entwickler die Hoffnung, daß diesem Bericht ein modernes, wohlüberlegtes, umfassendes und zukunftsweisendes Geoinformationssystem zugrunde liegt.

Die Leistungsfähigkeit jedes Computerprogrammsystems kann jedoch nicht auf Grund der Programmspezifikation, sondern muß in seiner Anwendung beurteilt werden. Es ist jedoch sowohl für die interne Planung und Konzipierung von DESBOD wie auch für den wirkungsvollen Dialog mit den späteren Nutzern, den Auftraggebern und Fachkollegen notwendig, frühzeitig eine Gesamtbeschreibung des geplanten Endproduktes vorzulegen. Dieses Dokument wird in weiteren Projektphasen durch erweiterte Darstellungen, detailliertere Benutzeranleitungen von Programmteilen und schließlich durch Erfahrungsberichte mit bestehenden Systemteilen ergänzt werden.

Der vorliegende erste Projektbericht ist eine Einladung an den Leser zu konstruktiver Kritik, welche zur Verbesserung des Gesamtkonzeptes und Einbindung in die bestehenden wissenschaftlich-technischen Strukturen Anlaß gibt.

TEIL 1

Automatisierte Geographische Informationssysteme =====

1.1 Definitionen

Ein Automatisiertes Geographisches Informationssystem (GIS) dient der Erfassung, Speicherung, Fortschreibung, Auswertung und Darstellung ortsbezogener Daten der Erde aus verschiedenen, miteinander zu verbindenden Fachgebieten, wobei diese Daten zum Großteil mit einem digitalen Computer verarbeitet werden.

Ein Benutzer kann das System selektiv nach den darin gespeicherten Daten abfragen, oder aus diesen Daten unter Benützung von EDV-Geräten Resultate ableiten.

Zu den Charakteristiken eines GIS zählen sein Inhalt, seine Datenstrukturen, seine möglichen Auswertungen, die dafür benutzten Algorithmen und die für sein Funktionieren benötigte Hardware (Weber, 1978).

Der interne Aufbau eines GIS erlaubt eine Unterscheidung in 5 Teilsysteme: (Siehe Fig. 1.1)

- a) Datencodierung
- b) Aufbereitung der Eingabe
- c) Datenverwaltungssystem
- d) Datenauswertung und Analyse
- e) Datenausgabe

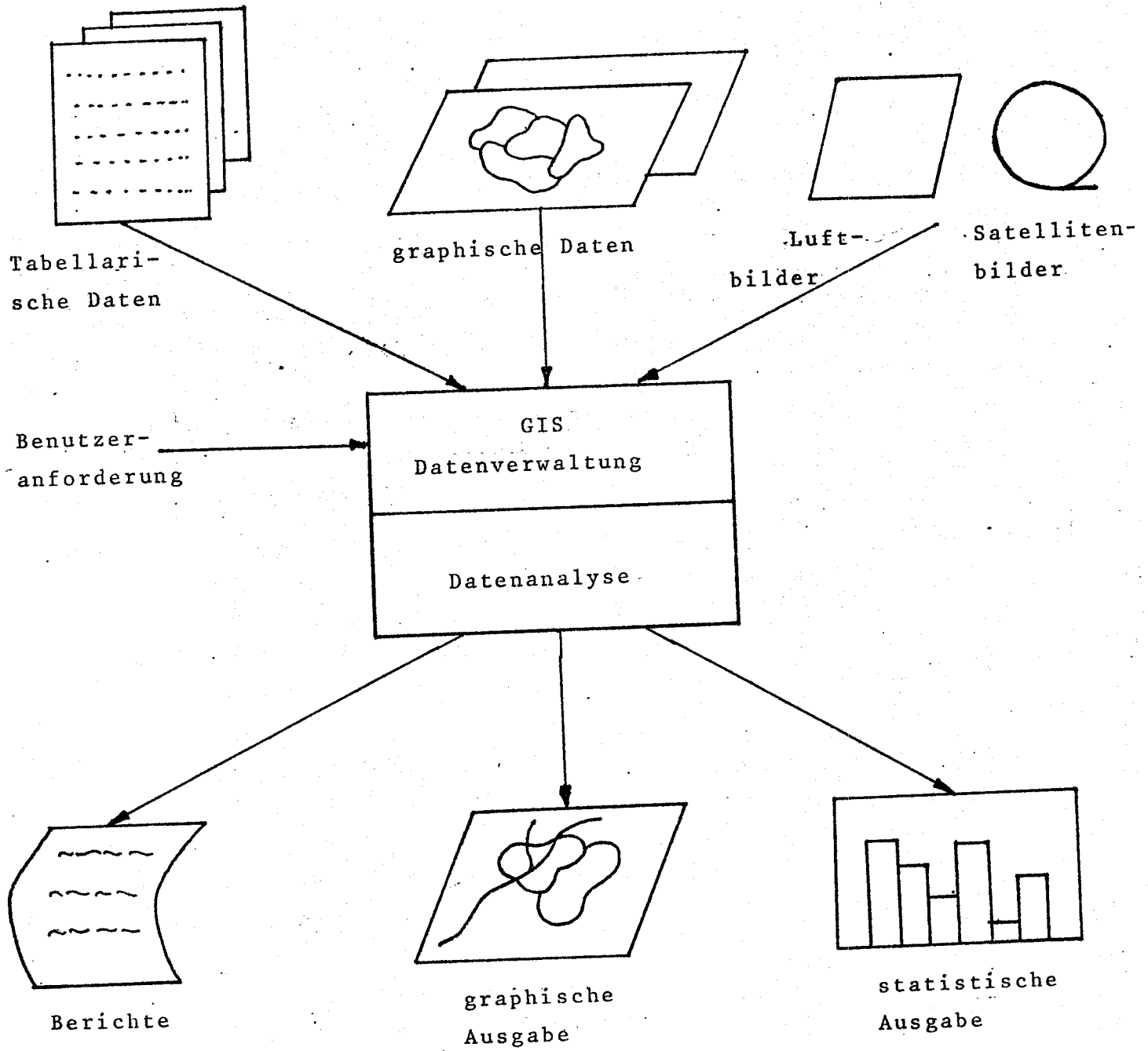


Fig. 1.1
Funktion eines GIS (Knapp, 1978)

1.2 Datencodierung und Aufbereitung der Eingabe

Als Eingabe für ein GIS kommen Daten verschiedener Art und aus verschiedenen Quellen in Frage. Um diese Daten in ein einheitliches Schema zu bringen wird ein mehr - (meist drei-, manchmal vier-) dimensionaler Datenraum definiert. Die Daten werden nun nach den Dimensionen gegliedert, denen sie angehören (Weber, 1978).

So sind die GRUNDRISSELEMENTE in den ersten beiden also planimetrischen Dimensionen (x,y) definiert, die THEMAELEMENTE gehören der dritten Dimension (z) und die ZEITELEMENTE der vierten (t) an.

Die Grundrißelemente lassen sich weiter in Punkt-, Linien- und Flächenelemente einteilen, die durch die auf ihnen liegenden Punktfolgen mit ihren ebenen oder geographischen Koordinaten dargestellt werden. Das Beispiel eines Grundrißelementes ist eine Fläche einheitlicher Vegetation oder eine Besitzgrenze.

Die Themaelemente kann man in solche alphanumerischer und numerischer Natur einteilen, wobei sich die numerischen Themaelemente weiter in solche kontinuierlichen oder diskreten Charakters untergliedern lassen. So sind etwa die Vegetationsart "Wald" oder der Name "Radkersburg" alphanumerische Themaelemente, eine genügend dichte Beschreibung des Höhenverlaufs der Erdoberfläche besteht aus kontinuierlichen und die Bevölkerungszahlen gewisser Gebiete aus diskreten numerischen Themaelementen.

Zeitelemente werden nur dann zusätzlich notwendig, wenn bestimmte Themaelemente zu verschiedenen Zeitpunkten verschiedene Werte annehmen (z.B. Temperaturmessungen zu verschiedenen Jahreszeiten, oder Bevölkerungsstatistiken einer bestimmten Region über mehrere Jahre).

Um alle diese Datenelemente in ein GIS einbringen zu können, müssen die Grundrißelemente von analoger in digitale Form umgewandelt und ihnen die entsprechenden Themelemente zugeordnet werden.

Die Erfassung der Themelemente ist eine einfache Aufgabe, wenn diese Daten entweder schon in maschinenlesbarer Form vorhanden sind oder leicht in diese - durch die Verwendung herkömmlicher Dateneingabegeräte für alphanumerische Daten - umgewandelt werden können. Allerdings können bei der Zuordnung von Thema - zu Grundrißelementen gewisse Probleme auftreten. In diesen Fällen müssen besondere technische Lösungen entwickelt werden. Ein Beispiel ist die Erfassung von topographischen Höheninformationen nach der analog-digital-Umwandlung von planimetrischen Höhenschichtenlinien (Grundrißelementen).

Die Verfahren bzw. die technischen Einrichtungen mit denen die Grundrißelemente in maschineninterne Form gebracht werden, zählen zu den wichtigsten Komponenten eines GIS, weil sie die spätere interne Darstellungsform der Daten und deren Genauigkeit - relativ zum Quellenmaterial - mitbestimmen.

Es gibt mehrere Möglichkeiten zur Erfassung der graphischen Datenelemente, welche im folgenden diskutiert werden.

1.3 Rasterverfahren

a) Manuelle Rasterung

(Hydrologic Engineering Center, 1978)

Über die zu erfassende Karte wird ein Raster mit beliebiger Gittergröße gelegt. Eine Rasterzelle repräsentiert

eine Dateneinheit, der - je nach den in ihr enthaltenen Grundrißelementen und den dazugehörigen Themavariablen - ein numerischer Wert zugewiesen wird. Dieser numerische Wert kann einem tatsächlich vorhandenen Wert entsprechen (z.B. 321.4 für die Höhe) oder ein für eine Themavariablen vereinbarter Wert sein (z.B. 14 bezeichnet die Vegetationsart "Wald").

Wenn der Wert bzw. die Klassifikation der Themavariablen für die gesamte Rasterzellenfläche einheitlich ist, gibt es keine Schwierigkeiten bei der Zuweisung eines Wertes für die Rasterzelle. Sehr oft wird jedoch mehr als ein Wert für eine Variable innerhalb der Rasterzelle vorhanden sein (z.B. zwei oder mehr verschiedene Bodentypen in derselben Rasterzelle) und dann muß eine Entscheidung getroffen werden, welcher Wert der Rasterzelle zugewiesen werden soll; für diese Fälle gibt es verschiedene Strategien, von denen hier nur eine stellvertretend genannt wird: Jeder Rasterzelle wird der Wert der Themavariablen zugewiesen, der genau im Mittelpunkt der Rasterzelle auftritt.

Die manuell gerasterten Daten werden dann wie eine Matrix in die Rechenanlage eingegeben, wobei entweder die Werte der einzelnen Rasterzellen zeilenweise angegeben werden, oder eine Rasterzelle nur dann angegeben wird, wenn sich ihr Wert gegenüber der vorigen Rasterzelle geändert hat (run encoding).

Da die manuelle Rasterung eine relativ einfache Art der Datenerfassung ermöglicht und außer einem herkömmlichen Computer keine speziellen Geräte benötigt, wird sie vor allem dort angewendet, wo sich wegen des relativ geringen Datenanfalls die Anschaffung zusätzlicher Hardware nicht lohnt.

Die manuelle Rasterung stand am geschichtlichen Anfang der GIS, wurde also in ersten Experimenten, etwa gegen Ende der sechziger Jahre verwendet. Ein bekanntes Beispiel ist das an der Harvard-Universität entwickelte SYMAP-System (Sheehan, 1979).

Die manuelle Rasterung hat heute noch eine gewisse Bedeutung bei der Erfassung von Daten aus Luftbildern oder direkt im Gelände, wo dieses Verfahren einem systematischen Entnehmen einer statistischen Stichprobe von zweidimensional verteilten Variablen entspricht.

b) Automatische Rasterung

Bei der automatischen Rasterung wird zur Erfassung der graphischen Daten ein Rasterabtastgerät (Rasterscanner) verwendet. Dieses Gerät tastet das Quellenmaterial ab und registriert das Vorhandensein oder Fehlen bestimmter Markierungen. Das Auflösungsvermögen - d.h. die Fähigkeit zwei Punkte als voneinander getrennt zu erkennen, entsprechend also der Rastergröße beim manuellen Rastern - ist je nach Geräte-Modell unterschiedlich, beträgt aber bis zu 40 Punkten pro Millimeter (Turek, Walker, 1979; Nagy 1981).

Manche Scanner können verschiedene Graustufen verarbeiten und produzieren etwa ein 8-Bit Byte für jedes Bildelement (engl. picture element, zusammengezogen zu "Pixel". Ein Pixel ist das Analogon zu einer Rasterzelle beim manuellen Rastern), andere sind auf Schwarz-Weiß-Operationen (binäre Daten) beschränkt. Einige Scanner können auch verschiedene Farben unterscheiden, mehrfarbige Linienzeichnungen in einem oder mehreren Durchgängen abtasten und in getrennte Farbauszüge zerlegen. (Nagy, Wagle, 1979; Nagy, 1981).

Im allgemeinen muß das Eingabematerial für den Scanner speziell vorbehandelt werden, sodaß die aufzunehmenden Elemente der Karte oder des Luftbildes in einer klaren Linienzeichnung vorgegeben sind, die dann vom Scanner verarbeitet wird. Es besteht die Wahl, entweder in einem Vorbereitungsschritt für den Scanner geeignete Vorlagen zu erstellen, oder aber nach dem Rastervorgang interaktiv eine Bereinigung der nunmehr digitalen Daten vorzunehmen. Die Entscheidung für das eine oder andere Vorgehen ist von Fall zu Fall zu treffen. Eine manuelle oder interaktive Datenbehandlung ist meist für verbleibende Fehler, aber auch für die Verknüpfung der Grundriß- mit Themaelementen notwendig.

Für den weiterverarbeitenden Prozeß erscheinen die Ausgabedaten vom Rasterscanner als eine nach Zeilen oder Spalten geordnete Matrix der reflektierten Werte. Bei einem Schwarz-Weiß-Scanner handelt es sich um eine binäre Matrix, die nur die Werte "0" für Weiß und "1" für Schwarz enthalten kann, während bei anderen Scannertypen die Matrizelemente mehrere unterschiedliche Werte annehmen können. Die so erhaltenen Matrizen werden jedoch zweckmäßigerweise vor ihrer endgültigen Speicherung einer Aufbereitung unterzogen werden (Herausfiltern der Störungen, Skeletonierung der Linien, Glättung). Die Größe der ursprünglichen Matrix hängt zwar von der jeweiligen Auflösung des Gerätes ab, der Datenanfall ist aber in jedem Fall sehr groß, sodaß es sinnvoll ist, in einem Nachverarbeitungsschritt datenreduzierende Programme einzusetzen (Nisen, 1980).

Als Beispiele eines Rasterabtastgerätes seien das Modell Kartoscan der Firma MBB - München (Fig. 1.2.a) und das Modell Photomation der Firma Optronics erwähnt (Fig. 1.2.b).

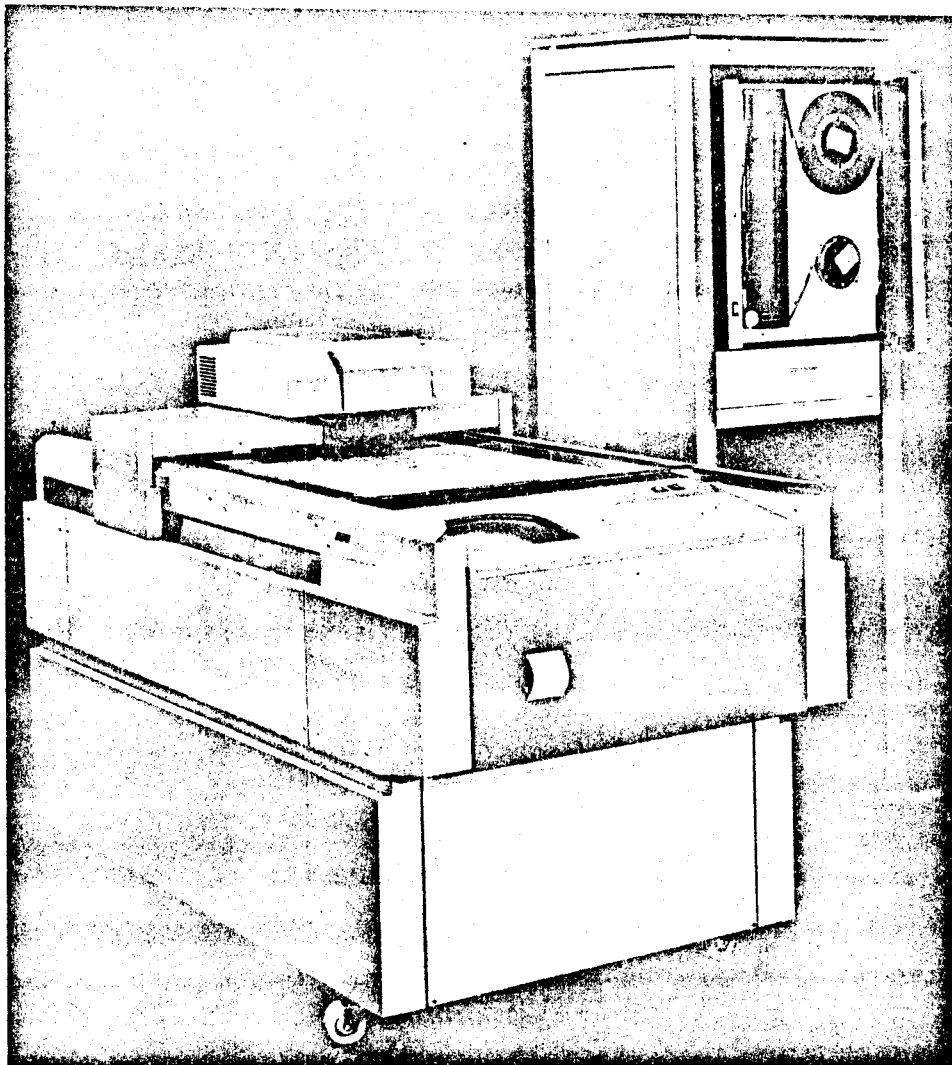


Fig. 1.2.a)

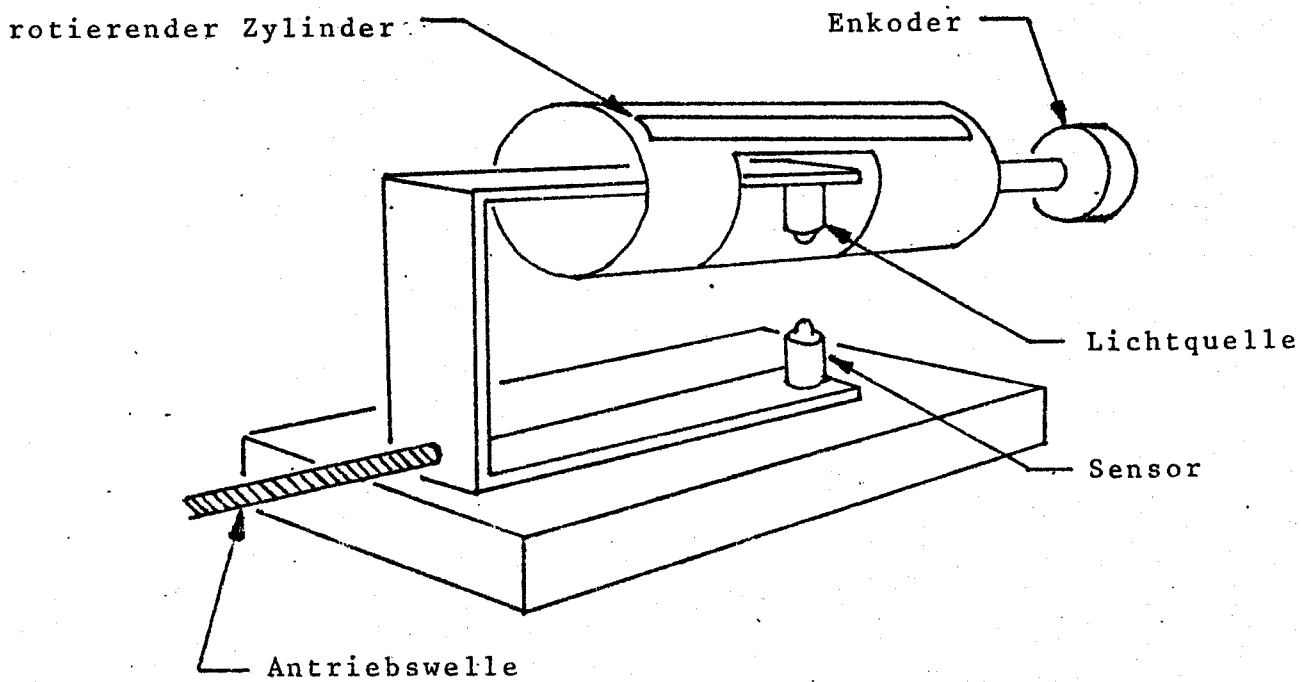


Fig. 1.2.b)

1.4 Vektorverfahren

Vektorverfahren beruhen darauf, daß nicht die gesamte Fläche der Karte oder des Bildes gleichmäßig zur Datenerfassung herangezogen wird, wie das bei Rastern der Fall ist, sondern daß nur die interessierenden Grundrißelemente selbst aufgenommen werden. Es werden die x,y -Koordinaten von Punkten oder von Punktketten registriert, um die relative Lage eines Elementes auf der Karte oder dem Bild festzulegen. Punktelemente werden als einzelne x,y -Koordinaten, Linienelemente als eine Folge von Koordinaten wiedergegeben; für Flächenelemente wird das sie umschließende Polygon als geschlossene Folge von Koordinaten registriert.

Flächenelemente können allerdings auch auf effizientere Art angegeben werden: Wenn für jedes Flächenelement das Begrenzungs-polygon erfaßt wird, werden die Grenzen zwischen den einzelnen Flächenelementen **d o p p e l t** erfaßt, was neben dem erhöhten Aufwand auch zur Folge hat, daß beim Auszeichnen diese Grenzen oft als unschöne, doppelte Konturen erscheinen. Um das zu vermeiden, werden zur Darstellung der Flächenelemente diese Grenzlinien nur **e i n m a l** angegeben, allerdings muß zusätzlich vermerkt werden, welche Flächenelemente durch welche Grenzlinien getrennt werden.

Gekrümmte Linien werden durch eine Folge von kurzen Geradenstücken angenähert, wobei die Genauigkeit von der Länge dieser Geradenstücke abhängt (Hydrologic Engineering Center, 1978).

Zur vektoriellen Erfassung von Daten stehen verschiedene Geräte zur Verfügung: Die am weitesten verbreitete Art ist der Digitalisierertisch (engl. coordinate digitizer, graphic tablet), der manuell betrieben wird (siehe Fig.1.3); der

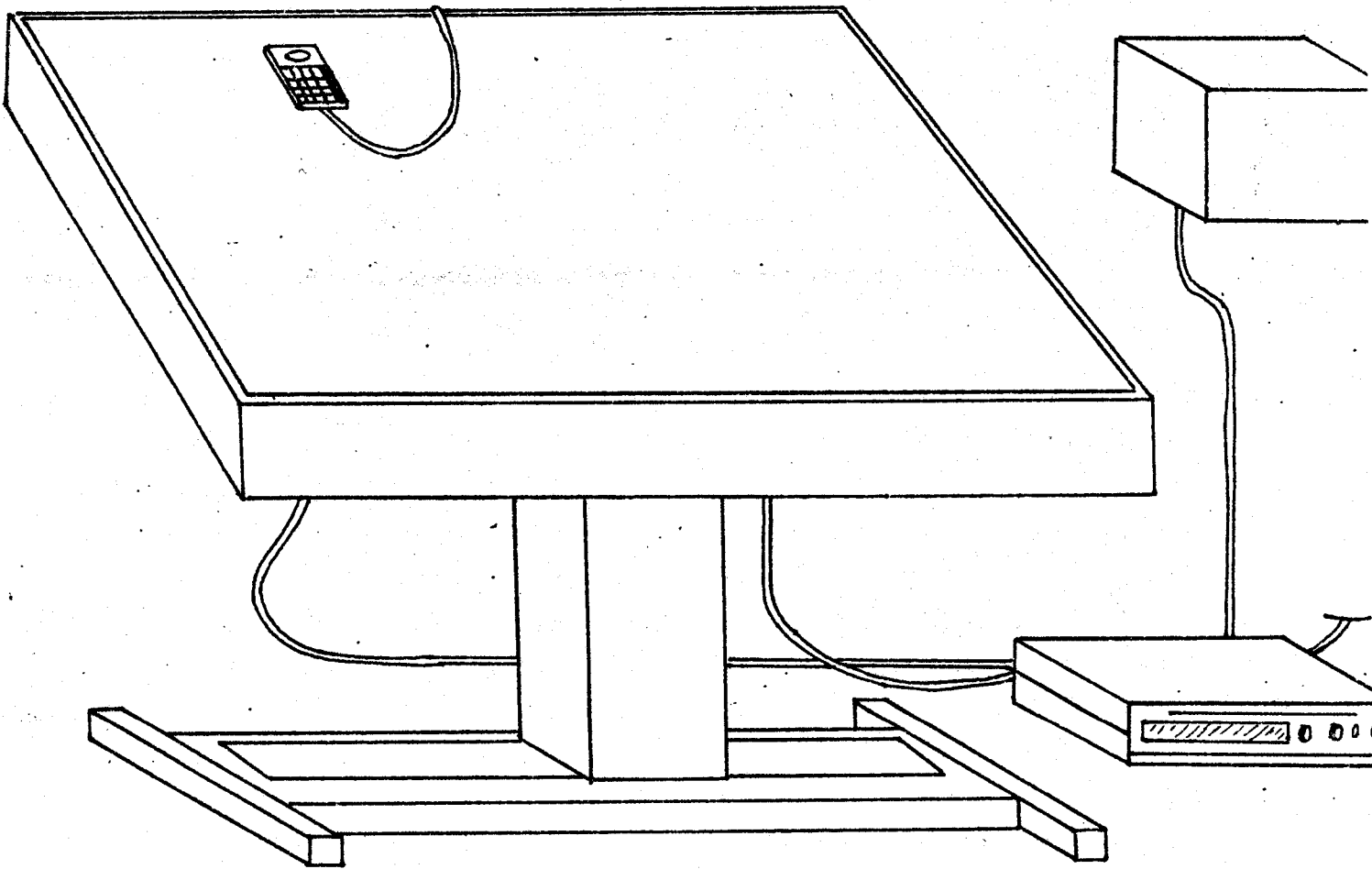


Fig. 1.3
Manueller Digitalisiertisch

Benutzer wählt selbst die Punkte (bzw. Linien) aus, deren Koordinaten aufgenommen werden sollen. Ein weniger bekanntes Gerät ist der automatische Linienvorfolger (engl. line follower, curve follower). Hier folgt ein suchender Sensor automatisch den Linien einer Zeichnung; der Sensor hat sich auch die Kreuzungspunkte (Knoten) von Linien zu merken, um alle von einem Punkt ausgehenden Linien zu erfassen (Nisen, 1980).

Da Digitalisierertische die am häufigsten verwendeten Geräte sind, sollen sie hier näher beschrieben werden.

Digitalisierertische können on-line betrieben werden, d.h. sie schicken die erfaßten Koordinaten direkt an den Computer, oder sie arbeiten off-line, indem sie die Koordinaten auf einem Speichermedium (Magnetband, Platte) ablegen.

Der Cursor (zu deutsch: "Schieber") - das ist die Einrichtung, mit der der Benutzer die zu digitalisierenden Linien abtastet - kann entweder durch einen beweglichen Arm mit mechanischen Vorrichtungen, die die Bewegung dieses Arms in horizontaler und vertikaler Richtung registrieren, mit dem Digitalisiergerät verbunden sein, oder der Cursor ist frei beweglich und die Koordinaten werden von einem elektronischen Netz erfaßt, das in den Digitalisierertisch eingebaut ist. Der Cursor selbst kann ein spezieller Stift sein oder ein Vergrößerungsglas mit Fadenkreuz. Die Größe der Fläche, die digitalisiert werden kann, ohne ein neues Blatt auflegen zu müssen, reicht von 20 x 26 cm bis zu 75 x 100 cm mit einer Auflösung von ca. 0.02 mm.

Die Position des Cursors wird entweder abgenommen, wenn ein Knopf auf der Tastatur des Cursors gedrückt wird (Punktmodus), oder die Koordinaten werden in bestimmten, vorher festgelegten Zeit- oder Wegintervallen abgenommen (kontinuierlicher Modus, "stream mode").

Für die weiterverarbeitenden Prozesse erscheinen die digitalisierten Daten als eine Folge von x- und y-Koordinaten.

Um die digitalisierten Grundrißelemente mit den dazugehörigen Themaelementen zu verbinden, ist es notwendig, neben dem Digitalisieren alphanumerische Informationen einzugeben. Diese Informationen können über die alphanumerische Tastatur eines Terminals oder direkt über die Tastatur des Cursors (wo aber nur wenige Zeichen zur Verfügung stehen) eingegeben werden. Eine andere Möglichkeit ist die Verwendung einer Software-Tastatur (oder -Keyboards), die aus einem Gitter auf dem Digitalisiertisch besteht, wobei jeder Gitterzelle eine bestimmte Zahl, ein bestimmter Buchstabe oder ein spezifisches Symbol (engl. "code" z.B. jeweils für eine Bodenart) zukommt. Statt ständig zwischen der Bildschirmtastatur und dem Digitalisiertisch wechseln zu müssen, kann man dann die zusätzlichen Informationen angeben, indem man einfach mit dem Cursor auf die entsprechende Gitterzelle zeigt. Die x,y-Koordinaten jeder Gitterzelle identifizieren eindeutig den gewünschten Begriff; im weiterverarbeitenden Prozeß können diese Koordinaten entsprechend umgewandelt werden. Für jede neue Anwendung können den einzelnen Gitterzellen des Software-Keyboards neue Begriffe zugewiesen werden (Nagy, Wagle, 1979). Das Software-Keyboard wird auch mit "Menüfeld" bezeichnet (Fig. 1.4).

Es besteht die Möglichkeit die beschriebenen Digitalisiergeräte durch Automationsunterstützung zu verbessern: Man erhält ein Mittelding zwischen Digitalisiertisch und Linienverfolger. Der Benutzer folgt mit dem Cursor nur grob dem Verlauf einer Linie, deren exakte Koordinaten dann vom Gerät selbst z.B. mit Hilfe einer Detektorenanordnung erkannt werden. Durch diese Kombination können bis zu viermal bessere Ergebnisse erzielt werden, als durch rein manuelles Digitalisieren (Anderson, 1980).



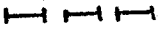







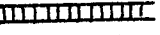


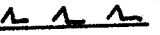
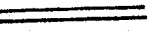

Gemeinde- grenze 	Bezirks- grenze 	Staats- grenze 	Grenz- signatur 1 
Grenz- signatur 2 	Grenz- signatur 3 	Grenz- signatur 4 	Grenz- signatur 5 
Hecke 	Zaun 	Mauer 	Stütz- mauer 
Baumreihe 1 	Baumreihe 2 	Begleit- linie 	Steilrand 

Fig. 1.4

Beispiel eines Menüfeldes für das manuelle Digitalisieren
von Katastermappen

1.5 Vergleich der Möglichkeiten für die Datenerfassung

Bevor man sich für eine Methode zur Datenerfassung entscheidet, ist es notwendig zu wissen, welche Arbeiten zurzeit und in Zukunft durchgeführt werden sollen. Die Fragen, die hierfür beantwortet werden müssen, sind etwa folgende:

Wieviele Kartenblätter sollen erfaßt werden?

Verteilt sich die Erfassung gleichmäßig über einen längeren Zeitraum oder ist zu einem bestimmten Zeitpunkt ein besonders großer Arbeitsanfall zu erwarten?

Wie oft müssen die erfaßten Daten auf den neuesten Stand gebracht werden?

Welche Genauigkeit wird verlangt?

Welche Größe haben die aufzunehmenden Kartenblätter?

Welches Personal steht zur Dateneingabe zur Verfügung?

Daraus ergibt sich eine Aufwandsermittlung für die verschiedenen Digitalisierverfahren, die natürlich sehr unterschiedliche Investitionen erfordern.

Die manuelle Rasterung ist dann einsetzbar, wenn eine statistische Probenentnahme gefordert wird oder ganz spezielle Anforderungen bestehen (Bodenwertschätzung etc.). Eine effiziente Digitalisierung, auch bei manueller Rasterung, erfordert spezielle Hardware. Von den besprochenen Geräten ist das manuelle Digitalisiergerät das billigste (ca. öS 30.000 - bis 300.000,-), der Preis für einen Rasterscanner liegt zwischen öS 300.000 und 1500.000,-, ein komplettes Rasterabtastsystem mit Scanner, Registrierung und Grundsoftware erfordert eine Investition von etwa öS 3000.000,- und ein automatisches Linienverfolgungssystem kostet ebenfalls mehr als öS 3000.000,- (J.Orr, 1978).

Wenn es erforderlich ist, daß bei der Datenerfassung ein hoher Grad an Genauigkeit erhalten bleibt (wie z.B. für Katasterkarten), dann ist ein Rasterscanner oder Linienverfolger zu bevorzugen. Das heißt nicht, daß manuelle Digitalisiergeräte ungenau sind, aber hier spielt der menschliche Faktor eine große Rolle. Hohe Genauigkeiten können auch mit manueller Digitalisierung erreicht werden, wenn nur die graphische Vorlage z.B. photographisch vergrößert wird. Tabelle 1.1 stellt die erreichbaren Genauigkeiten für verschiedene Digitalisierverfahren zusammen.

Tabelle 1.1: Erreichbare Genauigkeiten von Digitalisierverfahren

Gerät	Genauigkeit	Bemerkung
Manueller Digitalisiertisch	± 0.075 mm bis ± 0.150 mm	graphische Genauigkeit eines Zeichners
Rasterabtastung	± 1 Pixel, also bis zu ± 0.025 mm	abhängig von der Auf- lösung und Linien- qualität
Linienverfolgung	± 0.5 Linienbreite	abhängig von Linien- qualität

Wenn das GIS aber zum Beispiel hauptsächlich zur Erzeugung von thematischen Karten herangezogen werden soll, ist ein manuelles Digitalisiergerät hinreichend genau.

Zum Abschluß der Diskussion sei hier eine Tabelle wiedergegeben, die die Überlegungen zusammenfaßt, die für die Wahl der Hardware getroffen werden müssen. (Nisen, 1980).

Wenn die meisten der Angaben auf der linken Seite der Tabelle für das zukünftige System zutreffen, sollte an die Anschaffung eines Raster-scanners oder Linienverfolgers gedacht werden, die rechts stehenden Angaben lassen ein manuelles Digitalisiergerät als geeignet erscheinen.

Tabelle 1.2: Überlegungen zur Wahl eines Erfassungsgerätes für graphische Daten

	Digitalisierverfahren:	
	automatisch	Im wesentlichen nur manuell
Budget	hoch	niedrig
Anzahl der Kartenblätter, die pro Tag verarbeitet werden sollen	mehrere	nicht mehr als drei
Bedienungspersonal	gut ausgebildete Techniker	keine spezielle Ausbildung
verfügbare Computer-Ressourcen	großes Computersystem	größerer Mikrocomputer
Häufigkeit der Benutzung des Erfassungsgerätes	konstante Benutzung	sporadische Benutzung
Art des zu erfassenden Quellenmaterials	gut erkennbare Linien, keine Beschriftung, hohe Komplexität (sehr viele Linien)	verschiedene Kartentypen wenig komplex (nicht viele Linien)

Schließlich muß bei der Wahl des Datenerfassungssystems auch berücksichtigt werden, in welcher Struktur die eingegebenen Daten im GIS gespeichert werden sollen. Das wird unmittelbar einsichtig, wenn man die möglichen Datenstrukturen betrachtet (siehe Kapitel 1.6).

Grundsätzlich läßt sich zwar aus jedem Format, in dem die Daten nach der Erfassung vorliegen, jede interne Datenstruktur erzeugen, nur kann dies mit einem beträchtlichen Umrechnungs- und Umformungsaufwand verbunden sein und einen Verlust an Genauigkeit mit sich bringen.

Für die Datenerfassung zum GIS ist in jedem Falle manuelles Digitalisieren notwendig, auch wenn ein automatisches Verfahren zur Verfügung steht. Es ist immer effizient, für kleine Zusätze, für kleine Datenmengen und für gewisse spezielle Aufgaben eine manuelle Datenerfassung vorzusehen.

Die automatische Analog-Digitalumwandlung ist erst langsam im Begriffe, operationell eingesetzt zu werden. Die Nutzung ist mit der vorhergehenden Einrichtung eines umfassenden Systems verbunden: Erfassungsgerät, Zwischenspeicher, Computer, Software, interaktive Arbeitsplätze zur Datenbereinigung usw. sind einzurichten und stellen insgesamt ein sehr komplexes automatisches Produktionssystem dar.

Langfristig bietet die Rastererfassungstechnologie mehr Perspektiven als die Linienverfolgung (Boyle, 1980). Dies ist auch dadurch belegt, daß es weltweit nur einen Erzeuger von Linienverfolgungssystemen, aber mehr als 20 Erzeuger von Abtastsystemen gibt (Nagy, 1981).

Die Art der vorgegebenen Daten beeinflußt stark die Methode der Digitalisierung. In einer Vorlage, welche wenig graphische, aber viel thematische Daten enthält, kann die manuelle Methode wirtschaftlicher sein als die automa-

tische. Typische Anwendungen automatischer Verfahren sind im Falle dichter, jedoch einfacher Daten gegeben (Höhenlinien, Landnutzungsflächen).

1.6 Datenstrukturen in Geoinformationssystemen

Es gibt zwei Haupttypen von Strukturen, die in einem GIS Anwendung finden können; alle uns bekannten GIS verwenden eine von diesen beiden Strukturen oder eine Kombination aus beiden. Diese zwei Möglichkeiten sind die Rasterstruktur und die Vektorstruktur. An und für sich stellt die Rasterstruktur ein kontinuierliches und die Vektorstruktur ein diskretes Modell der in ihr enthaltenen Daten dar.

Jede der beiden Strukturen ermöglicht die Darstellung aller drei Grundrißelementsarten und kann auf alle Arten von Themalementen angewendet werden.

Im folgenden werden beide Möglichkeiten beschrieben und es wird versucht, ihre Vor- und Nachteile zu diskutieren.

1.6.1 Rasterstrukturen

Das zu bearbeitende Gebiet ist in eine Matrix von Zellen gleicher Größe unterteilt. Für jede Zelle wird eine bestimmte Anzahl von Speicherplätzen reserviert. Der Speicherplatz für jede Zelle enthält entweder die Information, die mit dieser Zelle assoziiert ist (z.B. Vorhandensein bzw. Fehlen eines bestimmten Grundrißelements oder Wert einer Themavariablen), oder einen Zeiger zu dieser Information.

Zu den Vorteilen dieser Struktur zählt ihre Übereinstimmung mit dem Format des Raster-Scanner-Inputs und dem der manuellen Rastereingabe, sowie mit dem Format von auf Rasterbasis arbeitenden Speichern, Bildschirmen und Matrixdruckern. Probleme, die mit der Entfernung verschiedener Grundrißelemente voneinander zu tun haben, können leicht

gelöst werden, da die Entfernung zweier Punkte direkt aus deren Speicheradresse errechnet werden kann. In gleicher Weise können Datenzugriffe, die sich auf die Lage eines Grundrißelements innerhalb eines Gebiets beziehen, ohne zusätzliche Datenmanipulationen realisiert werden. Im allgemeinen ist die Entwicklung von Software für fast jede Anwendung bei rasterstrukturierten Daten einfacher als bei Daten in Vektorstruktur. Die Rasterstruktur ist am geeignetsten für die Beschreibung von Themen, die eine kontinuierliche Bedeckung des Studiengebiets repräsentieren (z.B. Höhenverlauf) (Nagy, Wagle, 1979).

So wäre die Rasterdatenstruktur recht vielversprechend und unkompliziert. Grundsätzlich besteht jedoch das Problem der Datenmenge, wenn das Rasterformat gewählt wird. Die Verarbeitung der großen Zahl von Pixeln - insbesondere bei hoher Auflösung, z.B. bei Digitalisierung mit Raster-scanner - benötigt auf der heute verfügbaren Hardware für die Praxis viel Raum und Zeit. Möglicherweise wird die zukünftige Hardwareentwicklung hier zu einer Begünstigung der Rasterdatenverarbeitung führen, doch ist zu erwarten, daß dies nur in Sonderfällen sinnvoll sein wird.

Aus den oben genannten Gründen ist auch die interne Auflösung der gespeicherten Daten meist geringer als die Auflösung der Eingabe, was natürlich zu einem Genauigkeitsverlust bei der Analog-Digital-Umwandlung führt. Für Grundrißelemente, die den Studienbereich nicht dicht überdecken (wie z.B. Punktelemente), stellt die Abspeicherung in Rasterstruktur eine Speicherplatzverschwendung dar, da ja auch alle Rasterzellen, die keine Elemente enthalten, mit berücksichtigt werden. Für diese Fälle gibt es eine andere Form der Rasterstruktur, die "Laufängenerfassung" (engl. "run-length-encoding") die ein Analogon zur "run encoded" manuellen Rasterung ist. Ein Matrixelement wird nur dann abgespeichert, wenn es einen anderen Wert als das

vorliegende Element aufweist. Da hier bei jedem abgespeicherten Element eine Information über seine relative Lage innerhalb der Matrix hinzukommt, ist diese Struktur der herkömmlichen Rasterstruktur nur dann vorzuziehen, wenn nur wenige relevante Werte in der Matrix vorkommen, oder wenn mehrere (durchschnittlich 4) Matrixelemente hintereinander den gleichen Wert repräsentieren (MLMIS, Minnesota Land Management Information Systems; Tomlinson, 1976).

Eine andere interessante Weiterentwicklung der Rasterstruktur stellt der "Quadtree" dar, der den in der Informatik gebräuchlichen Baumstrukturen entspricht. Diese Struktur beruht auf einer stufenweisen Verfeinerung des Rasters; das gesamte Studiengebiet wird in vier gleich große Rechtecke geteilt, von denen jedes wieder viermal unterteilt wird (Figur 1.5). Dieser Prozess wird für jede Rasterzelle so lange fortgesetzt, bis sich in ihr nur noch Elemente mit dem gleichen Wert bzw. überhaupt keine relevanten Elemente befinden, oder bis die kleinstmögliche Rasterzellengröße erreicht ist.

Obwohl über den Quadtree und seine Einsatzmöglichkeiten in der Bildverarbeitung und in Geoinformationssystemen schon zahlreiche theoretische Veröffentlichungen erschienen sind (vgl. H. Samet, 1981; A. Rosenfeld u.a., 1980), ist er bisher in keinem GIS implementiert.

Die Datenstruktur des Quadtree wird heute im wesentlichen für eine Bearbeitung von Grautonbildern vorgesehen. Im Falle von linienhaften Vorlagen scheint die Verwendung von Vektorstrukturen doch vorteilhaft (Haberäcker, 1981^{*)}, wenn das Ziel die bleibende Speicherung ist (vergleiche Abschnitt 1.6.2).

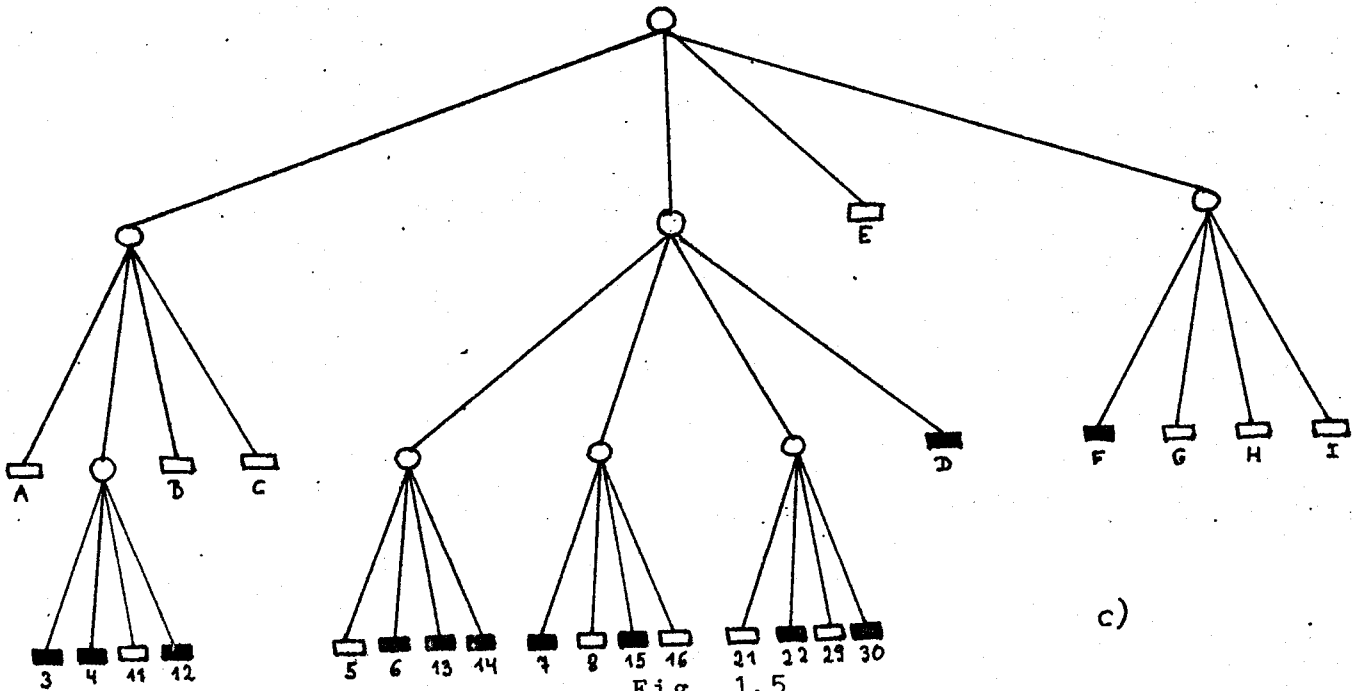
^{*)} Haberäcker P. (1981): Persönliche Mitteilung im Zusammenhang mit der Benutzung von Quadrees für die kurzzeitige Rastererfassung und -bearbeitung von Dokumentationsvorlagen (Computer-Aided Printing).

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64

a)

A	3	4	5	6	7	8
	11	12	13	14	15	16
B	C		21	22	D	
			29	30		
E			F		G	
			H		I	

b)



c)

Fig. 1.5
Quadtree

Bild b) zeigt die Zerlegung in kleinste, nicht mehr weiter unterteilbare Blöcke von der in Bild a) repräsentierten Fläche und Bild c) zeigt den entsprechenden Quadtree. Schwarz bzw. weiß gezeichnete Quadrate stehen für Blöcke, die ganz aus Einsen bzw. Nullen bestehen. Kreise bezeichnen die Verzweigungsknoten (H. Samet, 1981).

Das Beispiel der Figur 1.5 zeigt der Einfachheit halber eine Binärmatrix (besteht nur aus 0 und 1-Werten). Die "Wurzel" des Baumes^{*)} repräsentiert die ganze Matrix, die 4 Söhne der Wurzel stimmen - von links nach rechts gesehen - mit dem nordwestlichen, nordöstlichen, südwestlichen und südöstlichen Quadranten des Feldes überein und die Blattknoten stellen die Blöcke dar, die nicht weiter unterteilt werden müssen.

^{*)} "Bäume" (engl. tree) haben "Wurzeln", "Blätter" usw.. Ein Quadtree ist ein Baum, wobei jeweils 4 Söhne zur Wurzel gehören, und jeder Sohn wiederum 4 Söhne hat, bis ein Blatt erreicht wird.

1.6.2 Vektorstrukturen

In der Vektorstruktur werden "Vektoren" abgelegt. In der graphischen Datenverarbeitung sind dies geradlinige Verbindungen zwischen zwei durch (ebene) Koordinatenpaare x_1, y_1 und x_2, y_2 festgelegten Punkten.

Die einfachste Art der Vektorstruktur ist folgende, im Englischen mit "whole polygon structure" bezeichnete:

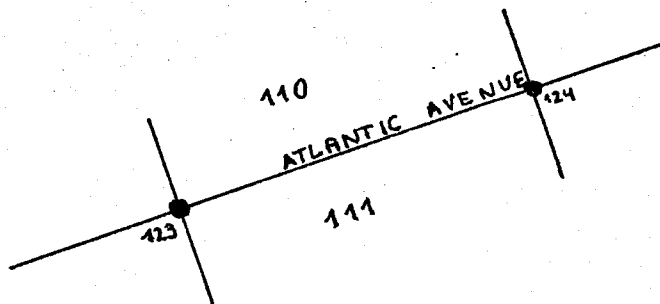
- Jedes Punktelement wird direkt durch seine x,y-Koordinaten beschrieben.
- Jedes Linienelement ist durch eine Folge von x,y-Koordinaten festgelegt.
- Jedes Flächenelement ist durch die x,y-Koordinaten seiner Begrenzungslinien definiert.
- Zu den Koordinateninformationen kommt gewöhnlich zu jedem Grundrißelement noch eine Information (Zeiger oder Referenz), die den Bezug zum entsprechenden Themaelement herstellt.

Diese Struktur hat für Flächenelemente den Nachteil, daß Grenzen zwischen Flächen doppelt gespeichert werden; daher besteht speziell für Flächenelemente eine Reihe anderer Strukturen, denen eines gemeinsam ist: Sie verwenden zur Speicherung von Flächenelementen Linieneinheiten, die untereinander so durch Zeiger verbunden sind, daß daraus die Begrenzungslinien einer Fläche zusammengesetzt werden können.

Die kleinste mögliche Linieneinheit ist ein Segment (auch engl.: segment), das aus einer geraden Linie zwischen 2 Punkten besteht. Ein Beispiel für die Verwendung von Segmenten ist die DIME-Struktur (Dual Independent Map Encoding; siehe DIME, 1974), die vom U.S. Bureau of Census

zu Darstellung von Straßen und Häuserblöcken im städtischen Bereich verwendet wird.

In jeder Stadt werden die Straßenkreuzungen beliebig durchnummeriert und mit ihren geographischen Koordinaten erfaßt. Die Straßenkreuzungen werden als Knoten eines Graphen betrachtet und die Straßensegmente zwischen den Kreuzungen entsprechen den Kanten eines Graphen. Die Häuserblöcke werden ebenfalls durch Nummern gekennzeichnet. Für jedes Straßensegment wird eine Eintragung angelegt, die die angrenzenden Straßenkreuzungen, sowie die zu beiden Seiten liegenden Häuserblöcke enthält. Gebogene Straßen werden in gerade Straßensegmente unterteilt und an den Schnittstellen der Segmente werden Pseudoknoten eingeführt.



110, 111 ... Nummern der rechts
und links liegenden Häuserblöcke

Fig. 1.6

DIME-Struktur (Nagy, Wagle; 1979).

Die DIME-Struktur ist allerdings für Flächen mit unregelmäßigen Begrenzungen schlecht geeignet, da diese Begrenzungen in viele gerade Segmente unterteilt werden müssen; bei einer sehr großen Anzahl von Segmenten ist es aber recht zeitraubend und schwierig, die Begrenzungen für eine Fläche herauszufinden.

Für Karten, die nicht so speziell wie die oben für Städte beschriebene sind, kann man also annehmen, daß an den meisten Punkten nur 2 Segmente zusammentreffen. Die Ketten-Repräsentation (engl. "chain-Code") einer Karte reduziert die Zahl von Linieneinheiten, indem sie nur die Punkte als Knoten berücksichtigt, an denen mehr als zwei Segmente aufeinandertreffen. Eine Anzahl von Segmenten zwischen zwei Knoten wird zu einer einzigen "chain" zusammengefaßt, die jetzt eine Linieneinheit bildet (Knapp, 1978). An dieser Stelle muß darauf hingewiesen werden, daß die Bezeichnungen für die verschiedenen Datenstrukturen in der Fachliteratur keineswegs einheitlich sind - für die chain-Struktur wird im Amerikanischen auch der Name 'Arc-and-Node Structure' gefunden; im Deutschen wird oft - in Anlehnung an die Graphentheorie - die Bezeichnung "Kanten-Knotenstruktur" verwendet.

Auf eine nähere Beschreibung der Kanten-Knotenstruktur wird hier verzichtet, da sie in Kapitel 2.4.3 ausführlich behandelt wird.

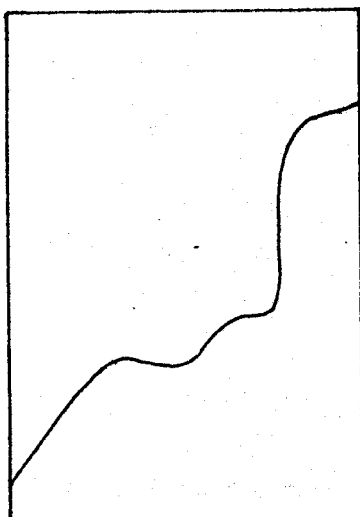
Zusätzlich können bei den beschriebenen vektoriiellen Datenstrukturen noch andere Relationen zwischen den Grundrißelementen berücksichtigt werden (wie z.B. die Angabe bei jedem Punktelement, in welchem Flächenelement es liegt, Angabe aller zu einem Flächenelement benachbarten Flächenelemente, etc.); es würde aber zu weit führen, wenn man alle diese Möglichkeiten hier auführen und für jede eine geeignete Implementierung angeben würde.

Es sei hier nur betont, daß Angaben von Relationen zu einer "relationalen Datenstruktur" führen (Weber, 1978). Ihr Aufbau ist zwar aufwendig, aber sie bieten sehr weitreichende Möglichkeiten der Auswertung (Kropatsch, Leberl, 1981).

Vektorstrukturen können im allgemeinen direkt aus dem Format, wie es von einem Digitalisiergerät geliefert wird, erzeugt werden. Es können natürlich auch Daten von einem Rasterscanner durch entsprechende Umformungen in eine vektorielle Form gebracht werden (Vektorisierung). Wie diese Vektorisierung der Rasterdaten vor sich geht, ist in Fig. 1.7 dargestellt.

Vektoriell organisierte Daten sind in so gut wie jedem Fall speicherplatzeffizienter als Rasterdaten (wenn man bei beiden die gleiche Genauigkeit zu Grunde legt), sodaß i.a. bei der Speicherung eine größere Genauigkeit als bei Rasterdaten beibehalten werden kann. Da für jedes Grundrißelement seine genaue Lage bekannt ist (bei Rasterdaten ist oft nur bekannt, in welcher Rasterzelle ein Grundrißelement liegt), lassen sich bei Auswertungs- und Analyseprozeduren weitaus exaktere Ergebnisse erzielen.

Nachteile der vektoriellen Struktur betreffen vor allem die Komplexität der Software, die für die Verarbeitung und Verwaltung der Daten notwendig ist. Dies gilt z.B. für das Editieren und Fortschreiben der gespeicherten Daten, für Berechnungen von Entfernungen zwischen einzelnen Grundrißelementen, sowie für die Durchführung von geometrischen Mengenoperationen (z.B. Vereinigung oder Durchschnitt von Flächenelementen).



a)

```

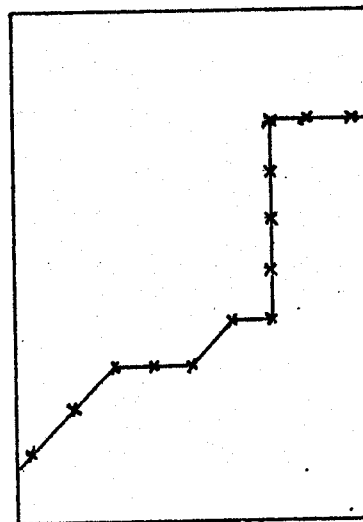
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 1 0 0
0 0 0 0 0 1 1 1 1
0 0 0 0 0 1 1 1 1
0 0 0 0 0 1 1 0 0
0 0 0 0 0 1 1 1 0
0 0 1 1 1 1 1 0 0
0 1 1 1 1 1 1 0 0
1 1 0 0 1 1 0 0 0
1 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
    
```

b)

```

0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 1 1 1
0 0 0 0 0 0 1 0 0
0 0 0 0 0 0 1 0 0
0 0 0 0 0 0 1 0 0
0 0 0 0 0 1 1 0 0
0 0 1 1 1 0 0 0 0
0 1 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
    
```

c)



d)

Fig. 1.7

- a) graphische Vorlage, b) Ergebnis der Rasterung,
 c) Skeletonierung, d) Erstellung der Vektordatenstruktur

1.6.3 "Hybride" Datenstruktur

Wie aus den vorherigen Abschnitten hervorgeht, sind die Vor- und Nachteile der Vektor- und der Rasterstruktur beinahe komplementär. Es liegt daher nahe, eine Vereinigung der beiden Strukturen anzustreben, die die Vorteile beider Datenorganisationen weitgehend ausnutzen kann.

Eine Möglichkeit, eine solche "hybride" Datenstruktur zu verwirklichen wird von Weber (1978) beschrieben:

Die Speicherung der Daten erfolgt hier in einem Baum ähnlich dem Quadtree, nur enthalten die Blattknoten Informationen in Vektorform (lagebezogene Struktur).

Das gesamte Studiengebiet wird wieder mit einem Netz geschachtelter Quadrate mit systematisch kleiner werdenden Seitenlängen überdeckt. Aus diesem Raster wird ein Baum (wie beim Quadtree hat jeder Verzweigungsknoten vier Söhne) aufgebaut, dessen Blattknoten Zeiger enthalten, die auf die Speicheradressen der Grundrißelemente zeigen, die in dem entsprechenden Quadrat liegen. Die Grundrißelemente sind im Speicher in Vektorform abgelegt. Analog zum reinen Raster-Quadtree wird ein Quadrat nur dann weiter unterteilt, wenn mehr als ein Grundrißelement in ihm liegt und die kleinste Quadratgröße noch nicht erreicht ist. Die kleinste Quadratgröße muß natürlich größer sein als die kleinste Rasterzelle bei einer Rasterdarstellung der Karte, da sonst kein Speicherplatzgewinn erzielt werden kann. (Weber, 1978).

Der Speicherplatzaufwand für die hybride Datenstruktur ist größer als für rein vektorielle Strukturen, aber geringer als für Rasterstrukturen; je nachdem, wie die kleinste Quadratgröße gewählt wird und wieviele Grundrißelemente das Studiengebiet enthält, wird er eher dem einen oder anderen Extrem zustreben.

Die beschriebene lagebezogene Datenstruktur ist bisher noch nirgends praktisch angewendet worden und es finden sich auch in der Literatur nur wenige Hinweise auf sie, sodaß über ihre etwaigen Vor- und Nachteile im Vergleich mit den anderen Strukturen nicht viel gesagt werden kann.

Abschließend muß noch folgendes festgehalten werden: Die Strukturierung der graphischen Daten ist nur ein Teil dessen, was beim Aufbau einer Datenbank für ein GIS berücksichtigt werden muß; ebenso wichtig ist die Organisation der alphanumerischen Daten (Themaelemente, allgemeine Beschreibungen etc.), jedoch sind die Möglichkeiten hierfür bereits hinlänglich bekannt.

1.7 Datenverwaltungssystem

Das Datenverwaltungssystem ermöglicht es dem Benutzer, auf die in der Datenbank gespeicherten Daten zuzugreifen.

Ein GIS sollte Datenverwaltungssoftware beinhalten, die folgendes gewährleistet:

- a) Die Fähigkeit des Systems, mehrere Benutzer gleichzeitig zu unterstützen.
- b) Eine Speicherung, die ein schnelles Auffinden und Ändern der Daten gewährleistet.
- c) Nicht-Redundanz der Daten (es werden alle Datenelemente nur einmal gespeichert).
- d) Datenunabhängigkeit, Datenschutz und Datenintegrität.

Eine wirkungsvolle Speicherung, die schnelles Auffinden und Ändern der Daten gewährleistet, ist größtenteils vom Speichermedium und von der Organisation der Daten (Datenstruktur) abhängig. Die Datenstruktur wiederum wird durch die Komplexität der zu bearbeitenden Daten und durch die auszuführenden Datenmanipulationen bestimmt.

Datenunabhängigkeit setzt voraus, daß die Daten und die Anwendungsprogramme, von denen sie verwendet werden, voneinander unabhängig sind, so daß jedes für sich geändert werden kann, ohne daß das andere davon beeinflusst wird.

Datenschutz beinhaltet die Sicherung gegen zufälligen oder gewollten Datenzugriff von dazu nicht berechtigten Personen, sowie die Sicherung gegen unerlaubte Änderung oder Vernichtung der Daten. Datenintegrität bezieht sich auf die Vorkehrungen, die die Daten vor Systemproblemen (sowohl Hardware - als auch Softwareproblemen) absichern.

Das Datenverwaltungssystem soll alle diese Konzepte beinhalten und außerdem die Grunderfordernisse - Schreiben, Lesen und Ändern der gespeicherten Daten - problemlos erfüllen.

1.8 Datenauswertung und Analyse

Die wichtigsten Auswertungen von Geoinformationssystemen werden im folgenden nach Themengruppen gegliedert angegeben; da es jedoch sehr vom Anwendungsgebiet des GIS abhängt, welche Auswertungsmöglichkeiten implementiert werden, ist hier auf eine detailliertere Beschreibung der einzelnen Punkte verzichtet worden, weil dies den Rahmen dieser Arbeit sprengen würde. Es sei hier auf die einschlägige Fachliteratur verwiesen, in der sich zahlreiche, ausführliche Beschreibungen von Auswertungsprozeduren finden lassen. (z.B. Harvard Library of Computer Graphics, 1979, 1980; Davis J.C. and Mc Cullagh M.J. 1975).

- Selektion:

Mit 'Selektion' ist die Bestimmung einer Untermenge des Inhalts eines GIS gemeint, die einem bestimmten Auswahlkriterium entspricht.

Je nach Art des Selektionskriteriums kann man verschiedene Selektionstypen unterscheiden:

- Selektion nach geometrischen Relationen:

Als Selektionskriterium können Ausdrücke wie "nahe bei", "angrenzend", "gelegen in", etc. auftreten.

- Selektion nach Themenrelationen:

Die Selektion erfolgt nach Vorgabe eines Relationstyps und eines Objekts (z.B. alle Orte, die ein Gymnasium besitzen).

- Zusammengesetzte Selektionen:

Aus den Ergebnissen der vorher genannten Selektionen können mit Hilfe von Mengenoperationen neue Objektmengen gewonnen werden.

- Berechnung von Skalaren zu den Grundrißelementen:
Z.B. die Länge von Linienelementen, Umfang und Inhalt von Flächen, etc..

- Geometrische Mengenoperationen:
Bildung des Durchschnitts, der Vereinigung oder der Differenz von Grundrißelementen. (Z.B. die Schnittfläche zweier Flächen).

- Thematische Mengenoperationen:
Äquivalent zu den geometrischen Mengenoperationen für die Themaelemente (z.B. welche Eigenschaften haben mehrere Objekte gemeinsam?).

- Statistische Berechnungen:
Z.B. Histogramme oder Häufigkeitsverteilung, Regression und Korrelation: Schaffung einer Schnittstelle für Standard-Statistik-Programmpakete.

- Generalisierung:
Generalisierung ist notwendig, wenn die Auflösung eines GIS verringert werden soll, d.h. der Gesamtinformationsinhalt wird so herabgesetzt, daß er für eine Darstellung in verkleinertem Maßstab geeignet ist. (Z.B. müssen Linien vereinfacht und relativ zueinander verlagert oder Objekte zusammengefaßt werden).

- Auswertungen von kontinuierlichen Themenüberdeckungen:
Z.B. Ermittlung der Höhe eines Punktes, Hangneigungskarten, Höhenprofile etc..

- Auswertung mittels Operations-Research-Methoden:
Optimierung von Funktionen mehrerer Variablen unter

Nebenbedingungen (z.B. die Ermittlung des kürzesten Weges zwischen 2 Orten).

Alle diese Auswertungs- und Analysemöglichkeiten arbeiten mit den aus der Datenbank gewonnenen Daten; es können damit aber auch neue Dateien innerhalb der Datenbank angelegt werden, um die oftmalige Durchführung der gleichen Routinen zu vermeiden.

1.9 Datenausgabe

Die Ausgabeoperationen eines GIS haben die Aufgabe, die Elemente eines GIS, bzw. die in den Auswertungsprozeduren gewonnenen Ergebnisse, in einer für den Betrachter leicht lesbaren und unmißverständlichen Form darzustellen. Die Darstellungsmöglichkeiten sind weitgehend von der vorhandenen Hardware abhängig, aber gerade für graphische Daten, wie sie in einem GIS vorhanden sind, sollte neben der Ausgabe von Tabellen und Listen auch die Ausgabe von Graphiken und Karten unterstützt werden. (Newman, Sproull, 1979).

Es sind Ausgabeformen als "Hardcopies" (auf Papier, Film usw.) und als "Softcopies" gegeben (auf Bildschirm, Magnetband etc.). Hardcopies sind für die Arbeit mit Papier und Bleistift, zur Lagerung und Publikation notwendig. Darstellungen auf Bildschirm werden hingegen zur schnellen Überprüfung der Daten, von Ergebnissen der Datenerfassung und -auswertung und zur effizienten Mensch-Maschine-Interaktion benötigt.

1.9.1 Geräte zur Erzeugung von Hardcopies

(a) Drucker

Bei beinahe jedem Computer steht als Ausgabegerät ein Drucker zur Verfügung, der primär für die Erstellung von Listen und Tabellen verwendet wird; es können über ihn aber auch graphische Daten in Rasterform ausgegeben werden. Hierbei kommt jedem Rastermatrixelement eine oder auch mehrere nebeneinander- und übereinanderliegende Druckposition(en) zu. Durch verschiedene Drucksignaturen werden die jeweiligen Werte der Matrixelemente wiedergegeben. Durch Überdrucken - mehrmaliges Ansprechen der gleichen Druckposition - kann auch die Darstellung verschiedener Grauwerte erreicht werden.

Vorteile:

- keine besondere Hardware notwendig;
- vom Benutzer einfach zu handhabende Steuerung (keine spezielle Software notwendig);
- beste Präsentation alphanumerischer Daten.

Nachteile:

- geringes Auflösungsvermögen;
- großflächige Darstellungen bedingen umfangreichen Output;
- mögliche Verzerrung des Verhältnisses zwischen Horizontale und Vertikale, weil eine Druckposition nicht gleich breit wie hoch ist (es gibt allerdings Geräte, wo Breite und Höhe einer Druckposition modifiziert werden können);
- durch die Leerflächen zwischen den einzelnen Druckzeilen entsteht ein künstlicher Raster, sodaß Flächen nicht zusammengehörend darstellbar sind;
- für die Erzeugung verschiedener Graustufen kann kein allgemein gültiges Muster angegeben werden, da dieses je nach Druckertyp (vor allem bei Matrix- und Zeilen-drucker) verschieden ist.

(b) Zeichengerät (Plotter)

Plotter benötigen als Eingabe Daten im Vektorformat. Der Zeichenstift (i.a. sind mehrere Zeichenstifte in verschiedenen Farben und Strichstärken vorhanden) des Plotters wird durch die Angabe von x,y-Koordinaten gesteuert. Die Schnittstelle zum Benutzerprogramm wird durch eine Programmbibliothek definiert, in der sich vom Benutzerprogramm aufrufbare Routinen befinden, die die Grundfunktionen, wie Zeichnen, Beschriften, Skalieren, etc., ausführen. Diese Programmbibliotheken enthalten i.a. eine

standardisierte Menge von Funktionen und werden vom Gerätehersteller zur Verfügung gestellt.

Daher ist die Erstellung einer Karte (oder auch anderer Zeichnungen und Graphiken) aus den zugrundeliegenden Vektordaten eine relativ einfache Aufgabe. Wenn die gewünschte Darstellung nicht direkt den Quelldaten entspricht, werden zusätzlich Operationen für das Abschneiden der Vektoren an den Rändern des Zeichenfensters benötigt (engl. "clipping"). Andere wichtige Funktionen sind z.B. das Zeichnen von Symbolen zur Beschriftung, das Anzeigen von Richtungen durch Pfeile und das Schraffieren von Flächen mit beliebigen Mustern.

Vorteile:

- optisch ansprechende Präsentation graphischer Daten;
- standardisierte Schnittstelle (vergleiche CALCOMP-Software); es ist bei Gerätewechsel keine Änderung der Benutzerprogramme notwendig.

Nachteile:

- oft Spezialpapier notwendig;
- hohe Gerätekosten und beschränkte Lebensdauer (8 Jahre?);
- geringere Geschwindigkeit der Plotter im Vergleich zu anderen Geräten.

(c) Elektrostatische Drucker

Eine preisgünstige Alternative zu x,y-Plottern ist der elektrostatische Drucker, bei dem Linien durch Punktfolgen angenähert werden. Die Schnittstelle zwischen Gerät und Benutzer ist analog zu der des Plotters aufgebaut. Der elektrostatische Drucker kann darüberhinaus aber auch für die Ausgabe tabellarischer Informationen verwendet werden und läßt bei der Schattierung von Flächen auch die Darstellung von Graustufen zu.

<u>Plotter</u>	<u>elektrostatischer Drucker</u>
beinahe beliebig große Auflösung (hohe Qualität)	Auflösung durch Rasterpunktdichte beschränkt: ca. 100 - 200 Punkte/Zoll (geringe Qualität)
mehrfarbige Darstellung	nur Schwarz-Weiß
Flächendarstellung durch Linienschraffur	Flächendarstellung auch mit Grautönen
Ausgabe alphanumerischer Daten sehr aufwendig	Verwendung als Drucker möglich
großes Zeichenformat	großes Format nur durch sehr teure großformatige Geräte oder bahnenweise Darstellung
Verwendung einzelner Blätter und Gravieren möglich	Derzeit nur mit speziellen Papierrollen verwendbar.

Tabelle 1.3: Unterschiede zwischen x,y-Zeichenautomaten und elektrostatischen Druckern.

(d) Filmschreiber

Aus Rasterdaten kann man mit Hilfe spezieller Geräte (z.B. OPTRONIX) auch Filme - sowohl Farb- als auch Schwarz-Weißfilme - belichten, wobei die Belichtungsstufen im Raster vorgegeben sein müssen. Dadurch erhält man eine für alle photographischen Vervielfältigungsformen gut geeignete Vorlage, die beliebig oft reproduzierbar ist. Die Qualität der Ausgabe ist hier die höchste im Vergleich zu allen anderen Hardcopyverfahren, aber auch die Kosten und der nötige Zeitaufwand sind maximal.

1.9.2 Softcopysichtgeräte

(a) Rasterbildschirme

Die naheliegendste Möglichkeit, graphische Daten am Bildschirm auszugeben, ist die Erweiterung vorhandener Datensichtgeräte durch Graphikmodule. Sie erlauben es, auf den wie ein binäres Schwarz-Weiß TV-Gerät (z.B. 256 x 256 Punkte) konzipierten Terminals einzelne Punkte des Bildschirms direkt anzusprechen.

Es kommen zur Zeit viele neue, schon für die Graphik konzipierte Rasterbildschirme auf den Markt, welche hohes Auflösungsvermögen haben (z.B. 1024 x 1024 und 4000 x 4000 Punkte!). Damit wird die Darstellung qualitativ wesentlich hochwertiger. Meist ist nur 1 Bit je Rasterpunkt vorgegeben, (Schwarz-Weiß), aber es nimmt auch das Angebot von Farbsystemen (4,6 oder 8 Bit je Rasterpunkt) zu.

(b) Vektorbildschirme

Die typischen Graphik-Bildschirmgeräte sind die Vektorbildschirme einerseits mit Speicherröhren (cathode ray storage tubes, z.B. TEKTRONIX) und andererseits mit Bildwiederholpeicher (engl. "refresh displays", z.B. GRAPHIC7 von SANDERS). Solche Geräte werden auch "kalligraphisch" genannt.

Die Darstellungsmöglichkeiten von allen diesen Geräten sind ähnlich denen des Plotters und auch die Schnittstelle zwischen Gerät und Benutzerprogramm besteht i.a. aus der gleichen standardisierten Software; bei einigen Geräten sind ebenfalls mehrfarbige Zeichnungen möglich. An diese Geräte können auch direkt Zeichengeräte angeschlossen sein, die die auf dem Bildschirm angezeigten Informationen (Softcopy) auf Papier wiedergeben (Hardcopy).

Unterschiede zwischen den zwei Arten der Vektorbildschirme sind in Tab. 1.4 angeführt.

Speicherbildschirmgeräte	Geräte mit Bildwiederhol- speicher
statische Bilddarstellung (einmal gezeichnete Vektoren können allein nicht verändert werden; Löschen des gesamten Bildschirms erforderlich)	dynamische Bilddarstellung (Änderung einzelner Vektoren bzw. Bildteile durch Änderung der Daten im Bildwiederhol- speicher möglich; Bildwieder- holung mit Videorate, d.h. 50 Mal pro Sekunde).
schlechte Editiermöglich- keiten	Editieren im Dialog gut un- terstützt (besonders einfach mit Lichtgriffel)
für komplizierte Zeich- nungen relativ langsam	Schnelligkeit unabhängig von der Komplexität der Zeichnung

Tab. 1.4: Unterschiede zwischen Speicherröhren und Bildwiederholungspeichern bei kalligraphischen Bildschirmgeräten.

(c) Andere Ausgabemöglichkeiten

Zur Darstellung von Rasterdaten gibt es Bildverarbeitungssysteme (z.B. DE ANZA), die neben einer mehrfärbigen Darstellung auch einfache Bildauswertungsoperationen (z.B. Überlagerung von Teilbildern, Kontrastverstärkung, usw.) ermöglichen.

Bildverarbeitungssysteme unterscheiden sich von gra-

phischen Rasterbildschirmen im wesentlichen durch den Zweck: Graphische Rasterbildschirme, welche unter (a) angeführt wurden, haben 1 Bit pro Pixel, um Linien darstellen zu können, oder einige Bit (4 bis 8), um Linien und Flächen färbig darzustellen. Bildverarbeitungssysteme haben zumindest 24 Bit pro Pixel, um 3 Grautonbilder simultan wiederzugeben, nämlich eines in Rot, eines in Grün, eines in Blau, und somit Echtfarbbilder darzustellen. Eine Bildverarbeitungsanlage kann daher immer auch für Vektordaten (Linien usw.) verwendet werden, nicht aber umgekehrt: Graphikrasterschirme sind für Bildverarbeitung ungeeignet.

Bei Vorhandensein eines Videoausgangs an der Datenverarbeitungsanlage kann man handelsübliche Farbfernsehgeräte anschließen und ähnlich wie bei Datensichtgeräten mit Graphikmodul Informationen darstellen. (CHIP, Jänner 1980, Seite 53-57).

1.9.3 Diskussion

Operiert ein GIS mit Vektordaten, so ist als Ausgabegerät zumindest ein elektrostatischer Drucker notwendig, um eine hinreichend gute Darstellung der graphischen Daten zu gewährleisten. Soll das GIS hauptsächlich für statistische Zwecke und zur Ausgabe tabellarischer Daten verwendet werden, kann mit einem herkömmlichen Drucker schon Wesentliches erreicht werden.

Die Datenerfassung und die Arbeit mit den gespeicherten Daten erfordert jedoch die schnelle Sichtbarmachung der erfaßten und möglicherweise fehlerhaften Daten und der daraus abgeleiteten Ergebnisse, bevor eine Hartkopie aufgezeichnet wird. Somit sind Bildschirmsichtgeräte für die Arbeit mit dem GIS unerlässlich.

TEIL 2

Überblick über das zu erstellende System DESBOD =====

2.1 Einführung

DESBOD steht für "digitale Erfassung, Speicherung und Bearbeitung ortsbezogener Daten". Daten mit Ortsbezug bestehen heute in einer Vielzahl von fachspezifischen Einrichtungen des öffentlichen und privaten Sektors. DESBOD richtet sich auf Anwendungen in Umweltfragen und den Geowissenschaften. Die Vielfalt in Form und Art der Daten erschwert ihre Verwendung in der Planung der Umwelt und Nutzung der natürlichen Hilfsquellen. In der Entwicklung von neuartigen Informationssystemen liegt daher ein Schlüssel für die Beseitigung dieser Hindernisse.

Die Nutzung digitaler Verfahren ermöglicht die wirkungsvolle Bearbeitung größter Datenmengen und erlaubt eine Objektivierung des Planungsvorganges. Diese Möglichkeiten erweitern über neue Planungsmethoden die Transparenz von Planungsvorgängen. Somit betrifft solch ein Vorhaben neben der Vielzahl von Geowissenschaften als Datenzulieferer und der Computerwissenschaft als Durchführungsgrundlage die sozialwissenschaftliche Umweltplanung und hat Einfluß auf eine Reihe von durch die Planung betroffenen Sektoren der Gesellschaft, Wirtschaft und Wissenschaften.

In dieser Beschreibung wird versucht, ein maximales System in bezug auf die derzeit verfügbare Hardware zu entwerfen und dabei die spätere Integration von Hardware-Erweiterungen vorzusehen. Bei der Anwendung des Systems wird tatsächlich immer nur eine Teilmenge der vorhandenen Möglichkeiten verwendet werden. Was die Software betrifft, so

beschreiben wir ein Programmsystem, das sich von einer Grundausbaustufe bis zu einem alle wesentlichen Teile der heutigen GIS-Technologie umfassenden Gesamtsystem entwickeln soll.

Im Gegensatz zu anderen Informationssystemen, die für einen bestimmten Anwendungsbereich geschaffen wurden (wie z.B. Bodennutzung), wird hier versucht, ein universell verwendbares System zu entwickeln. Dabei werden auch bereits bestehende Systeme und die daraus gewonnenen Erfahrungen berücksichtigt. Das ist umsomehr notwendig, weil diese Ergebnisse und Erfahrungen durch jahrelange Forschung und mit enormen finanziellen Aufwand erreicht wurden; wegen der beschränkten Mittel ist es auch gar nicht möglich, eine eigenständige Entwicklung von Grund auf durchzuführen, vielmehr soll den Teilen von DESBOD, die für eine kontinuierliche Weiterentwicklung des Systems (z.B. Datenbank) besonders wichtig sind, besondere Aufmerksamkeit geschenkt werden.

2.2 Hardware

Als Rechenanlagen stehen uns eine PDP 11/34 mit 256 K Byte und eine VAX 11/750 mit 1 M Byte zur Verfügung (siehe Fig. 2.1). Es ist geplant, DESBOD an der VAX zu installieren. Als Massenspeicher verwenden wir eine 127 M Byte Festkopffplatte und ein 9-Spur Magnetbandgerät. Alle diese Geräte sind von Digital Equipment Corporation.

Für die Datenerfassung benutzen wir derzeit einen manuellen Digitalisiertisch der Firma ALTEK, Bildschirmterminals und in einer späteren Ausbaustufe einen analytischen Plotter von KERN. Weiters ist die Verwendung von Daten eines kartographischen Rasterscanners vorgesehen (Karto-

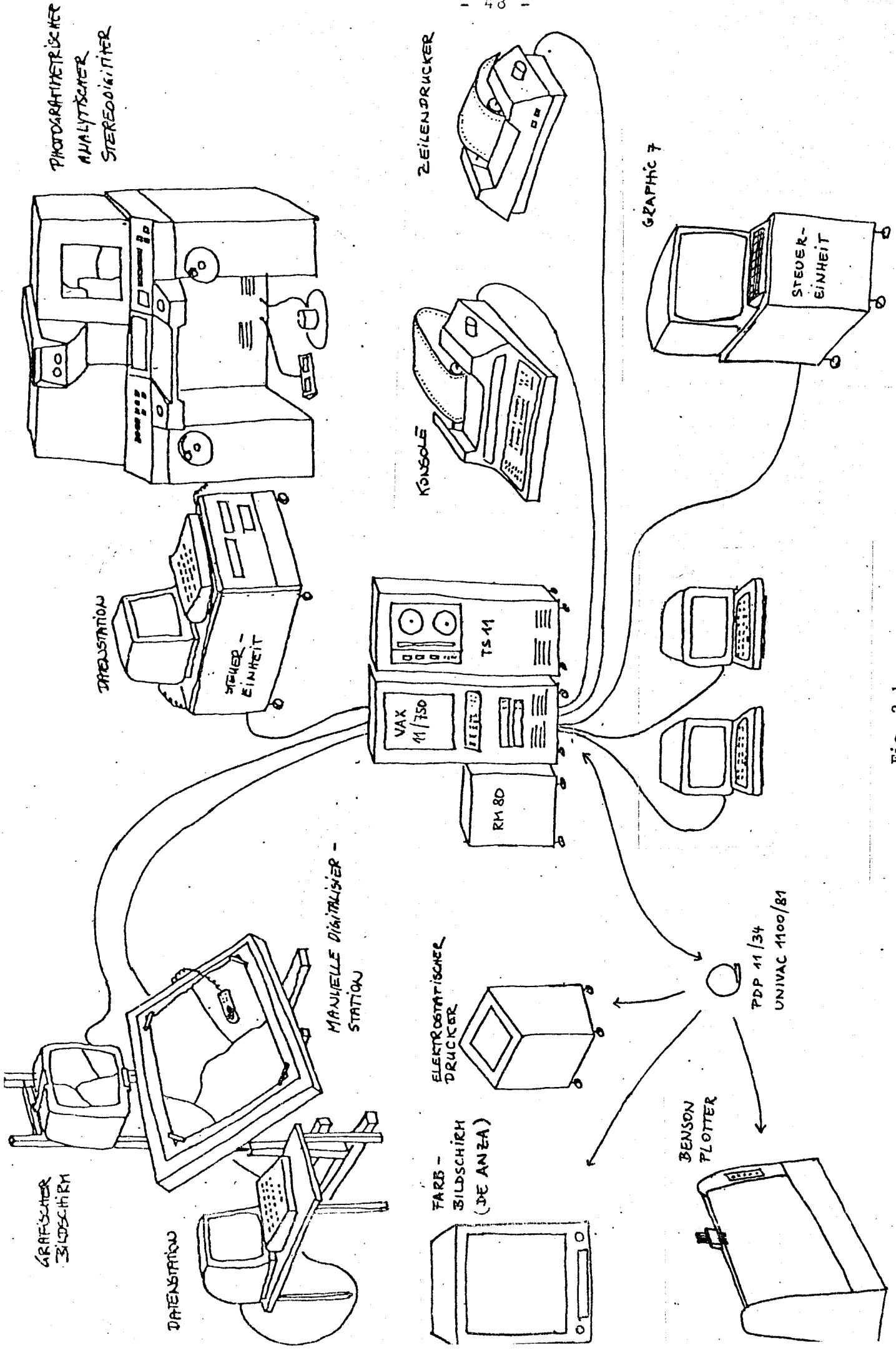


Fig. 2.1

Hardwarekomponenten von DESBOD

scan der Firma MBB). Die Ausgabe der Daten erfolgt auf graphischen Bildschirmen (TEKTRONIX 4014, GRAPHIC 7 von Sanders, DE ANZA Rasterfarbbildcomputer), Zeilendruckern und einem elektrostatischen Printer/Plotter von VERSATEC. Weiters besteht die Möglichkeit, der Ausgabe auf einem x,y-Zeichenautomaten (derzeit noch über Magnetbänder am BENSON Plotter der UNIVAC 1100/81 des Rechenzentrum Graz).

Der Digitalisiertisch, ein graphischer Bildschirm und ein Kontrollterminal sind zu einem graphischen Arbeitsplatz kombiniert. Ein graphischer Bildschirm mit Wiederhol-speicher (GRAPHIC 7) wird insbesondere für das interaktive Editieren der digitalisierten Karten verwendet, wobei diese manuell oder über den Rasterscanner erfaßt worden sein können. Mit Hilfe des analytischen Plotters (von einer PDP 11/03 gesteuert) ist es möglich, aus zwei Luft-bildern dreidimensionale Koordinaten des Geländes zu er-halten.

2.3 Gliederung von DESBOD

(a) Allgemein

Das Programmsystem DESBOD gliedert sich im wesentlichen in drei Teile, nämlich (i) ein Datenerfassungssystem, (ii) ein Datenbanksystem und (iii) ein Analyse- und Ausgabe-system. Jedes dieser Systeme ist wiederum in Teilsysteme unterteilt. Die drei Hauptsysteme werden von einer überge-ordneten Kommandoprozedur gesteuert (Fig. 2.2).

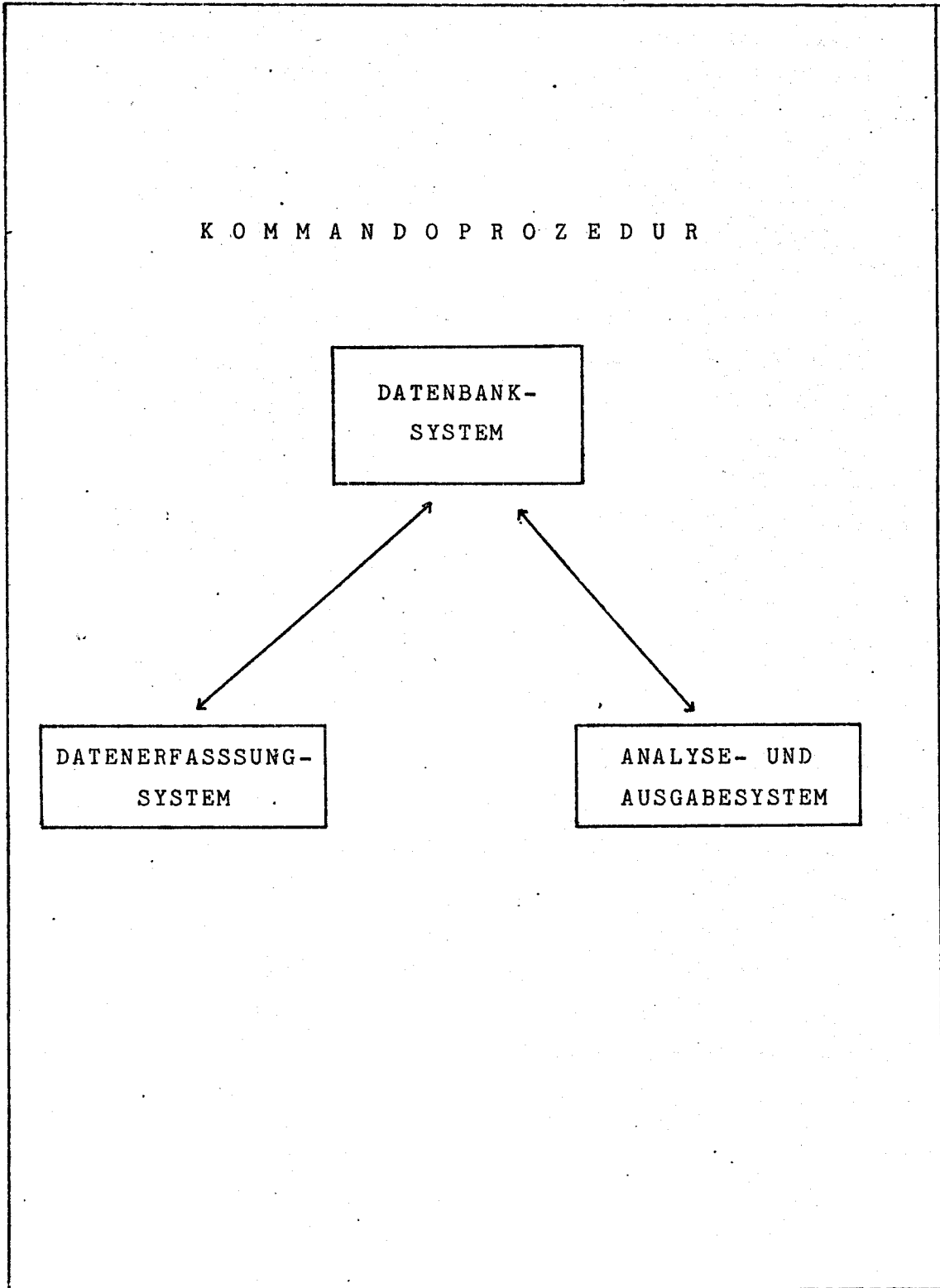


Fig. 2.2

(b) Zur Datenerfassung

Der Fluß der Daten in DESBOD wird in Figur 2.3 veranschaulicht. Für die Eingabe der Grundrißelemente sind manuelle und automatische Verfahren vorgesehen. Die manuelle Erfassung (Digitalisieren) erfolgt on-line über eine manuelle Digitalisierstation. Isolinien (wie z.B. Höhen-schichtenlinien) und dichte Polygonvorlagen können außer Haus mittels eines Scanners automatisch erfaßt und die so gewonnenen Rohdaten über Nachverarbeitungsschritte in das System eingebracht werden. Dieses Verfahren wird gewählt, wenn es sich bei den Daten um für den Scanner leicht erkennbare Strukturen handelt und außerdem die manuelle Erfassung einen größeren Aufwand an Zeit und Mitteln erfordern würde als die Abtastmethode. Alle erfaßten Daten werden dann einer interaktiven Prüfung unterzogen, bevor sie in die DESBOD-Datenbank aufgenommen werden können.

Die Themalelemente (Themabeschreibungen), die sowohl in numerischer als auch alphanumerischer Form vorliegen, werden auf der alphanumerischen Tastatur eines Bildschirmgerätes eingegeben. Diese Zeichenketten (Kartennamen, Legenden, Höhenangaben, etc.) werden in ein DESBOD-kompatibles Textformat gebracht und können dann in die Datenbank aufgenommen werden.

Luft- oder Satellitenbilder, welche im Digitalformat auf Magnetbändern gespeichert sind, stehen für die Verwendung in DESBOD zur Verfügung. Sie müssen jedoch mit einem unabhängig von DESBOD bestehenden Bildverarbeitungssystem DIBAG verarbeitet, analysiert und in ein geeignetes Format gebracht werden, bevor sie in die DESBOD-Datenbank aufgenommen sind.

Der letzte Schritt vor der Aufnahme von Daten in die Datenbank ist das Inbeziehungsetzen der graphischen Daten (Karten, Bilder) mit den Textdaten (Namen, Legenden, Attri-

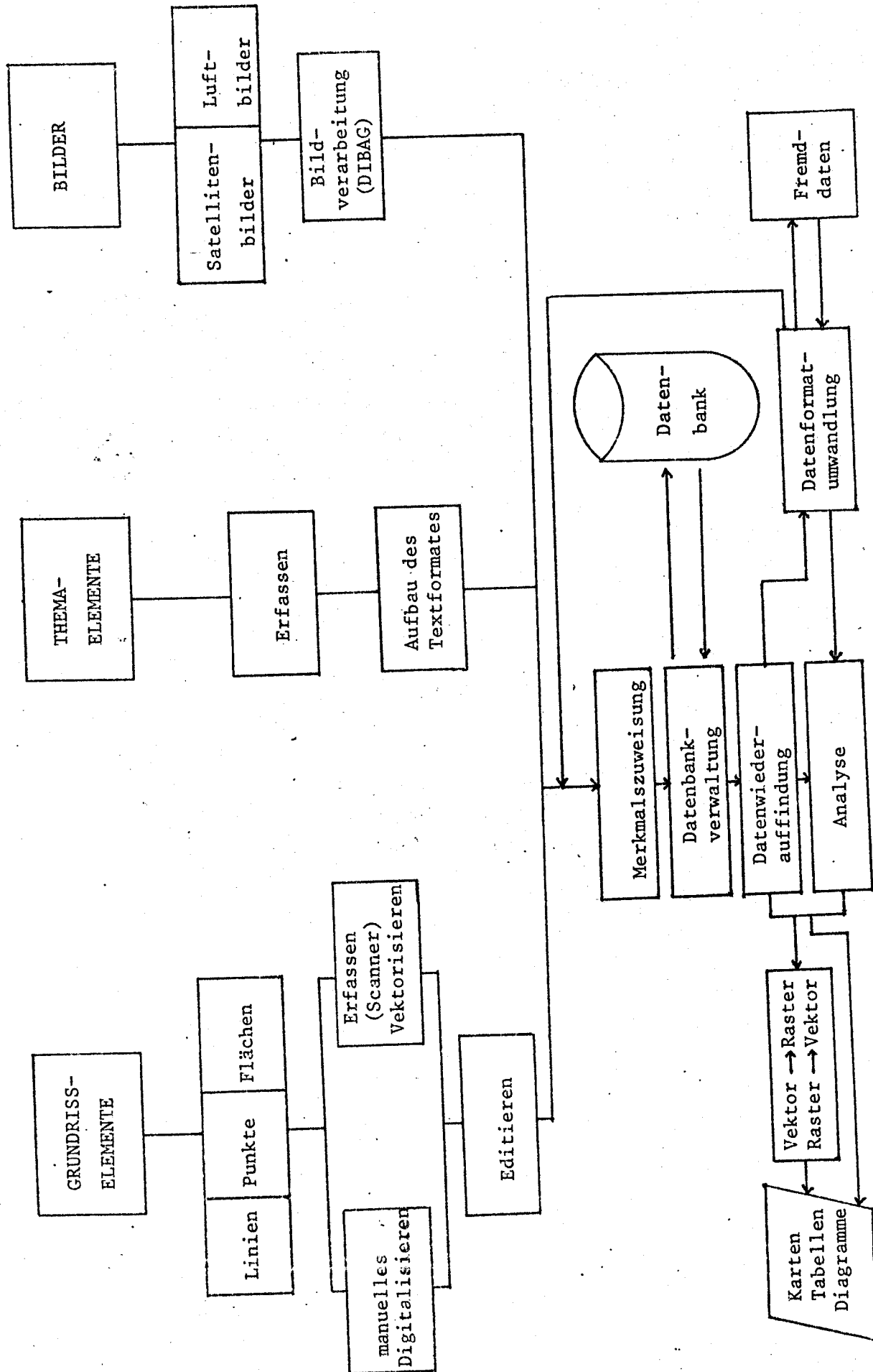


Fig. 2.3
Datenfluß in DESBOD

buten). Dieser Schritt wird mit "Merkmalszuweisung" bezeichnet und erfolgt interaktiv unter Zuhilfenahme graphischer Bildschirme. Der Vorgang kann, obwohl vom Menschen gesteuert, von der Maschine wesentlich unterstützt werden.

(c) Zur Datenbank

Die Datenbank ist so konzipiert, daß sie in ihrer Struktur alle wesentlichen logischen und topologischen Relationen widerspiegelt, die zwischen den Elementen einer Karte existieren. Aus Sicherheitsgründen ist es bei der Datenerfassung und -auswertung nicht möglich, direkt auf die Datenbank zuzugreifen. Dem Benutzer steht ein Arbeitsbereich für solche Manipulationen zur Verfügung. Über eine Schnittstelle ist eine Veränderung (Speichern, Löschen) der Datenbank möglich, allerdings ist dazu eine besondere Berechtigung erforderlich.

(d) Zur Auswertung

Für die Auswertung der Daten werden die interessierenden Teile der Datenbank vom Datenwiederauffindungssystem auf einen Arbeitsbereich gebracht und stehen dort für weitere Manipulationen bereit. Die Kartendaten werden entweder direkt, d.h. ohne Analyse, oder nach einer Analysephase auf den verschiedenen Geräten ausgegeben. Dabei ist es möglich, die graphischen Daten von Raster- in Vektorform oder umgekehrt von Vektor- in Rasterform umzuwandeln. Numerische und alphanumerische Daten werden für statistische Auswertungen verwendet. Diese können in Form von Tabellen, Histogrammen oder Listen ausgegeben werden.

Um Daten, die nicht von DESBOD stammen oder die in andere Systeme eingebracht werden sollen, bearbeiten bzw. bereitstellen zu können, ist eine Schnittstelle für Formatumwandlungen von und in Fremddatenformate vorgesehen.

2.4 Datenerfassungssystem

2.4.1 Gesamtablauf

Das Datenerfassungssystem gewährleistet das Erfassen und Editieren der graphischen Kartendaten, der Textdaten und Bilder (Fig. 2.4). Diesen Daten entsprechen drei Datentypen: Grundrißelemente (Punkte, Linien, Flächen), Themaelemente und Bilder. Themakarten (thematische Karten) bestehen aus Grundrißelementen und Themaelementen (z.B. Kartennamen, Legenden). Themaelemente können unabhängig von den Grundrißelementen eingegeben und verarbeitet werden. Bilder werden wie Rasterdaten behandelt.

Es wird im vorliegenden Abschnitt nur die manuelle Datenerfassung diskutiert. Sollten Rasterscanner zu Verfügung stehen, so ist eine getrennte Schnittstelle zur Einspeicherung in DESBOD zu entwickeln.

Grundsätzlich sind aber für jede Themakartenerfassung bzw. Themakartenerstellung eines bestimmten Untersuchungsgebietes folgende Informationen notwendig:

- Die genaue Begrenzung des Untersuchungsgebietes (der gesamte Raum, der von Interesse ist).
- Die verschiedenen Grundrißelemente innerhalb des Untersuchungsraums, die für den Untersuchungsgegenstand in Frage kommen.
- Die Werte (oder Attribute) die für den Untersuchungsgegenstand in den einzelnen Grundrißelementen in Frage kommen.

Sind diese Informationen vorhanden, so muß die Eingabe einer Themakarte folgender Ablauf eingehalten werden:

1. Für jedes Thema muß - bevor die Grundrißelemente der Karte eingegeben werden - eine Themabeschreibung er-

stellt werden, die die Bezeichnung des Themas und einen Themacode enthält und für jeden Wert bzw. für jedes Attribut des Untersuchungsgegenstandes den notwendigen Text und einen eindeutigen Klassifizierungscode beinhaltet.

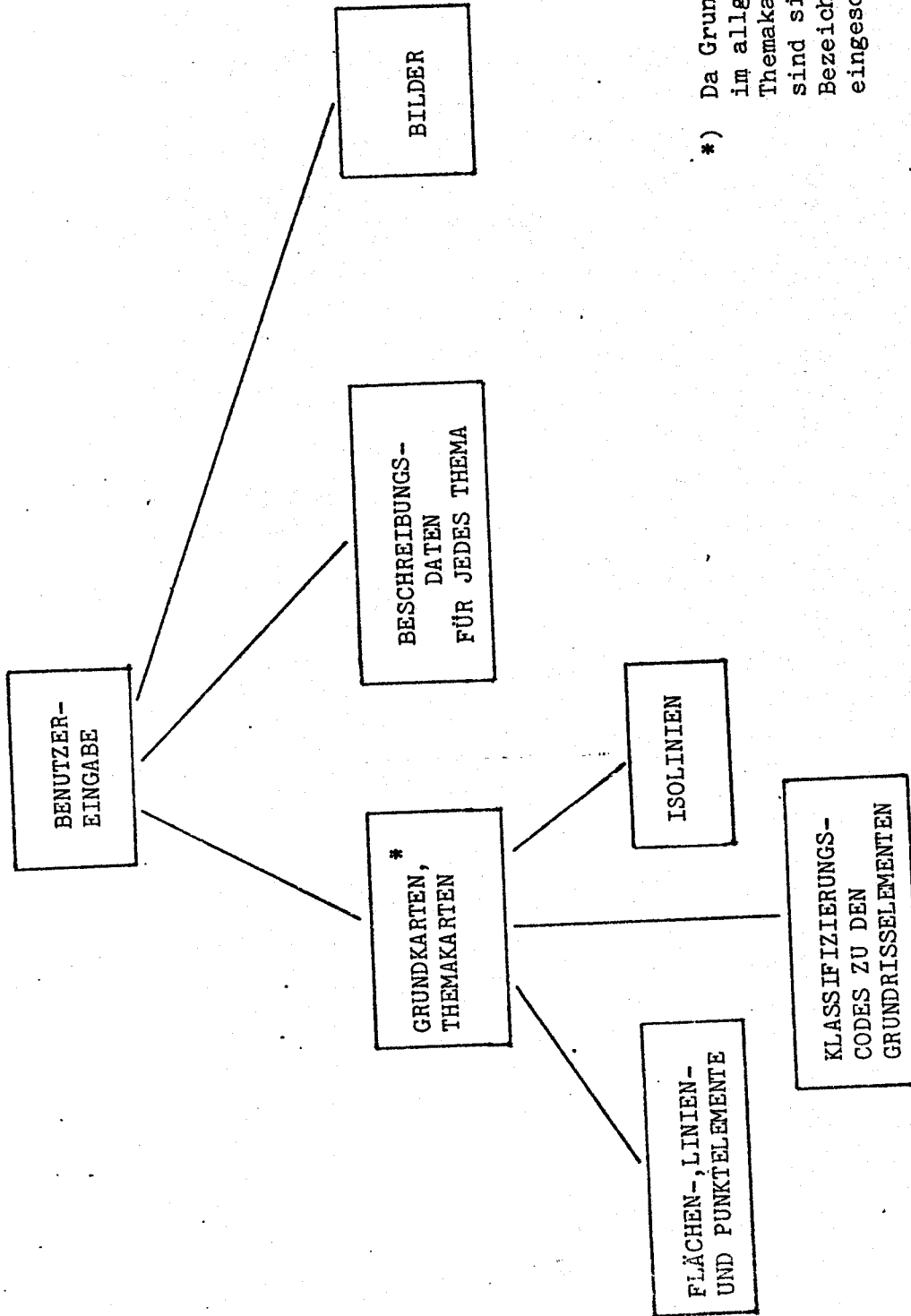
2. Der Benutzer kann das Studiengebiet in Einheiten unterteilen, die er dann getrennt behandelt. Ist das Studiengebiet z.B. die Steiermark, so kann man dieses Gebiet durch mehrere Blätter der ÖK 50 oder ÖK 25 überdecken. Jedes dieser Kartenblätter repräsentiert dann eine Einheit. Der Eingabemaßstab ist unwesentlich für die Datenerfassung; es muß dem Benutzer allerdings klar sein, daß durch die Wahl eines kleineren Maßstabes bei der Ausgabe als bei der Erfassung nur eine Vergrößerung, nicht aber auch eine Zunahme an Genauigkeit erreicht werden kann. 2
3. Durch die Eingabe der Koordinaten des linken unteren und des rechten oberen Eckpunktes der gewählten Einheit (in geographischer Länge und Breite) und eines Themacodes ist das zu digitalisierende Gebiet eindeutig bestimmt und es kann - nachdem das Kartenblatt auf dem Digitalisierertisch eingemessen wurde - mit dem Digitalisieren der Grundrißelemente begonnen werden. Der genaue Ablauf des Digitalisierens wird in Abschnitt 2.4.3 erläutert. Die erhaltenen Koordinaten werden als geographische Länge und Breite abgespeichert.
4. Beim manuellen Digitalisieren werden auch die Klassifizierungs-codes für die einzelnen Grundrißelemente eingegeben.
5. Durch die Verwendung von Editier- und Zeichenroutinen ist es dem Digitalisierenden jederzeit möglich, die bereits eingegebenen Grundrißelemente auf dem graphi-

schen Ausgabegerät auszuzeichnen und die dadurch erkannten Fehler zu beseitigen.

6. Nach Abschluß der Digitalisierung eines Gebietes wird ein interaktives Korrekturprogramm gestartet, das logische und topologische Fehler erkennt und diese - soweit es möglich ist - automatisch beseitigt. Die dann noch bestehenden Fehler werden im Dialog mit dem Benutzer korrigiert.

7. Nachdem noch die Kanten der Einheit an die Kanten der schon in der Datenbank bestehenden, angrenzenden Einheiten desselben Themas angeglichen wurden, wird die neue Themakarte mit Hilfe des Datenbank-Verwaltungssystems in der Datenbank abgelegt.

Im folgenden werden die sieben Arbeitsschritte einzeln beschrieben.



*) Da Grundkarten in diesem System im allgemeinen gleich wie Themakarten behandelt werden, sind sie in Hinkunft bei der Bezeichnung "Themakarte" mit eingeschlossen.

Fig. 2.4
Aufschlüsselung der Eingabedaten

2.4.2 Erfassung der Themabeschreibungen

Bevor die Grundrißelemente eines Studiengebietes erfaßt werden können, muß eine Themabeschreibung eingegeben werden. Diese Themabeschreibung besteht aus numerischen, alphanumerischen und symbolischen Themacodes. Jedoch entspricht jedem symbolischen Themacode ein numerischer oder ein alphanumerischer Code. Jedem alphanumerischen Code ist ein numerischer Referenzcode zugeordnet. Symbolische Codes werden über ein sogenanntes Softwarekeyboard (siehe Abschnitt 1.4) eingegeben. Jedes Thema hat einen alphanumerischen Themanamen, dem ein numerischer Themacode entspricht, welcher als Referenz für die Eingabe der dazugehörigen Grundrißelemente, Zusatzbeschreibungen und für die Datenausgabe und -analyse dient. Es ist möglich, zu denselben Grundrißelementen mehrere Themen einzugeben. Diese sogenannten Zusatzbeschreibungen werden an die übergeordnete Themabeschreibung angehängt. Ein Beispiel dafür wäre die Einwohnerzahl der steirischen Bezirke als Zusatzthema zu der Themakarte der politischen Grenzen und Namen der steirischen Bezirke. Fig. 2.5 zeigt das Konzept der Themabeschreibungserfassung.

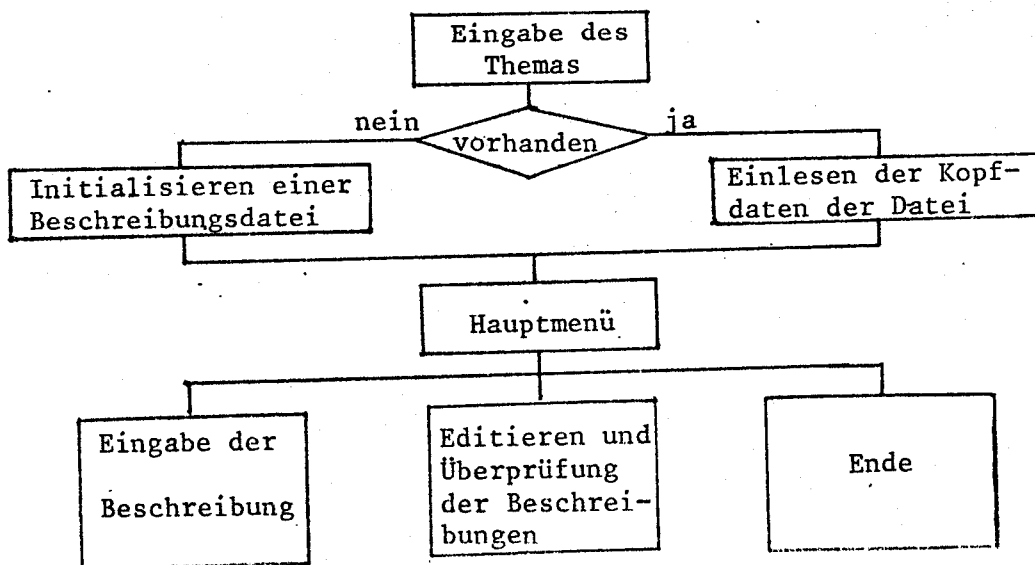


Fig. 2.5

Ist das Thema noch nicht beschrieben, muß eine entsprechende Beschreibungsdatei initialisiert werden. Dabei werden der Themacode und der Themaname festgelegt und die Themaelemente mit den dazugehörigen Codes definiert, d.h. es wird die Kartenlegende eingegeben. Kommen in der Legende Symbole vor (z.B. das Symbol . . . für eine archäologische Fundstätte), so können sie in dieser Programmphase über das Digitalisiergerät eingegeben werden. Dazu muß sie der Benutzer auf ein Blatt Papier zeichnen, welches dann als Teil des Softwarekeyboards verwendet wird. Dabei müssen die Konturen der Symbole mit dem Cursor nachgefahren und ein Themacode für das digitalisierte Symbol eingetippt werden. Später, beim Digitalisieren der Grundrißelemente, wird einem Objekt ein bestimmtes Symbol zugewiesen, indem man mit dem Cursor zu dem Symbol am Softwarekeyboard hinfährt und einen Punkt im Bereich des Symbols abnimmt. Fig. 2.6 zeigt Beispiele für die Themabeschreibungen.

Themaname:	steirische Bezirke	Code:	003
Themaelement:	Radkersburg	Code:	615

Fig. 2.6 a
alphanumerische Beschreibungen und Themacodes

Themaname:	Höhenschichtenlinien	Code:	010
Themaelement:	400	Code:	400

Fig. 2.6 b
numerische Beschreibung eines Themaelementes









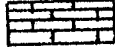
 Bohrloch	 Baum	 Ruine
 archäolog. Fundstätte	 Kirche	 Haus
 Bildstock Kapelle	 Meßpunkt	 Mauer
Menüfeld zum Thema		

Fig. 2.6 c
Softwarekeyboard

Numerische und alphanumerische Themabeschreibungen und die entsprechenden Themacodes werden anschließend über die alphanumerische Tastatur eines Bildschirmterminals eingegeben.

Alle Themabeschreibungen können geändert werden. Das wird durch ein Editier- und Überprüfungsmodul gewährleistet. Dabei können ganze Themabeschreibungen (eine vollständige Kartenlegende) gelöscht und neu angelegt werden; es lassen sich aber auch nur einzelne Themalemente löschen, ändern oder neu aufnehmen.

Der Vorteil dieses Konzeptes für die Themabeschreibungen liegt darin, daß Themabeschreibungen leicht mit den graphischen Grundrißelementen verknüpft werden können und außerdem lassen sich auch unabhängig von graphischen Elementen reine Themadateien anlegen, d.h. Dateien ohne graphischen Bezug.

2.4.3 Erfassung von Grundrißelementen

(a) Chain-Format

Aufgrund der in Anhang B dargelegten mathematischen Beziehungen zwischen den Elementen einer Karte verwenden wir für das digitale Erfassen der Grundrißelemente und für ihre computergerechte Darstellung ein sogenanntes Chain-Format. (Harvard, 1980; Voxland, 1981). Dabei unterscheidet man drei grundlegende Datenelemente:

- Kanten (Chains, 1-Zellen)
- Knoten (0-Zellen)
- Attribute (2-Zellen)

Fig. 2.7 veranschaulicht diese Elemente und

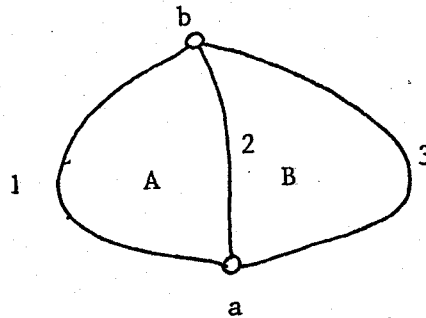


Fig. 2.7

zeigt zwei Flächen, A und B. Diese werden nach außen von den Kanten 1 und 3 begrenzt. Kante 2 ist die gemeinsame Grenze von A und B. Die Knoten a und b kennzeichnen die Punkte, wo die Kanten beginnen (Anfangsknoten) bzw. enden (Endknoten). Durchläuft man die Kante 2 von a nach b, so hat sie ein linkes Attribut A und ein rechtes Attribut B. Für linienhafte Strukturen, wie z.B. Flüsse oder Straßen, wird nur das linke Attribut verwendet. Bei Punktdaten,

z.B. Berggipfel oder Bohrlöchern, fallen Anfangs- und Endpunkte zusammen.

(b) Attribute

Die Kanten, Knoten und Attribute erhalten Nummern. Dabei werden die Nummern, die die Attribute darstellen, schon für Themacodes verwendet. Auf der politischen Karte kennzeichnet z.B. das linke Attribut einer Kante, die eine Gemeindegrenze darstellt, die links von der Grenze gelegene Gemeinde mit ihrer Gemeindenummer und das rechte Attribut die rechts gelegene Gemeinde mit ihrer Gemeindenummer.

(c) Zwischendatei, Verifizierung

Alle manuell digitalisierten Daten werden zuerst im Chain-Format auf Zwischendateien abgelegt. Erst nach der Editier- und Verifikationsphase werden diese Chain-Files abgearbeitet und ihr Inhalt in die DESBOD-Datenbank aufgenommen. Die Verwendung des Chain-Formates bietet sich deshalb an, weil diese Struktur alle Informationen enthält, um entscheiden zu können, ob ein aus mehreren Kanten bestehendes Gebilde ein Gebiet ganz umschließt, und es folglich auch möglich ist, geschlossene Gebiete (Polygone) zu extrahieren bzw. darzustellen.

(d) Punkterfassungsmodus

Beim manuellen Digitalisieren unterscheidet man zwei Digitalisierungsmethoden:

- punktweises Digitalisieren
- kontinuierliches (stream)- Digitalisieren

Beim punktweisen Digitalisieren werden die Punkte, die eine Kante bestimmen, einzeln abgenommen. Beim Zeichnen dieser Kante werden die Punkte durch gerade Linien verbunden. Diese Methode ist zu empfehlen, wenn wenig ge-

krümmte, geradlinige Gebilde zu digitalisieren sind.

Das Stream-Digitalisieren verwendet man für stark gekrümmte, verworrene Strukturen. Bei dieser Methode werden in bestimmten Abstands- oder Zeitintervallen Punkte automatisch abgenommen.

Um von einer Methode zur anderen überwechseln zu können, braucht man nur einen Schalter an einem Kontrollgerät umlegen. Dieses Kontrollgerät ist Teil der Digitalisierhardware. Während des Digitalisierens kann man jederzeit zwischen den zwei Methoden umschalten.

Werden diese zwei Methoden überlegt angewandt, so kann der Datenanfall auf ein Minimum reduziert werden. Eine hohe Punktdichte wird man bei komplexen Gebilden benötigen. In allen anderen Fällen findet man mit dem Punktdigitalisieren das Auslangen.

Die x,y-Koordinaten der Punkte werden in geographische Koordinaten (Länge, Breite) transformiert und gespeichert.

Der Digitalisiervorgang ist in Fig. 2.8 dargestellt.

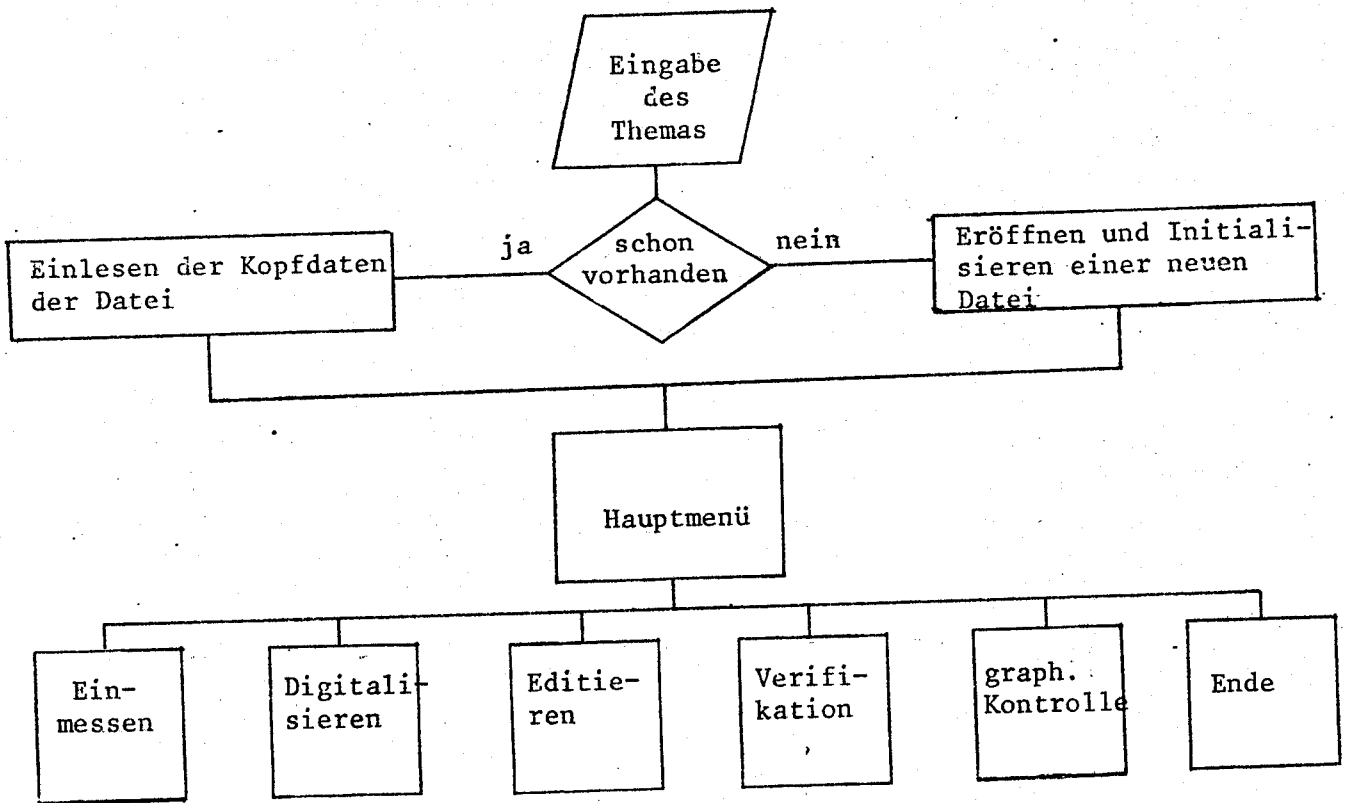


Fig. 2.8

(e) Geometrischer Bezug

Bevor mit dem Digitalisieren begonnen werden kann, muß das Kartenblatt eingemessen werden, d.h. es muß eine Beziehung zwischen den vom Digitalisiertisch gelieferten Tischkoordinaten x, y und den Benutzerkoordinaten λ, φ (Länge, Breite) hergestellt werden. Dies geschieht bei DESBOD zunächst mit einer mehrparametrischen Polynomtransformation über mindestens 2, besser 6, und im allgemeinen N Passpunkte (eine genauere Beschreibung der mathematischen Grundlagen findet man im Anhang C). Mit Hilfe statistischer

Kenngößen entscheidet man über die Güte der Transformation. Ist sie zu schlecht (Meßfehler bei der Abnahme der Passpunkte vom Digitalisiertisch), verwirft man sie und wiederholt den Einmeßvorgang. Das Transformationsverfahren wird so konzipiert, daß es zu einem späteren Zeitpunkt durch andere Methoden ersetzt werden kann oder wahlweise verschiedene Verfahren zur Verfügung stehen.

(f) Kantendigitalisierung

Das Digitalisieren einer Kante geschieht in vier Schritten:

1. Der Digitalisiermodus (punktweises- oder stream-Digitalisieren) wird gewählt.
2. Die Kantenummer, Anfangsknotennummer, Code der linken Region (linkes Attribut) und Code der rechten Region (rechtes Attribut) werden am Cursor des Digitalisiergerätes eingetippt.
3. Die Kante wird vom Anfangsknoten bis zum Endknoten digitalisiert. Dabei kann man falsch abgenommene Punkte wieder löschen. Dieser Vorgang kann auf einem graphischen Bildschirm mitverfolgt werden.
4. Durch Drücken einer Funktionstaste am Cursor des Digitalisiergerätes wird das Abnehmen von Punkten beendet und die Nummer des Endknotens eingegeben.

Dieses Verfahren wird solange wiederholt, bis alle Kanten der gewählten Karteneinheit digitalisiert sind.

Nach jeder Kante kann der Benutzer den Digitalisiervorgang verlassen, um folgende Funktionen durchzuführen:

- einen Bildausschnitt am graphischen Bildschirm zu vergrößern,

- Kanten editieren,
- Starten des Verifizierungsprogrammes,
- noch einmal einmessen,
- das Digitalisiersystem verlassen.

Ist eine Einheit fertig, so müssen die Daten auf ihre topologische Richtigkeit überprüft werden. Dabei werden z.B. die Knotennummern überprüft, ob sie eindeutig sind; es wird geprüft, ob die Koordinaten der End- bzw. Anfangspunkte von zusammenhängenden Kanten übereinstimmen, usw..

2.4.4 Erfassung von Bildern

Bilddaten werden auf Magnetbändern in bestimmten standardisierten Formaten angeboten. Das Bildverarbeitungssystem DIBAG verfügt über Leseroutinen für die verschiedenen Bilddatenformate und wird deshalb für das Erfassen von Bildern verwendet werden. Das Datenbankverwaltungssystem überträgt dann diese Bilder als eine Rasterdatenstruktur in die DESBOD-Datenbank. Der geometrische Bezug zwischen den Vektordaten und Bildern muß durch eine geometrische Transformation von φ, λ und den Rasterzeilen/Spalten hergestellt werden. Hierzu ist entweder:

- in DIBAG eine geometrische Bildentzerrung,
oder
 - in DESBOD eine geometrische Vektortransformation
- durchzuführen.

Andererseits kann DIBAG als Ergebnis einer Bildanalyse ein Rasterbild erzeugen, das nur Kanten enthält. Als Beispiel sei auf das Ergebnis einer Waldklassifizierung verwiesen, welche in einem Ergebnisbild resultiert, das nur die Wald-

grenzen beinhaltet. Diese im Rasterformat vorgegebenen Ergebnisbilder müssen einer Vektorisierung unterworfen werden, um sodann nach einem Editiervorgang in die DESBOD-Datenbank eingespeichert zu werden.

2.5 Die DESBOD-Datenbank

2.5.1 Begriffsdefinitionen

Da sich bis heute in der Datenverarbeitung keine einheitliche Nomenklatur durchgesetzt hat, werden hier einige Begriffe festgelegt, mit denen im folgenden operiert wird.

Erklärung 1:

Ein n -Tupel A von $n \geq 0$ Elementen a_1, a_2, \dots, a_n ist eine Folge von Elementen der Länge n und wird wahlweise als $A = (a_1, a_2, \dots, a_n)$ oder auch nur $A = a_1, a_2, \dots, a_n$ geschrieben.

Erklärung 2:

Das (Kartesische) Produkt zweier Mengen M und N , symbolisch $M \times N$, ist eine durch $M \times N = \{(x, y) \mid x \in M \text{ und } y \in N\}$ definierte Menge von 2-Tupeln oder Paaren. Man nennt jede Teilmenge von $M \times N$ eine Relation (über $M \times N$ oder kurz über M , wenn $M = N$ gilt).

Den Datenbestand in einer Datenbank kann man sich als eine Ansammlung von Datenelementen vorstellen, die als Werte $n \geq 0$ Tupel von Zeichenfolgen besitzen (Typische Werte von Datenelementen sind z.B. Tripel von Zeichenfolgen, die Namen, vorrätige Stückzahlen und Identifikationsnummern von Geräteteilen darstellen).

Zusammenhänge zwischen den Datenelementen lassen sich mit Hilfe von Relationen bequem darstellen. Unter der Datenstruktur der Datenbank soll deshalb hier die Gesamtheit der Datenelemente und der aus ihnen gebildeten Relationen verstanden werden.

Erklärung 3:

Eine Datenstruktur B ist ein Paar $B = (D, R)$. Dabei ist D eine endliche Menge von Datenelementen und R eine endliche Menge von Relationen über D . (Maurer, 1974).

Eine Relation kann in der Datenbank implizit oder explizit enthalten sein. "Explizit" heißt, daß die in einer Relation zusammengefaßten Datenelemente durch Datenbank - Adressen gekennzeichnet und somit im Direktzugriff zugänglich sind. "Implizit" bedeutet, daß die Bildung der Relation indirekt aus den gespeicherten Datenelementen durch Berechnung oder Vergleich abgeleitet werden kann. (Z.B. Erkennung von geometrischen Nachbarschaftsrelationen durch Abstandsberechnung aus den Punktkoordinaten).

Erklärung 4:

Graphische Darstellung einer Datenstruktur: Sei $B = (D, R)$ eine Datenstruktur. Eine graphische Darstellung von B wird folgendermaßen erhalten: Für jedes Datenelement $d \in D$ wird eine Figur gezeichnet: ein Kreis, wenn der Wert von d null ist, anderenfalls ein Rechteck. Zeigt ein Datenelement d auf ein Datenlement d' (bezüglich einer Relation r), dann werden die d und d' entsprechenden Figuren durch eine gerichtete Linie verbunden.

2.5.2 Allgemeine Überlegungen zum Datenbank-Entwurf

Bevor ein Datenbankentwurf angefertigt werden kann, müssen verschiedene Fragen geklärt werden:

- a) Welche verschiedenen Arten von Datenelementen bilden den Datenbestand der Datenbank?
- b) Welche grundsätzlichen, anwendungsunabhängigen Beziehungen bestehen zwischen diesen Datenelementen?
- c) Welche Beziehungen sollen explizit in der DB realisiert werden?
- d) Eng verbunden mit der letzten Frage ist folgende:
Welche Arten von DB-Zugriffen sind zu erwarten?
- e) In welche Einheiten werden die Datenelemente für die Abspeicherung zusammengefaßt?

zu a) Arten von Datenelementen

Zwei große Gruppen von Datenelementen sind in der Datenbank (DB) zu unterscheiden: Grundrißelemente (GE) und Themaelemente (TE).

Während die verschiedenen Grundrißelemente von vorneherein genau definiert werden können (nämlich Flächen-, Linien- und Punktelemente), gilt das für die Themaelemente nicht. Da es möglich sein muß, letztere nach Bedarf in einem relativ freien Format aufzubauen, muß der Benutzer in der Lage sein, Themaelemente nach Bedarf zu definieren. Grundriß- und Themaelemente werden deshalb in der DB verschieden behandelt.

Es wird sich für die Implementierung als praktisch erweisen, bei den Themaelementen noch eine Unterteilung in lokale und globale Elemente vorzunehmen. Als globale Themaelemente sind solche zu verstehen, die allgemeine

Beschreibungen zu einem Thema darstellen und unabhängig von einem bestimmten Kartenblatt bei jeder Behandlung des Themas betrachtet werden (z.B. Vegetationsarten, Gesteinsarten). Im Gegensatz dazu werden lokale Thamaelemente immer in bezug auf ein bestimmtes Kartenblatt existieren und stellen somit Beschreibungen lokaler Gegebenheiten dar (z.B. topographische Namen oder Daten zu Meßpunkten).

zu b) Beziehungen zwischen Datenelementen

Ohne auf spezielle Relationen zwischen den einzelnen Datenelementen einzugehen, können 3 Gruppen von Relationen unterschieden werden:

- Relationen innerhalb der Menge der Grundrißelemente (z.B. Nachbarschaftsrelation: Ordne einem Flächenelement alle benachbarten Flächenelemente zu)
- Relationen innerhalb der Menge der Thamaelemente (z.B. Hierarchische Relation: Ordne der Vegetationsart "Wald" alle möglichen Waldarten zu)
- äußere Relationen zwischen der Menge der Grundrißelemente und der Menge der Thamaelemente (z.B. ordne der Vegetationsart "Wald" alle Flächenelemente zu, die von Wald bedeckt sind; oder ordne jedem Flächenelement die Vegetationsart zu, die es darstellt).

Idealerweise sollen alle Beziehungen, die in einer Karte zwischen den einzelnen Elementen bestehen, in der Datenbank erhalten bleiben, d.h. all die Informationen, die der Nutzer durch die visuelle Interpretation einer Karte gewinnen kann, sollte man auch durch Anfragen an die DB erhalten können. In der Praxis jedoch erweist sich diese Forderung als beschränkt sinnvoll, da der Informationsgehalt eines Kartenblattes die Speicherkapazitäten einer Datenverarbeitungsanlage überschreiten kann.

Es ist daher sinnvoller, die Forderung zu erfüllen,

daß aus der zur Verfügung stehenden Urdatenmenge eine Teilmenge derart auszuwählen ist, daß alle wesentlichen Beziehungen erhalten bleiben bzw. mit minimalen Fehlern rekonstruiert werden können.

zu c) und d): Explizit realisierte DB-Beziehungen,
DB-Zugriffe

Grundsätzlich sind die meisten Relationen zwischen den Datenelementen der Datenbank schon im Datenbestand selbst implizit vorhanden und können aus diesem durch Angabe einer Konstruktionsvorschrift abgeleitet werden. Die Bildung einer Relation auf diese Weise kann allerdings recht zeitaufwendig sein; wenn daher zu erwarten ist, daß eine gewisse Klasse von Anfragen, die dieselbe Relation benutzen, recht häufig ist, sollte diese Relation explizit in der Datenbank gespeichert sein.

Für Grundrißelemente wird in der DESBOD-Datenbank die explizite Darstellung aller topologischen Relationen gewählt, weil diese bei allen *l a g e b e z o g e n e n* Anfragen Verwendung finden und ihre Rekonstruktion besonders komplexe Algorithmen erfordert.

Äußere Beziehungen zwischen den Grundriß- und Themalementen werden in der Regel schon bei der Datenerfassung durch ein geeignetes Numerierungssystem explizit angegeben; diese Relationen bleiben auch in der Datenbank explizit bestehen, um *t h e m a b e z o g e n e* Anfragen effizient bearbeiten zu können.

Wir haben bereits dargestellt, daß Themalemente vom Benutzer frei definiert werden. Es muß ihm daher auch die Entscheidung überlassen bleiben, welche Relation er in der Datenbank explizit festhält. Letzten Endes muß also er abschätzen, welche Anfragen bei der Bearbeitung eines Themas häufig zu erwarten sind.

zu e): Zusammenfassung zu Einheiten

Da in der DB eine sehr große Menge an Datenelementen zu verwalten ist, ist es sinnvoll, diese große Menge in Teilmengen zu zergliedern. Es ist zu überlegen, in welche Einheiten diese Datenelemente für die Abspeicherung zusammengefaßt werden sollen. Diese Einteilung soll einerseits die Verwaltung der Daten erleichtern und andererseits eine effiziente Suche nach bestimmten Datenelementen ermöglichen.

Für die Grundrißelemente wählen wir eine Einteilung nach Kartenblättern, denn:

- Die Erfassung der Daten erfolgt sinnvollerweise nach Kartenblättern.
- Die Menge an Informationen (auf ein Thema bezogen) pro Kartenblatt variiert wenig, es werden also Dateien von ähnlicher Größe aufgebaut.
- Mehrere Kartenblätter lassen sich entlang ihrer Ränder leicht zu größeren Einheiten zusammenfügen; dies setzt voraus, daß sich die einzelnen Blätter nicht überlappen.
- Ein Kartenblatt läßt sich einfach beschreiben, etwa durch die Angabe der Koordinaten der Eckpunkte.
- Die zu einem vorgegebenen Studiengebiet gehörenden Kartenblätter lassen sich unmittelbar bestimmen.

Lokale Themalelemente werden ebenfalls kartenblattweise zusammengefaßt. Globale Themalelemente werden themawise zusammengefaßt.

Für den in der Datenbank vorhandenen Datenbestand ergibt sich also das in Figur 2.9 dargestellte Speicherschema.

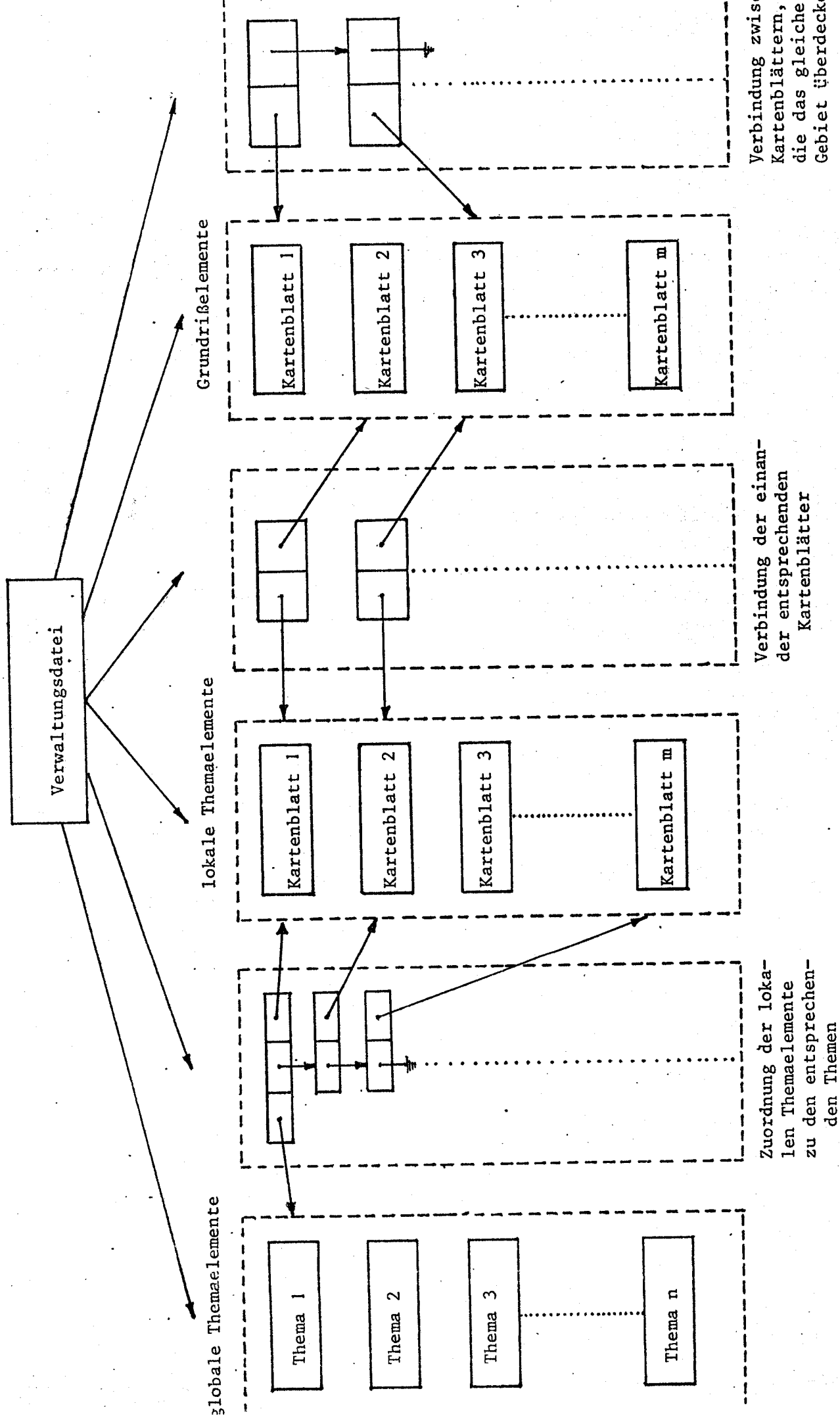


Fig. 2.9
Gliederung des Datenbestandes in der DESBOD-Datenbank

2.5.3 Struktur für Grundrißelemente

Da genau bekannt ist, welche verschiedenen Grundrißelemente es gibt, können auch die zwischen diesen Grundrißelementen bestehenden topologischen Relationen fix festgelegt werden, da sie ja für alle Karten gleich bleiben. In der DESBOD-Datenbank werden daher die inneren Relationen zwischen den Grundrißelementen fix definiert und automatisch aus den Eingabedaten erzeugt; der Benutzer hat daher auf die Festlegung dieser Relationen keinen Einfluß mehr.

Folgende topologische Relationen zwischen den Grundrißelementen existieren:

- 1) Zwischen Flächenelementen $R \subset FE \times FE$ (FE Menge der Flächenelemente)
 - 1 a) Ordnet jedem Flächenelement die benachbarten Flächenelemente zu $(x,y) \in R \iff x$ grenzt an y ;
 $x \in FE, y \in FE$
 - 1 b) Ordnet einem Flächenelement die Flächenelemente zu, aus denen es gebildet wird (hierarchische Flächen-gliederung, z.B. Bundesland - Bezirke)
 $(x,y) \in R \iff x$ beinhaltet y ; $x \in FE, y \in FE$
- 2) Zwischen Flächen- und Linienelementen $R \subset FE \times LE$ (LE Menge der Linienelemente)
ordnet einem Flächenelement die Linienelemente zu, die durch es führen
 $(x,y) \in R \iff y$ führt durch x ; $x \in FE, y \in LE$
- 3) Zwischen Flächen und Punktelementen $R \subset FE \times PE$ (PE Menge der Punktelemente) ordnet einem Flächenelement die Punktelemente zu, die in ihm liegen
 $(x,y) \in R \iff y$ liegt in x ; $x \in FE, y \in PE$

- 4) Zwischen Linienelementen $R \subset LE \times LE$
 ordnet jedem Linienelement die Linienelemente zu, die in es münden $(x,y) \in R \iff y$ mündet in x ; $x \in LE, y \in LE$.
- 5) Zwischen Linien- und Punktelementen $R \subset LE \times PE$
 ordnet jedem Linienelement die Punktelemente zu, die auf ihm liegen
 $(x,y) \in R \iff y$ liegt auf x ; $x \in LE, y \in PE$.

Zu jeder dieser Relationen muß auch noch die jeweilige Umkehrrelation berücksichtigt werden (z.B. Umkehrrelation 3' zu 3 : $R \subset PE \times FE$: ordnet jedem Punktelement das Flächenelement zu, in dem es liegt;
 $(y,x) \in R \iff y$ liegt in x ; $y \in PE, x \in FE$). Die Umkehrrelationen zu 1a) und zu 4) sind mit diesen Relationen ident.

Die beschriebenen Relationen sind in Figur 2.10 graphisch dargestellt.

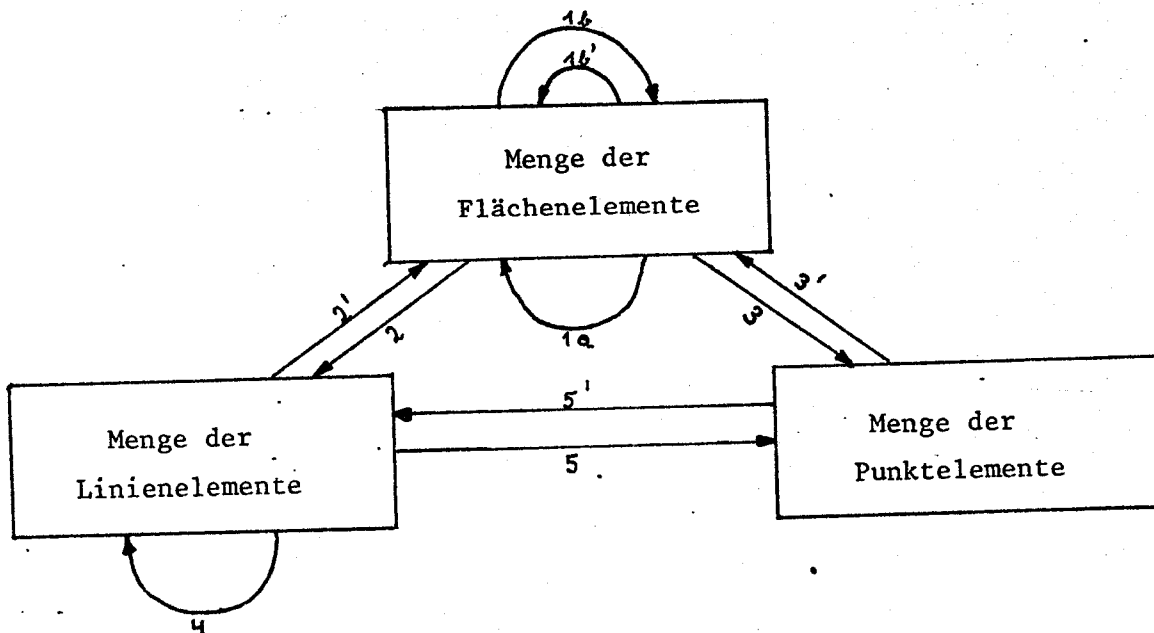


Fig. 2.10

Weitere Relationen - wie z.B. zu jedem Punktelement das ihm am nächsten liegende Punktelement in Beziehung zu setzen - sind nicht mehr zu berücksichtigen, da sie nicht rein lagebezogen sind, sondern noch andere Informationen - wie z.B. die Angabe von Entfernungen - benötigen. Es fragt sich nun, ob alle oben definierten Relationen explizit in der Datenbank vorhanden sein sollen. Die Entscheidung muß davon bestimmt werden, welchen Aufwand es jeweils bedeutet, eine Relation aus dem vorhandenen Datenbestand zu rekonstruieren und wie oft diese Relation bei DB-Zugriffen benötigt wird. Weiters ist die Speicherstruktur der Grundrißelemente selbst - d.h. in welcher Form die Grundrißelemente abgelegt werden - in Betracht zu ziehen.

Ist als Speicherstruktur eine Rasterstruktur vorgesehen, so sind die meisten der obigen Relationen schon explizit vorhanden. Wird hingegen eine Vektorstruktur verwendet, so liegt der Fall anders: "Liegt in" - oder "Liegt neben" - Relationen können durch Vergleich der x,y-Koordinaten der einzelnen Grundrißelemente errechnet werden; dies ist allerdings bei einer großen Menge von Grundrißelementen - von denen ja jedes mit jedem verglichen werden muß - recht zeitaufwendig und mühsam. Es ist daher besser, diese Relationen explizit zugänglich zu machen.

In der DESBOD-Datenbank ist die Speicherung der Grundrißelemente sowohl in Raster- als auch in Vektorstruktur möglich. In welcher Form die Speicherung geschieht, kann vom Benutzer bestimmt werden. Die verschiedenen Kartenblätter eines Themas müssen allerdings in der gleichen Struktur gespeichert sein, weil sonst das Zusammenfügen der Kartenblätter an den Blatträndern nicht durchgeführt werden kann.

Die Implementierung der oben beschriebenen Relationen wird nun für Vektor- und Rasterstruktur getrennt beschrieben.

2.5.3.1 Topologische Relationen für Grundrißelemente in Vektorstruktur

(a) Beschreibung der Relation

Als Ausgangspunkt für unsere Überlegungen nehmen wir die expliziten Relationen, die durch die Kanten-Knotenstruktur geliefert werden:

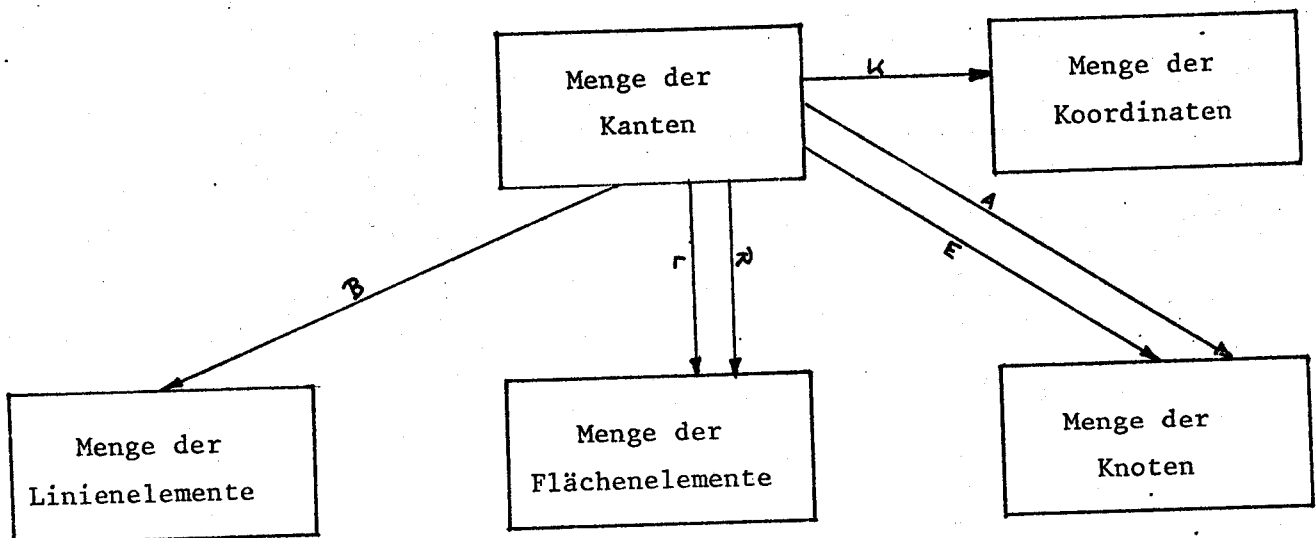


Fig. 2.11

Relationen in der Kanten-Knotenstruktur

Für Fig. 2.11 gilt die folgende Legende:

- K: ordnet jeder Kante die Folge ihrer Koordinatenpaare zu
- A: ordnet jeder Kante ihren Anfangsknoten zu
- E: ordnet jeder Kante ihren Endknoten zu
- L: ordnet einer Kante das links von ihr liegende Flächen-
element zu
- R: ordnet einer Kante das rechts von ihr liegende Flächen-
element zu
- B: ordnet einer Kante das Linienelement zu, zu dem sie ge-
hört.

Bei der Kanten-Knotenstruktur sind allerdings die Punkt-
elemente nicht berücksichtigt, da diese in einer eigenen
Datei eingegeben werden.

Weiters sind Flächen- bzw. Linienelemente nicht explizit
vorhanden, sondern müssen aus dem zugrunde liegenden Da-
tenmaterial aufgebaut werden. Flächenelemente werden durch
die sie begrenzenden Kanten dargestellt, dies stellt die
Umkehrrelation zu den oben angeführten Relationen "L" bzw.
"R" dar. Da jede Kante eine Durchlaufrichtung besitzt, die
Flächenbegrenzung aber einen einheitlichen Durchlaufsinne
haben muß, wird für jede Kante zusätzlich angegeben, ob
sie in oder entgegen der ursprünglichen Richtung durch-
laufen wird.

In analoger Weise wird jedes Linienelement durch die Kan-
ten, aus denen es besteht, charakterisiert.

Unter Berücksichtigung der beschriebenen Relationen ergibt
sich das in Figur 2.12 wiedergegebene Bild:

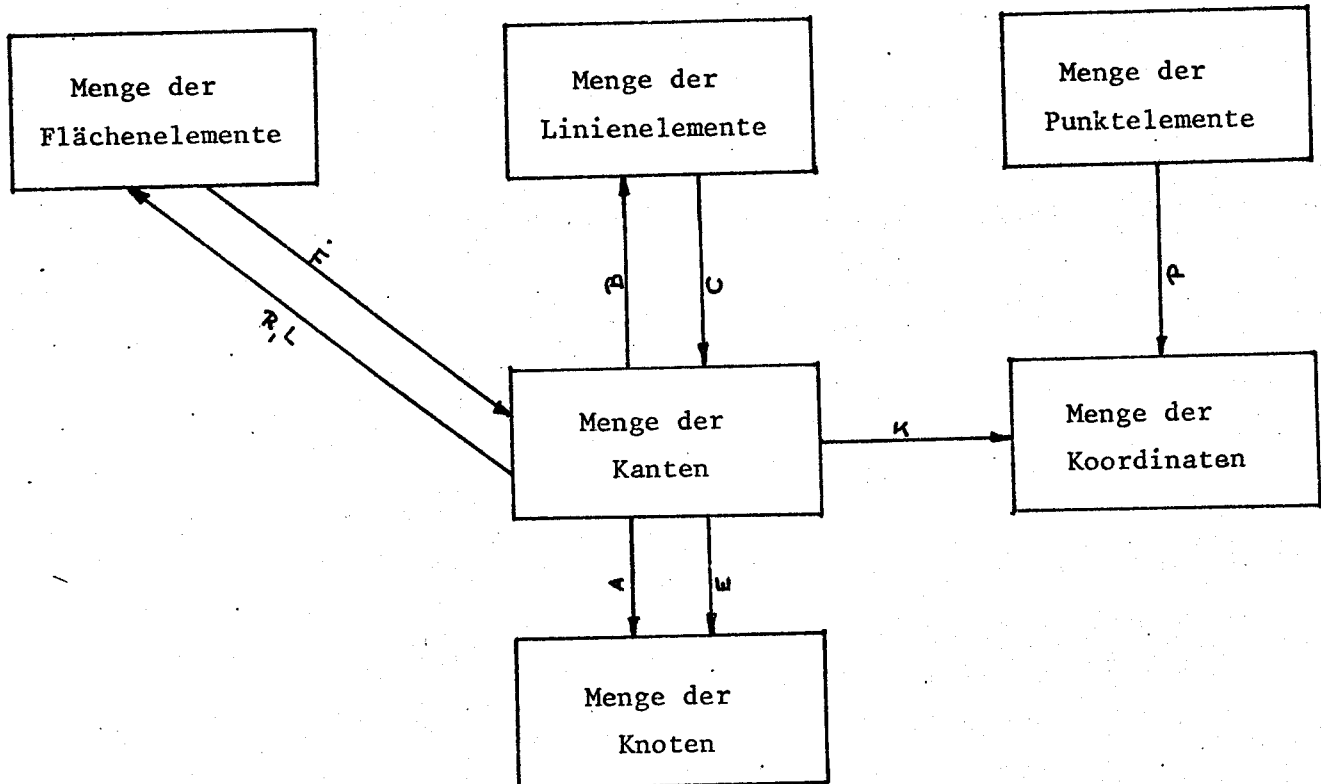


Fig. 2.12

F: ordnet jedem Flächenelement die begrenzenden Kanten zu

C: ordnet jedem Linienelement die Kanten zu, aus denen es besteht

P: ordnet jedem Punktelement seine Koordinaten zu.

Auf diese Weise sind jetzt die Grundrißelemente selbst eindeutig festgelegt.

Die Relationen 1a) - 5) aus Abschnitt 2.5.3 werden folgendermaßen behandelt:

- 1a): Fläche - Fläche Nachbarschaft:

Diese Relation ist bereits explizit vorhanden, da zu jedem Flächenelement die begrenzenden Kanten bekannt und zu jeder Kante die auf der anderen Seite liegenden Flächenelemente gespeichert sind.

- 1b): Fläche aus Teilflächen:

Es wurde bereits ausgeführt, daß bei der Datenaufnahme durch geeignete Wahl eines Numerierungssystems Relationen explizit angegeben werden können. Hierarchische Relationen lassen sich dadurch charakterisieren, daß jeder Hierarchiestufe eine entsprechende Anzahl von Stellen in der Nummer zukommt. (Ein Beispiel dafür ist die Dezimalklassifikation von DEWEY (Knuth, 1973)). Wird dem DESBOD-Datenverwaltungssystem bekanntgegeben, auf welche Weise das Numerierungssystem gebildet wurde, so werden die entsprechenden Hierarchierelationen explizit aufgebaut. Es werden "zusammengesetzte" Flächenelemente gebildet und diese werden zu den "elementaren" Flächenelementen, aus denen sie aufgebaut sind, in Beziehung gebracht.

- 2): Fläche - Linie:

Diese Relation wird vorerst nicht explizit berücksichtigt, sondern muß durch Vergleich der entsprechenden Koordinaten errechnet werden. Da für die Abspeicherung der expliziten Relationen ein flexibles Format gewählt wird, ist es aber ohne weiters möglich, auch diese Relation später einmal explizit einzuführen.

- 3): Fläche - Punkt:

Auch bei dieser Relation wird auf eine explizite Darstellung verzichtet, da solche Beziehungen besser bei den entsprechenden Thamaelementen realisiert werden können.

- 4): Linie - Linie:

Diese Relation wird explizit gemacht, indem zu jedem Knoten die Kanten angegeben werden, die in ihm zusammenlaufen. Alle in ein Linienelement mündende andere Linienelemente sind dann über die Knoten bekannt.

- 5): Linie - Punkt:

Hier gilt das gleiche wie für 3).

Figur 2.13 stellt alle Relationen zwischen den Grundrißelementen dar.

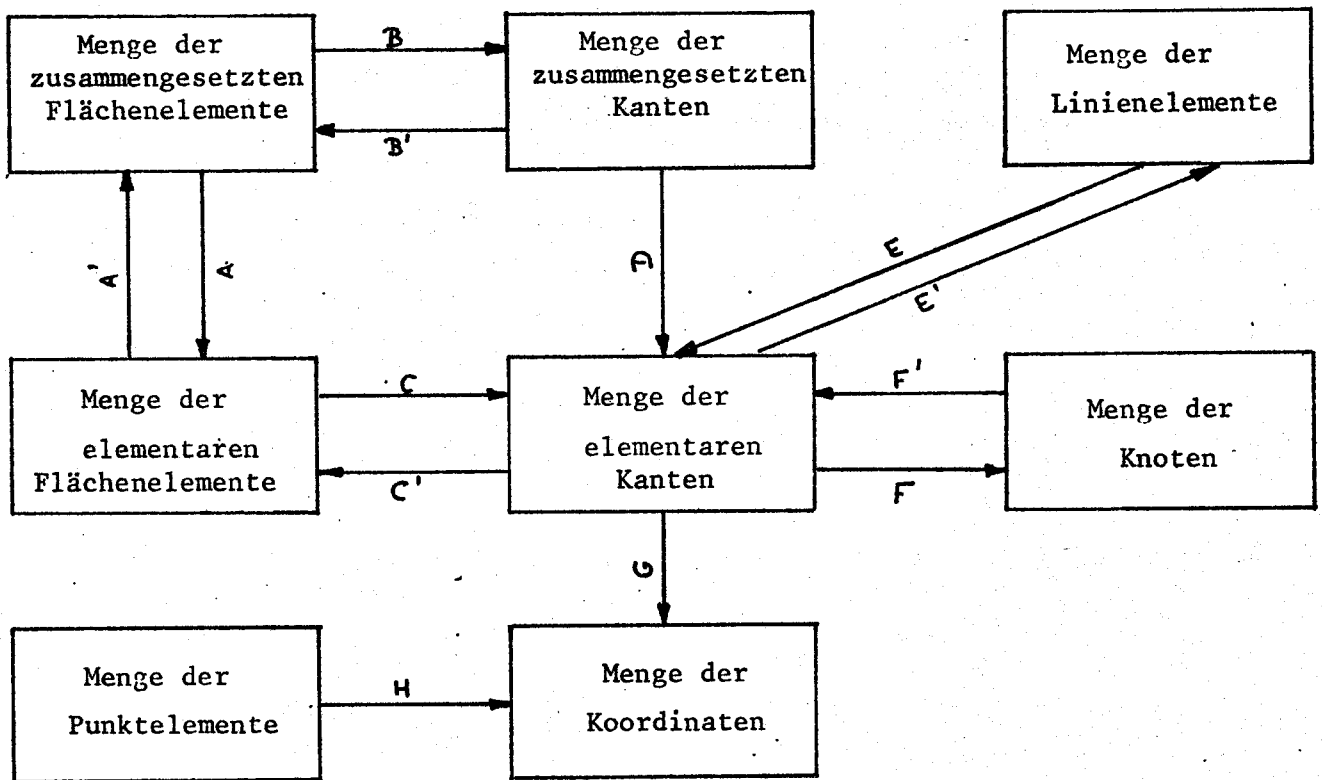


Fig. 2.13

Für Figur 2.13 gilt die folgende Legende:

- A: ordnet einem zusammengesetzten Flächenelement die elementaren Flächenelemente zu, aus denen es besteht
- A': Umkehrrelation zu A: ordnet einem elementaren Flächenelement das übergeordnete zusammengesetzte Flächenelement zu
- B: ordnet einem zusammengesetzten Flächenelement dessen begrenzende Kanten zu
- B': ordnet jeder zusammengesetzten Kante die rechts und links liegenden zusammengesetzten Flächenelemente zu
- C: analog zu B für elementare Flächenelemente
- C': analog zu B' für elementare Flächenelemente / *Kanten*
- D: ordnet jeder zusammengesetzten Kante die elementaren Kanten zu, aus denen sie besteht
- E: ordnet einem Linienelement die Kanten zu, aus denen es gebildet wird
- E': ordnet einer Kante das Linienelement zu, zu dem sie gehört
- F: ordnet einer Kante deren Anfangs- und Endknoten zu
- F': ordnet einem Knoten alle Kanten zu, die in ihm zusammenlaufen
- G: ordnet jeder Kante ihre Koordinatenfolge zu
- H: ordnet jedem Punktelement dessen Koordinaten zu.

(b) Darstellung der expliziten Relationen im Speicher
Die Relationen werden so gespeichert, daß die Dateien der Grundrißelemente dadurch nicht verändert werden, d.h. es hat keinen Einfluß auf die Darstellung eines Datenelementes selbst, ob es an einer Relation beteiligt ist oder nicht.

Die Abspeicherung der expliziten Relationen geschieht in Form von Zeigerdateien:

Die einfachste Zeigerdatei besteht aus Elementen, deren erste Komponente auf ein Datenelement aus der ersten und deren zweite Komponente auf ein Datenelement aus der zweiten an der Relation beteiligten Menge zeigt. Diese Methode ist sinnvoll, wenn einem Element der einen Menge jeweils nur ein Element der anderen zugeordnet wird. Sind einem Element der einen Menge unterschiedlich viele Elemente der anderen zugeordnet, so werden diese Relationen in Form von Listen dargestellt (Maurer, 1974; Knuth 1973). Durch die Invertierung dieser Listen können auch Umkehrrelationen realisiert werden.

2.5.3.2 Topologische Relationen für Grundrißelemente in Rasterstruktur

Im Rahmen von DESBOD wird die Rasterdarstellung für die Speicherung von Funktionswerten auf Pixelmengen verwendet (z.B. digitales Höhenmodell; jede Rasterzelle enthält dabei die ihr entsprechende durchschnittliche Höhe). Topologische Beziehungen lassen sich über Vergleiche der Matrixindizes der Rasterzellen bestimmen.

Vorerst sollen die Rasterzellen in einer Matrix gespeichert werden, aber es ist daran gedacht, die Rasterzellen

in einer Baumstruktur (Quadtree, siehe Kapitel 1.6.1) abzulegen. Lagebezogene Relationen lassen sich dann durch Vergleich der Blattknoteninhalte feststellen.

2.5.4 Struktur für Thamaelemente

Wie bereits erwähnt, müssen Format und Anzahl der verschiedenen Thamaelemente für ein Thema vom Benutzer festgelegt werden. Auch die zwischen diesen Elementen bestehenden Zusammenhänge müssen vom Benutzer in geeigneter Form angegeben werden.

Bei der Darstellung wird grundsätzlich zwischen globalen und lokalen Thamaelementen unterschieden. Globale Thamaelemente stellen im wesentlichen die Beschreibung des Aufbaus von lokalen Thamaelementen dar, ergänzt durch die graphische Darstellungsform für jedes Element.

Ein globales Thamaelement besteht aus einem Elementtyp, der gleichzeitig einen Code für dieses Element darstellt, einem Elementnamen und der Beschreibung der einzelnen Datenfelder für diesen Elementtyp.

Das lokale Thamaelement hierzu besteht aus dem Elementtyp und enthält die Werte zu den einzelnen Datenfeldern.

Um zu verdeutlichen, wie globale und lokale Thamaelemente aussehen können, folgen dazu zwei einfache Beispiele.

Beispiel 1:

Themaname: politische Karte

einige globale Thamaelemente:

Elementtyp: 1 Elementname: STAAT

Datenfelder: BEZEICHNUNG EINWOHNER FLÄCHE IN KM²
Format (alphanumerisch) (numerisch) (numerisch)

Elementtyp: 2 Elementname: BUNDESLAND

Datenfelder: BEZEICHNUNG KENNZAHL HAUPTSTADT
Format (alphanumerisch) (numerisch) (alphanum.)

Elementtyp: 3 Elementname: BEZIRK

Datenfelder: BEZEICHNUNG KENNZAHL
Format (alphanumerisch) (numerisch)

Elementtyp: 4 Elementname: STADT

Datenfelder: BEZEICHNUNG EINWOHNER FLÄCHE IN KM²
Format (alphanumerisch) (numerisch) (numerisch)

(Diese Beschreibung der globalen Thamaelemente kann neben dem Elementnamen noch Symbole für die graphische Darstellung des Thamaelements enthalten, wobei bei jedem Symbol angemerkt werden muß, ob es für die Darstellung von Flächen, Linien oder Punkten verwendet werden soll).

Lokale Thamaelemente:

<u>Elementtyp</u>	<u>Werte für die Datenfelder</u>		
1	ÖSTERREICH	7,510000	83853
2	STEIERMARK	6	GRAZ
2	TIROL	7	INNSBRUCK
2	BURGENLAND	1	EISENSTADT
.			
.			
.			
3	RADKERSBURG	617	
3	LIEZEN	612	
.			
.			
.			
4	GRAZ	251.000	20

Beispiel 2:

Themaname: Landnutzung

Einige globale Themalelemente:

Elementtyp: 1 Elementname: LANDWIRTSCHAFTLICH
GENUTZTE FLÄCHEN

Datenfelder: _____

Elementtyp: 11 Elementname: WEIDELAND

Datenfelder: NUMERIERUNG
Format (numerisch)

Elementtyp: 12 Elementname: ACKERLAND

Datenfelder: NUMERIERUNG
Format (numerisch)

·
·
·

Elementtyp: 2 Elementname: VERBAUTE FLÄCHEN

Datenfelder: _____

Elementtyp: 22 Elementname: INDUSTRIEGEBIET

Datenfelder: NUMERIERUNG
Format (numerisch)

·
·
·

usw.

Einige dazugehörige lokale Thamaelemente:

<u>Elementtyp</u>	<u>Datenfelder</u>
11	1
11	2
11	3
12	1
12	2
22	1
.	
.	
.	
usw.	

2.5.4.1 Relationen zwischen Thamaelementen

Man kann folgende Arten von Relationen zwischen Thamaelementen unterscheiden:

- zwischen globalen und lokalen Thamaelementen eines Themas
- zwischen globalen Thamaelementen eines Themas
- zwischen lokalen Thamaelementen eines Themas
- zwischen Thamaelementen verschiedener Themen.

Die Beziehungen zwischen globalen und lokalen Thamaelementen werden durch die Angabe des Elementtyps bei den lokalen Thamaelementen hergestellt.

Relationen zwischen globalen Thamaelementen eines Themas müssen vom Benutzer definiert werden. Hierfür gibt es zwei Möglichkeiten:

- Hierarchische Beziehungen innerhalb eines Themas werden

durch geeignete Wahl der Nummern für den Elementtyp festgelegt (wie in Beispiel 2: Die Elementtypen 11, 12, 13 sind dem Elementtyp 1 untergeordnet, die Elementtypen 121, 122 sind wiederum dem Elementtyp 12 untergeordnet, usw.).

- Hierarchierelationen können aber auch mit Hilfe eines Strukturbaums aufgebaut werden. Die Nummern für die Elementtypen können dann beliebig gewählt werden. Die Wurzel des Strukturbaumes bildet das in der obersten Hierarchieebene gelegene Themaelement (bzw. der Thema-name selbst, wenn sich die Hierarchie auf das gesamte Thema bezieht), die Söhne eines Knotens bilden die Themaelemente, die dem in diesem Knoten gespeicherten Themaelement untergeordnet sind.

Der Strukturbaum für Beispiel 2 ist in Figur 2.14 abgebildet.

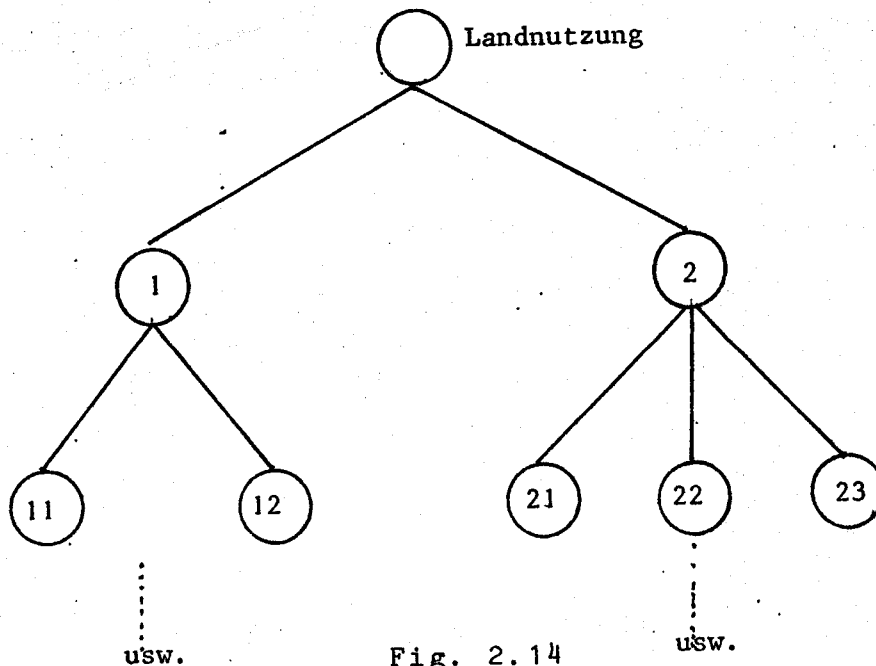


Fig. 2.14

In der DESBOD-Datenbank werden zu jedem Thema, das Hierarchierelationen enthält, die entsprechenden Struktur-bäume abgespeichert, auch wenn die Eingabe der Relationen nicht über Struktur-bäume erfolgt ist.

Für Relationen, die nicht hierarchisch sind, gilt: Die Relation muß einen Namen haben und alle Paare von globalen Thamaelementen (bzw. von Datenfeldern globaler Thamaelemente), die an dieser Relation beteiligt sind, müssen explizit angegeben worden sein. Dann wird in der DESBOD-Datenbank für diese Relation eine Zeigerdatei erstellt, auf die über den Relationsnamen zugegriffen werden kann. Auf diese Weise können auch Beziehungen zwischen globalen Thamaelementen verschiedener Themen angegeben werden, da die an der Relation beteiligten Thamaelemente ja nicht zu demselben Thema gehören müssen.

Die Speicherung von Relationen zwischen lokalen Thamaelementen ist zur Zeit nicht vorgesehen, aber zu einem späteren Zeitpunkt wohl möglich.

Das Schwergewicht bei der Implementierung der Datenbank wird auf der Darstellung der Thamaelemente liegen, weil der Benutzer in der Regel über diese auf den Datenbestand zugreifen wird.

2.5.5 Verbindung von Grundriß- und Thamaelementen

Die Verbindung von Grundriß- und Thamaelementen stellen die äußeren Relationen zwischen diesen beiden Mengen her. In der DESBOD-Datenbank gibt es nur eine äußere Relation, die über der Menge der lokalen Thamaelemente und der Menge der Flächen-, Linien- und Punktelemente definiert ist.

Diese Relation stellt aber eine bijektive Abbildung dar, d.h. es kommt jedem lokalen Thamaelement genau ein Grundrißelement zu und jedem Grundrißelement genau ein lokales Thamaelement.

Da also jedes Grundriß- und Themalement in der Relation genau einmal vorkommt, wird in diesem Fall von der Regel abgegangen, daß die Speicherung von Relationen die Speicherung der Elemente selbst nicht verändern darf. Es wird bei jedem Grundriß- und bei jedem lokalen Themalement ein Zeiger vorgesehen, der auf das entsprechende Grundriß- bzw. Themalement zeigt.

Somit sind alle Relationen, die in der DESBOD-Datenbank realisiert werden, beschrieben. Das Bild 2.15 vermittelt einen Überblick über den Aufbau der Datenbank.

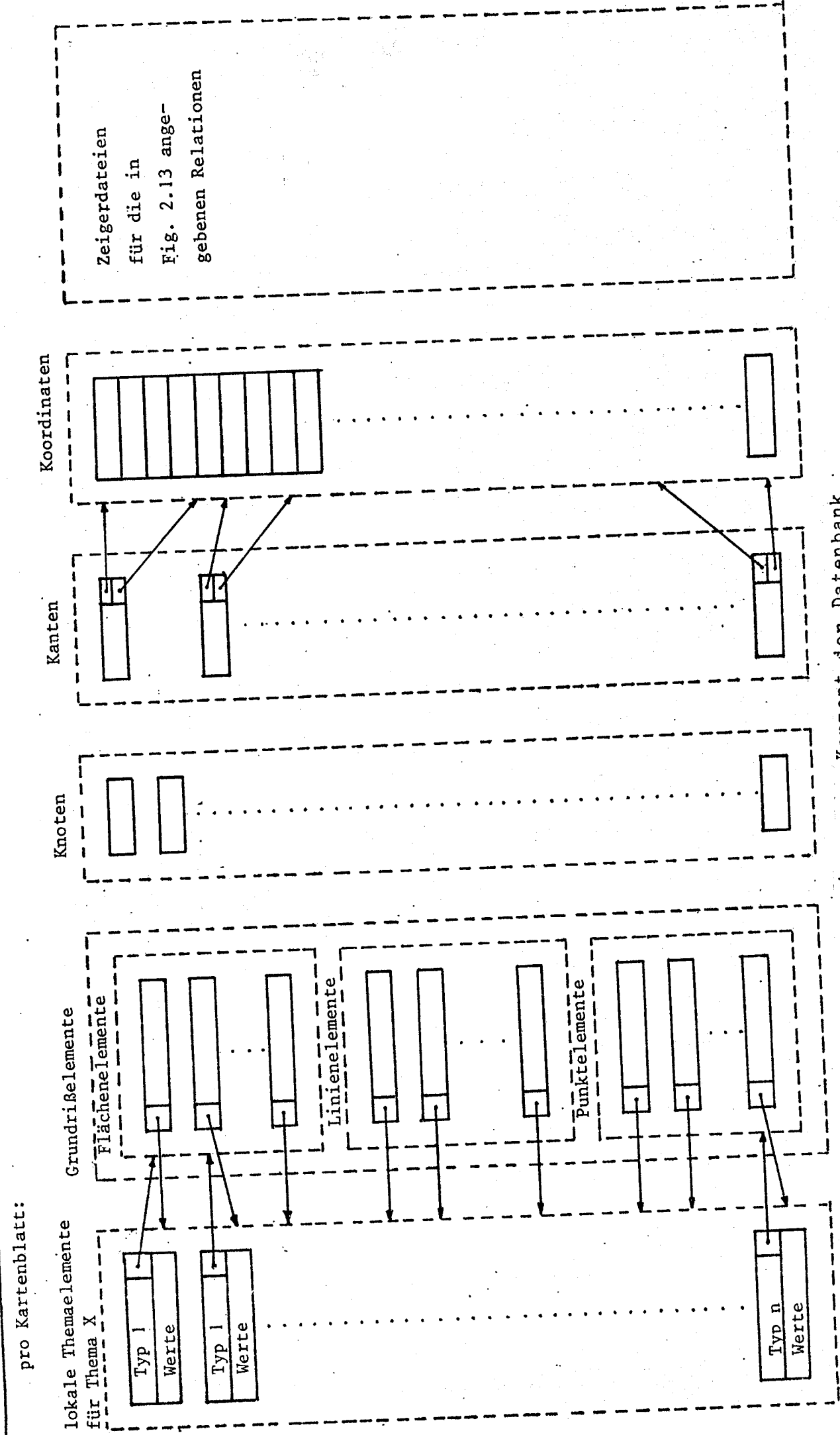
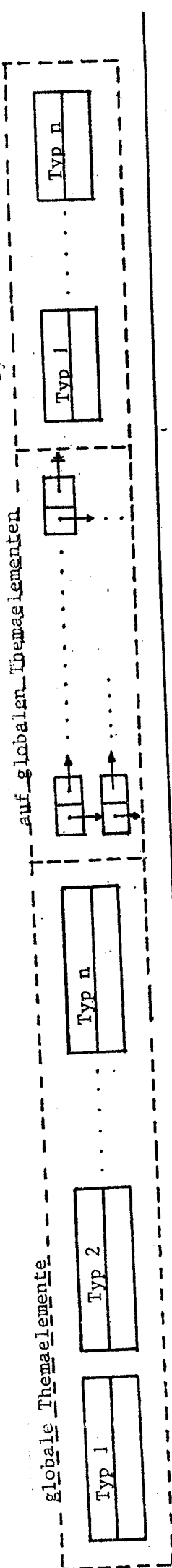


Fig. 2.15 Konzept der Datenbank

2.5.6 Datenverwaltung

Das Datenverwaltungssystem wird so aufgebaut, daß es allen in Kapitel 1.7 genannten Anforderungen gerecht wird:

Es sind Funktionen vorgesehen, die

- das Ändern und Löschen vorhandener Datenelemente ermöglichen,
- das Einfügen einzelner Datenelemente, sowie das Speichern ganzer Dateien ermöglichen,
- dem mißbräuchlichen Datenzugriff vorbeugen sollen,
- es erlauben, Datenelemente nach bestimmten Selektionskriterien aus der Datenbank zu extrahieren.

Insbesondere wird es notwendig sein, eine Datenbanksprache zu entwickeln, mit Hilfe derer die Grundfunktionen wie Ändern, Löschen und Einfügen von Datenelementen durchgeführt werden können. Weiters können mit der Datenbanksprache Strukturen auf den Datenelementen definiert werden und Informationen über den Datenbestand der Datenbank eingeholt werden.

Selektionen von Datenelementen aus der Datenbank werden auch mit Hilfe der Datenbanksprache vorgenommen. Man kann themabezogene, lagebezogene, sowie thema- und lagebezogene Selektionen unterscheiden, außerdem kann auch auf die gespeicherten Relationen zugegriffen werden.

Lagebezogene Selektion: z.B. alle Grundrißelemente eines bestimmten Kartenblattes, oder nur die Punktelemente eines Kartenblattes, etc.

Themabezogene Selektion: z.B. alle globalen Thamaelemente

eines Themas, oder die lokalen Themaelemente eines bestimmten Elementtyps, etc.

Lage- und Themabezogene Selektionen: z.B. alle Thema- und Grundrißelemente eines Kartenblattes, oder nur die Grundrißelemente, die einem bestimmten Themaelementtyp entsprechen, etc.

Selektion nach Relationen: z.B. alle Flächenelemente, die einem bestimmten Flächenelement benachbart sind, oder Angabe der Hierarchierelation zwischen globalen Themaelementen, etc..

Um den Datenbestand vor unberechtigten Zugriffen zu schützen, wird jede Datei mit einem Schutzkennwort versehen. Dieses Schutzkennwort gibt an, welche Privilegienstufe notwendig ist, um eine der Funktionen Lesen, Schreiben oder Löschen auf dieser Datei durchzuführen.

2.6 Analyse und Ausgabe von Daten

Das Analyse- und Ausgabesystem von DESBOD ist ein interaktives Programmsystem, das die Weiterverarbeitung der in der DESBOD-Datenbank gespeicherten Daten ermöglicht. Das Programmsystem besteht wiederum aus mehreren Teilsystemen entsprechend Figur 2.16, die durch entsprechende Benutzerkommandos aktiviert werden:

- a) Übersetzer für die Kommandosprache
- b) Auskunft- und Kontrollsystem
- c) Datenauffindung- und Manipulationssystem (dieses System bildet die Schnittstelle zur Datenbankverwaltung)
- d) Datenanalyse
- e) Ausgabe der Daten

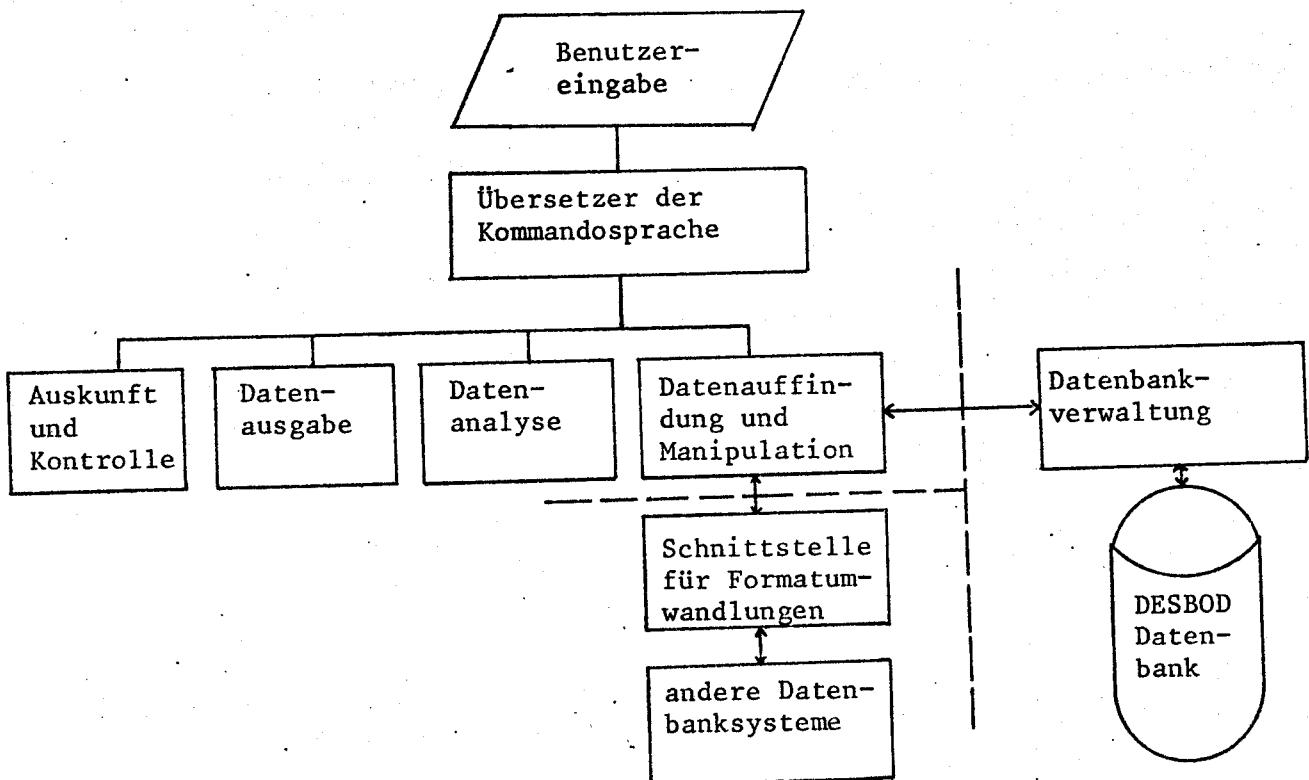


Fig. 2.16
Diagramm der Teilsysteme

Um einen optimalen Schutz für die in der Datenbank gespeicherten Daten zu erreichen, können diese von einem Benutzer des Auswertungssystems weder geändert noch gelöscht werden. Jeder Benutzer bekommt einen eigenen Arbeitsspeicherbereich, in dem er durch spezielle Selektionskommandos Dateien aus der Datenbank ablegen kann. Diese Dateien im Arbeitsspeicher stehen ihm dann für seine Auswertungen zur Verfügung und können auch nach Belieben geändert oder gelöscht werden.

Für eine sinnvolle Datenauswertung müssen im allgemeinen mindestens eine Themabeschreibung, sowie die Grundrißelemente für dieses Thema in den Arbeitsspeicher geladen werden. Für die Speicherung der Grundrißelemente im Arbeitsspeicher gibt es - abhängig von der Art der zu erfolgenden Auswertungen - 2 Möglichkeiten: Speicherung in Rasterformat oder im Vektorformat, wobei es keine Rolle spielt, in welchem Format die Elemente in der Datenbank abgelegt sind.

Jeder Arbeitsspeicher besitzt außerdem einen Index, in dem alle in diesem Arbeitsspeicher enthaltenen Dateien aufgeführt sind; ein übergeordneter Hauptindex ermöglicht die Verwaltung aller Arbeitsspeicher, die von den verschiedenen Benutzern angelegt wurden. Dadurch ist auch die Übertragung von Daten von einem Arbeitsspeicherbereich in einen anderen möglich (Fig. 2.17).

Gerade das Analyse- und Ausgabesystem eines GIS ist jener Teil, der am meisten von anwendungsorientierten Überlegungen geprägt ist und oft nur für ganz besondere Anforderungen und Anfragen geeignet ist. Es wird hier versucht, das DESBOD-Analyse- und Ausgabesystem so allgemein und flexibel zu halten, daß mit einem Grundstock an Möglichkeiten durch Kombinationen derselben auch komplexe und komplizierte Auswertungen möglich sind.

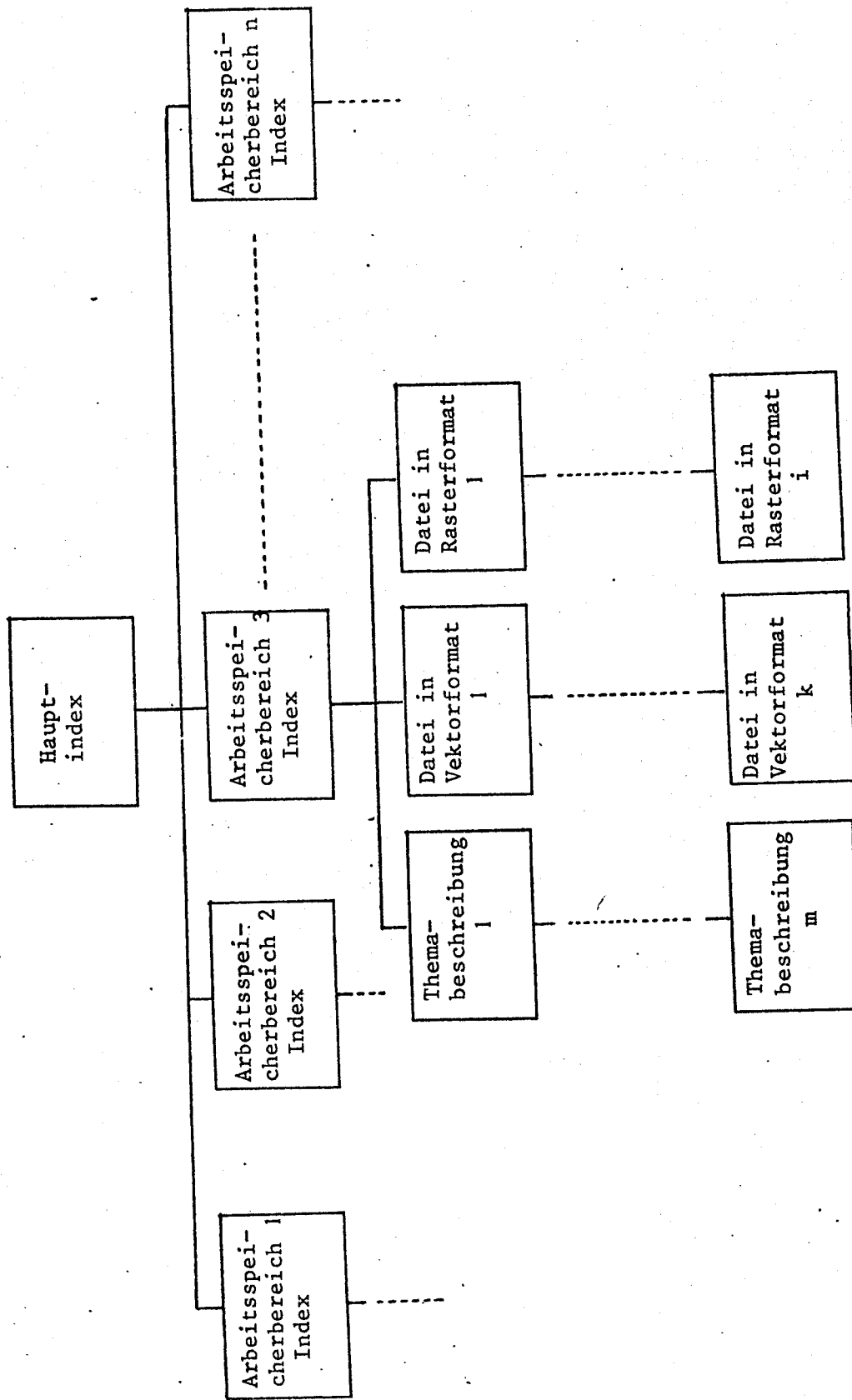


Fig. 2.17

Aufteilung des Arbeitsspeichers für das Analyse- und Ausgabesystem

2.6.1 Übersetzer für die Kommandosprache

Die Steuerung der Analyse- und Ausgabe der Daten könnte im wesentlichen mit einer Menütechnik oder Kommandosprache erfolgen (Autometric, 1979; Autometric 1978). Es wird für DESBOD eine Kommandosprache gewählt, weil im Gegensatz zur Menütechnik die Flexibilität der Eingabe gegeben ist, die man bei den oben erwähnten Anforderungen verlangt. Außerdem bietet eine Kommandosprache die Möglichkeit, batch-orientierte Verarbeitungen durchzuführen.

Die Syntax eines Kommandos ist bestimmt durch einen Befehl (Kommando) und dazugehörige Parameter und Optionen. Kommandos werden immer auf der Tastatur eines Bildschirmgerätes eingegeben; Parameter und Optionen können manchmal auch über graphische Eingabegeräte (Lichtgriffel, Fadenkreuz am graphischen Bildschirm) eingegeben werden.

Der Übersetzer interpretiert jedes vom Benutzer eingegebene Kommando, d.h. er prüft, ob es sich um ein gültiges Kommando handelt und ob alle notwendigen Parameter und Optionen für dieses Kommando vorhanden sind; er aktiviert dann die entsprechenden Programme. Er fordert den Benutzer zur Eingabe von Parametern auf, wenn diese vergessen wurden, aber zur Durchführung eines Kommandos unbedingt notwendig sind.

2.6.2 Auskunfts- und Kontrollsystem

Das Auskunfts- und Kontrollsystem dient der Unterstützung des Benutzers, aber auch der Überwachung und Kontrolle seiner Aktivitäten. Es bietet Informationen über verbrauchte Betriebsmittel (z.B. verbrauchte Rechenzeit), Zeit des Sitzungsbeginns, etc.. Man kann auch ein Inhalts-

verzeichnis der DESBOD-Datenbank, eines Arbeitsbereiches oder eine Liste der vorhandenen Themen erhalten. Eine HELP-Funktion ermöglicht dem Benutzer, die Syntax von Kommandos, eine Beschreibung von Kommandos und eine Erläuterung der benötigten Optionen und Parameter abzufragen. Das Kontrollsystem prüft Zugriffsberechtigungen (z.B. für das Löschen von Sätzen aus der DESBOD-Datenbank) und beendet das Programm bei Eingabe eines bestimmten Kommandos.

2.6.3 Datenauffindungs- und Manipulationssystem

Das Datenauffindungs- und Manipulationssystem stellt die Schnittstelle zwischen der DESBOD-Datenbank und den Benutzerbereichen dar. Da ein Benutzer keinen direkten Zugriff auf Daten in der Datenbank hat, muß er diese erst in einen Arbeitsbereich bringen. Dieses System erlaubt es nun, solche Arbeitsbereiche anzulegen und Daten aus der Datenbank nach bestimmten Gesichtspunkten in diese Arbeitsbereiche zu bringen. Will z.B. ein Benutzer die Karte der steirischen Gemeinden bearbeiten, so muß er einen Arbeitsbereich anlegen und kann dann die gewünschte Karte aus der Datenbank extrahieren. Erst jetzt kann diese Karte bearbeitet werden. So ist es etwa möglich, das Format von Vektor- in Rasterstruktur zu ändern; auch Maßstabsänderungen gehören zu diesem System.

Die Auffindung der Daten ist in Abhängigkeit von lage- oder themabezogenen Bedingungen möglich, oder auch in Abhängigkeit von Entfernungen zu einem bestimmten Kartenobjekt.

Mehrere Arbeitsbereiche lassen sich zu einem Arbeitsbereich verschmelzen, neue Arbeitsbereiche - um z.B. Ergebnisse von Verknüpfungen mehrerer Karten aufzunehmen - können erzeugt werden.

Natürlich gibt es auch Kommandos zum Löschen oder Überschreiben bestimmter Karten oder ganzer Arbeitsbereiche. Arbeitsbereiche können gesichert werden, so daß sie für weitere Bearbeitungen erhalten bleiben. Dadurch ist es nicht notwendig, jedesmal alle Operationen zu wiederholen, wenn die Bearbeitung eines Arbeitsbereiches für längere Zeit unterbrochen wurde.

Eine Schnittstelle ist vorgesehen, um Daten in "fremde" Datenformate umzuwandeln oder von solchen zu dekodieren. Dadurch soll in einer späteren Ausbaustufe die Interaktion zwischen DESBOD und anderen Informationssystemen ermöglicht werden, d.h. Daten aus der DESBOD-Datenbank können in andere Informationssysteme übertragen werden, und Daten aus anderen Systemen können in die DESBOD-Datenbank eingebracht werden. (Fig. 2.18).

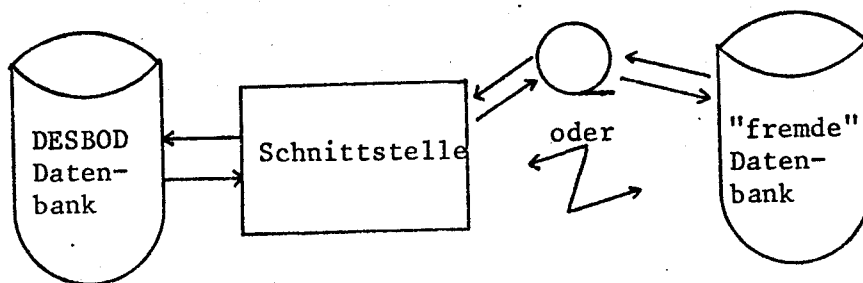


Fig. 2.18

Interaktion von DESBOD mit anderen Informationssystemen

2.6.4 Datenanalyse

(a) Einige Beispiele

Die Datenanalysekommandos machen den größten Teil aller verfügbaren Kommandos aus. Sie geben dem Benutzer die Möglichkeit, mit den Daten etwas zu tun. So kann man zum Beispiel Antwort auf folgende Fragen erhalten:

- Wir haben einen Flächenwidmungsplan eines bestimmten Studiengebietes und Anforderungen, die für einen bestimmten Zweck an ein Grundstück gestellt werden. Welche Grundstücke erfüllen diese Anforderungen?
- Wie viele Kilometer Wasserwege der Wassergüte X gibt es in der Steiermark?
- Bestimme alle Gebiete der Steiermark mit Vegetationstyp Y, die höher als 1000 m liegen.
- Wieviele Kilometer Luftlinie sind es von Graz nach Wien?

Zur Beantwortung der ersten Frage sind folgende Schritte erforderlich:

1. Anlegen eines Arbeitsbereiches mit der Karte des Flächenwidmungsplanes.
2. Selektives Aufsuchen aller Objekte mit den geforderten Eigenschaften.
3. Eventuell Ausgabe der Ergebnisse auf dem graphischen Bildschirm.

Für die zweite Frage wird folgender Weg sinnvoll sein:

1. Aufsuchen des Themas "Wassergüte der Steiermark".
2. Selektives Aufsuchen aller Objekte der Wassergüte X und Ablegen des Ergebnisses auf einem Arbeitsbereich.
3. Berechnen der Längen und Summierung der Längen aller Objekte dieses Arbeitsbereiches.

Zur drittten Frage:

1. Holen der Vegetationskarte (in Vektorformat) der Steiermark in den Arbeitsbereich.
2. Holen der Höhenschichtenlinienkarte (in Rasterformat) der Steiermark in den Arbeitsbereich.
3. Extrahiere alle Objekte mit Vegetationsart Y aus der Vegetationskarte und lege sie im Arbeitsbereich ab.
4. Umwandeln der extrahierten Objekte in Rasterform.
5. Extrahiere alle Gebiete aus dem Höhenraster, die höher als 1000 m liegen und lege sie im Arbeitsbereich ab.
6. Bilde den Durchschnitt beider Bereiche.

Zur Beantwortung der vierten Frage schließlich ist es notwendig, ein entsprechendes Kommando zur Berechnung von Entfernungen einzugeben. Mit dem Fadenkreuz oder dem Lichtgriffel des graphischen Bildschirms fährt man dann beide Orte am Bildschirm an. Das System berechnet dann die Luftlinie zwischen den zwei Orten.

(b) Zu realisierende Kommandos

Es gibt drei Haupttypen von Analysekommandos. Meß-Kommandos berechnen Flächeninhalte, Weglängen und Entfernungen. Statistik-Kommandos ermöglichen die Berechnung von Häufigkeitstabellen, Histogrammen und statistischen Parametern wie Mittelwert, Standardabweichung, Median etc.. Schließlich erlauben die Synthese-Kommandos mehrere Karten zu verknüpfen, um daraus neue Karten zu erzeugen.

Mit Hilfe von Meß-Kommandos können sowohl von Vektor- als auch Rasterstrukturen Flächeninhalte von Objekten, Prozentanteil an der Gesamtfläche eines Gebietes und Längen von Strecken berechnet werden. Bei der Streckenberechnung soll sowohl die Luftlinie als auch der Weg entlang eines

bestimmten Pfades betrachtet werden können. Weiters ist es möglich, die Länge oder Fläche aller Elemente einer bestimmten Kategorie (z.B. Autobahnen, Bundesstraßen, Eisenbahnen, Seen) zu bestimmen. Damit kann z.B. die Frage nach der Länge des Autobahnnetzes der Steiermark und nach dem prozentuellen Anteil am Gesamtstraßennetz der Steiermark beantwortet werden.

Die einfachsten statistischen Analysen der Daten sind die Berechnungen von statistischen Kenngrößen, wie Mittelwert, Median, Varianz und Standardabweichung von Kartenobjekten. Beispiele dafür wären etwa die mittlere Fläche der steirischen Bezirke oder die Variabilität der Flächen der Bezirke um den mittleren Wert aller Bezirke. Dazu kommen noch Häufigkeitsdiagramme, Histogramme und Veranschaulichungen von Verteilungen und Vorkommen der Objekte einer Karte. Zu diesen einfachen beschreibenden Verfahren kommen später auch noch Methoden der analytischen Statistik. Damit wird der Benutzer die Möglichkeit erhalten, aus vorhandenen Datensätzen Aussagen und Prognosen herzuleiten. Dazu werden statistische Methoden wie Regressionsanalyse, Varianzanalyse etc. verwendet. (Ripley, 1981; Bahrenberg und Giese, 1975; Journel and Huijbregts, 1978).

Synthese-Kommandos erlauben die Durchführung von Mengenoperationen sowohl auf der Menge der Thamaelemente als auch der Menge der Grundrißelemente. Mengenoperationen stellen die Verknüpfung von Datenelementen durch die logischen Operationen "und", "oder" und "nicht" dar.

Durch Verknüpfung von Thamaelementen des gleichen Themas oder verschiedener Themen lassen sich Fragen beantworten wie:

"Welche steirischen Orte liegen höher als 1000 m"? oder
"In welchen steirischen Bezirken wird Wein angebaut"?

Mit mengentheoretischen Operationen auf Grundrißelementen (Durchschnittsbildung, Vereinigung, Differenz zweier Mengen) können sowohl Raster- als auch Vektordaten verknüpft werden. Fig. 2.19 zeigt ein Beispiel einer Rasterverknüpfung, Fig. 2.20 zeigt eine Vektordatenverknüpfung (Polygonschnitte).

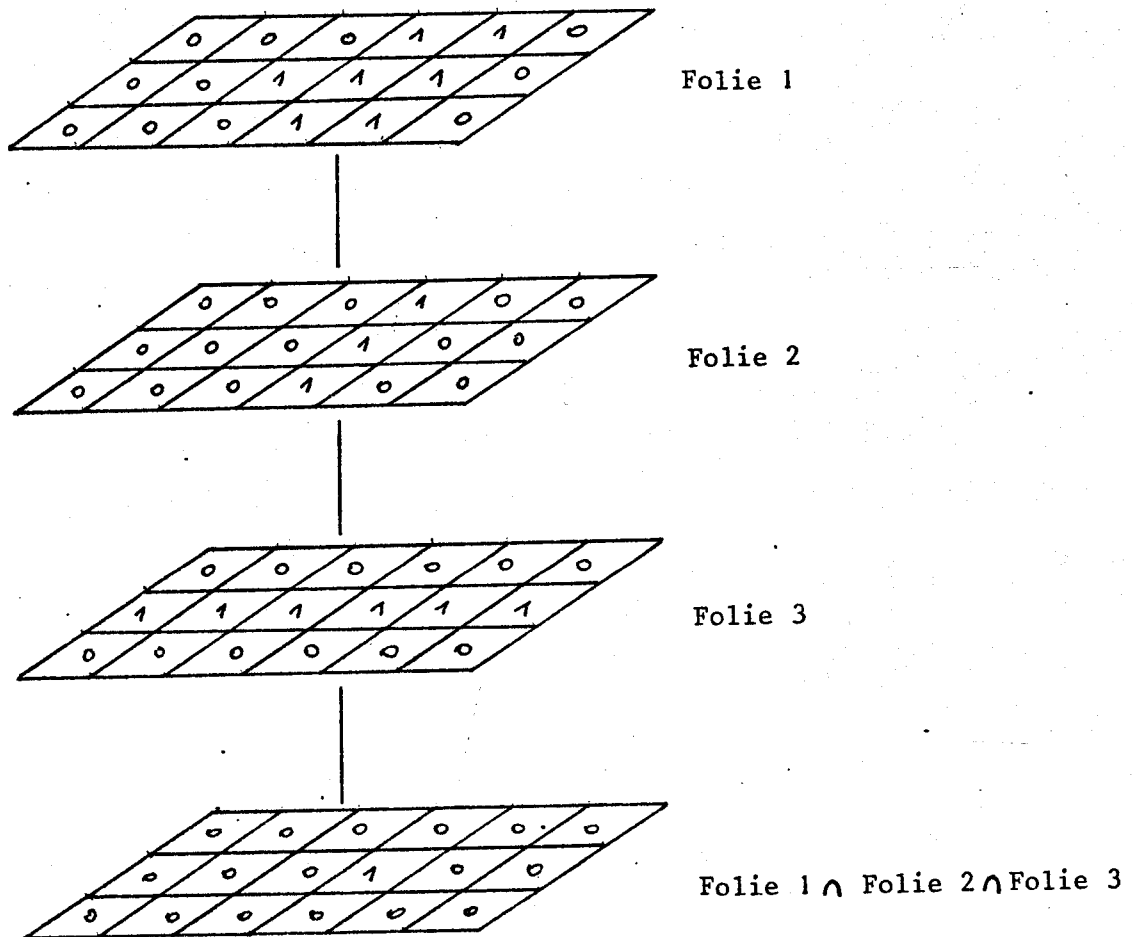


Fig. 2.19
Rasterverknüpfung

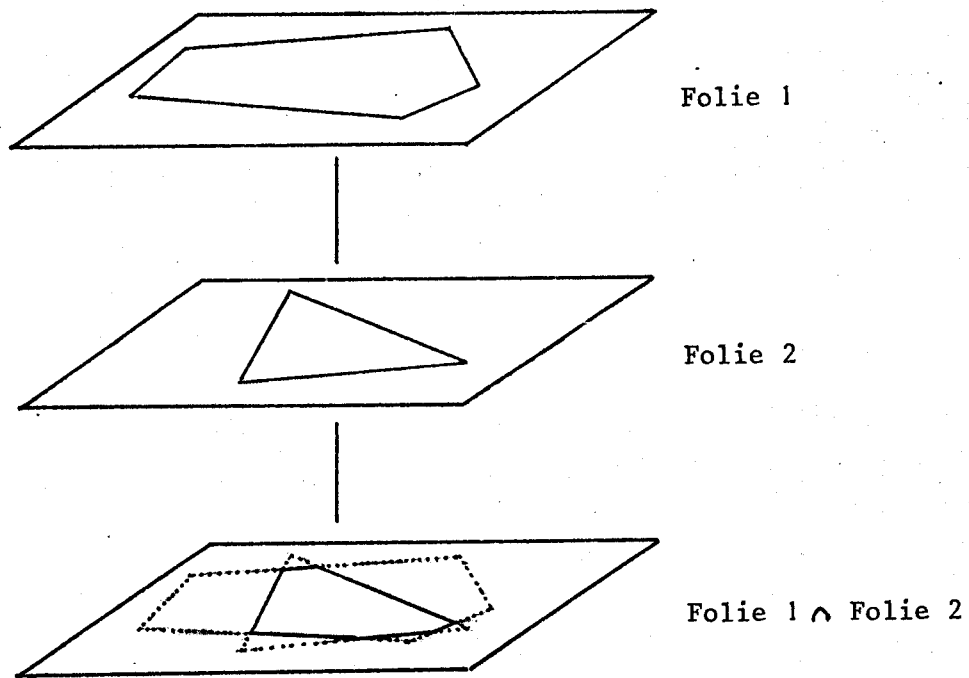


Fig. 2.20
Vektordatenverknüpfung

Die Verknüpfung (Übelagerung) von mehreren Folien einer Karte ist wegen der Kommutativität und der Assoziativität von Durchschnitt- und Vereinigungsbildung von Mengen durch Hintereinanderausführen der Verknüpfungen von je zwei Kartenblättern leicht durchführbar. Wegen der Möglichkeit

der Umwandlung von Vektor- in Rasterdatenstruktur ist auch die Verknüpfung von Karten, die in verschiedenen Strukturen gespeichert sind, möglich (Fig. 2.21).

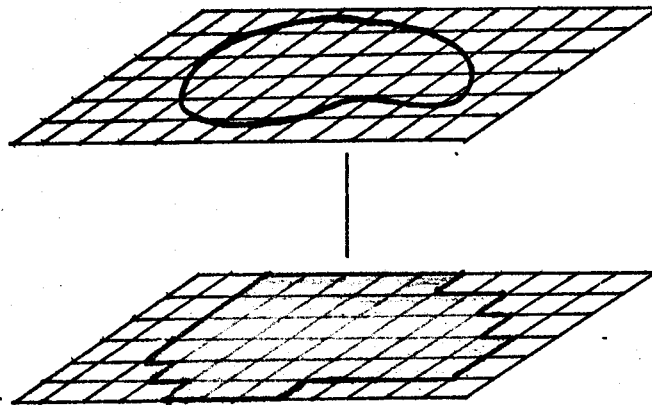


Fig. 2.21
Vektor in Rasterumwandlung
("Scan-Konversion")

2.6.5 Ausgabe der Daten

Das Ausgabesystem hat die Aufgabe, die Ergebnisse von Analysen und Anfragen an das DESBOD-System in tabellarischer oder graphischer Form aufzubereiten und auszugeben (Fig. 2.22).

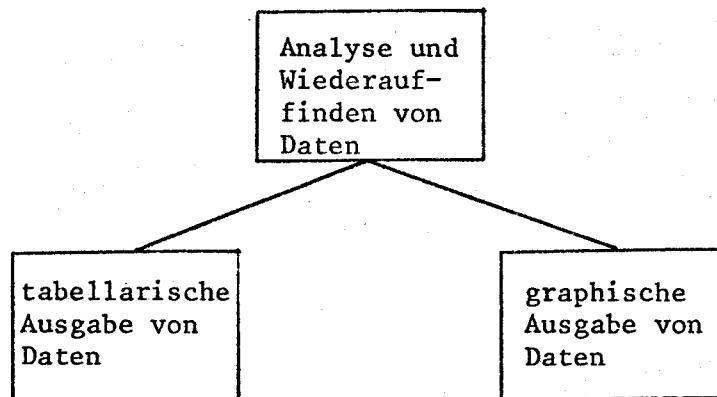


Fig. 2.22
Ausgabe der Daten

Die tabellarische Ausgabe der Daten umfaßt alle Darstellungen von Daten, für die keine Graphikgeräte erforderlich sind. Es sind Listen, Tabellen und Texte (z.B. die Liste der Namen aller steirischen Bezirke, die Beschreibung eines bestimmten Kartenthemas, Häufigkeitstabellen, Statistiken). Für diese Ausgaben werden Zeilendrucker und Bildschrimgeräte verwendet.

Die graphische Ausgabe der Daten erfolgt mit Hilfe von graphischen Bildschirmen, einem Zeilendrucker und einem Plotter. Durch die Überdruckmethode können Zeilendrucker für die einfache und billige Darstellung von Rasterdaten herangezogen werden; schönere Darstellungen von Raster- und Vektordaten gelingen mit dem elektrostatischen Drucker/Plotter oder Filmschreiber, und für die Ausgabe von Vektordaten wird ein Zeichenstiftplotter verwendet.

Die graphischen Ausgabekommandos erlauben die Darstellung von Karten in jedem beliebigen Maßstab. Dazu ist vorgesehen, die Kartenlegenden in Form von Texten und Symbolen mitauszugeben. Objekte einer Karte können nach Belieben schraffiert, beschriftet oder mit bestimmten Symbolen aufgefüllt werden (Fig. 2.23).

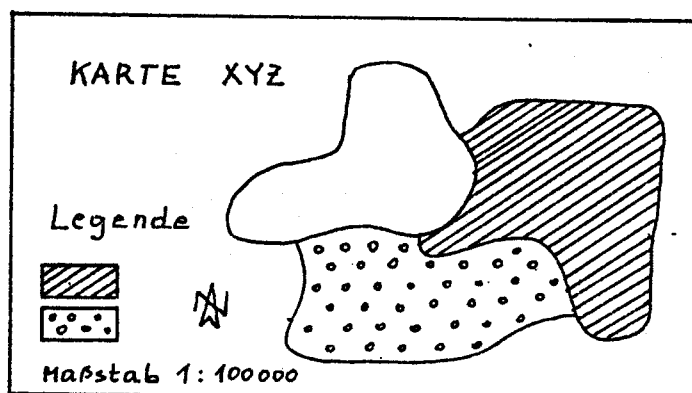


Fig. 2.23
Darstellung einer Karte

Durch Festlegen eines Zeichenfensters ist es möglich, auch nur Ausschnitte aus einer Karte zu zeichnen ("Lupenfunktion"). Kommandos zur Umwandlung von Vektor- in Rasterform und umgekehrt von Raster- in Vektorform für Darstellungszwecke gestatten eine Vielzahl von Möglichkeiten, um ortsbezogene Daten anspruchsvoll zu präsentieren.

Aus interpolierten Höhendaten lassen sich dreidimensionale Geländemodelle von beliebigen Raumpunkten aus betrachten und zeichnen. Mehrfarbige Darstellungen gehören ebenso zu diesem System wie farbige Rasterdarstellungen in Fernsehqualität.

Bei allen graphischen Ausgabeoperationen kann das Ausgabe-medium frei gewählt werden. Allerdings sind geräteabhängige Darstellungsmöglichkeiten (z.B. kann eine Karte in Vektorform nicht direkt auf einem Zeilendrucker ausgegeben werden) zu beachten.

ANHANG A

Darstellung von Zwischenergebnissen der Systemrealisierung

A.1 Einleitung

In diesem Abschnitt sollen einige Ergebnisse einer beginnenden Systemrealisierung präsentiert werden, die neben der, oder parallel zur Konzeptentwicklung erreicht wurden. Dabei ging es vor allem darum, erste Möglichkeiten zur Erfassung von Grundrißelementen zu schaffen, um damit einen dingenden Bedarf zur Nutzung der graphischen Datenverarbeitung zu befriedigen. Dazu wurde ein Digitalisierungssystem entwickelt, welches sich zurzeit noch in der Grundstufe befindet und nach und nach ausgebaut werden soll. Das Programm ist menügesteuert und bietet folgende Möglichkeiten:

- Einmessen eines Kartenblattes,
- Digitalisieren der Grundrißelemente,
- Editieren der digitalisierten Grundrißelemente,
- automatische Fehlerprüfung (Verifikation) der Daten,
- graphische Kontrolle.

Die gegenwärtigen Möglichkeiten werden im folgenden kurz diskutiert und mit Beispielen erläutert.

A.2 Beschreibung der Funktionen

Das Einmessen, Digitalisieren und Editieren wurde in dem Abschnitt von Teil 2 besprochen, der das Digitalisierungssystem behandelt. Für die Speicherung der Daten wurde das

Chain-Format (siehe Abschnitt 2.4.3) gewählt. Das Verifizieren der Daten beschränkt sich vorerst auf die Überprüfung, ob Kanten mehrmals vorkommen, ob Anfangs- und Endknoten aneinanderliegender Kanten zusammenpassen; es lassen sich von Kanten umschlossene Regionen extrahieren.

Die graphische Kontrolle erlaubt eine Darstellung der digitalisierten Kanten mit den Kantennummern und Knotennummern, und zwar derzeit sowohl auf dem TEKTRONIX-Bildschirm als auch auf dem VERSATEC-Plotter.

Diese Daten können mit Hilfe eines Zeichenprogrammes (Voxland, 1981) in verschiedenen geographischen Projektionen und Maßstäben gezeichnet werden. Es existiert auch die Möglichkeit, Flächen zu schraffieren und einfache Kartenüberschriften anzubringen.

Für Testzwecke wurden bis jetzt die Bezirksgrenzen Österreichs und die wichtigsten Flüsse des Landes digitalisiert.

A.3 Beispiele

Figuren A.1 bis A.4 zeigen einige Beispiele, welche die derzeitigen Ausgabemöglichkeiten illustrieren.

Hiezu wurde als Digitalisiervorlage die Übersichtskarte 1:500 000 von Österreich verwendet. Der Bezirksfile umfaßt 14241 Punkte und 305 Kanten, der File des Flußnetzes hat 2877 Punkte und 69 Kanten. Für das Digitalisieren der Bezirksgrenzen wurden ca. 20 Stunden, für die Flüsse ca. 5 Stunden benötigt.

DIE BEZIRKE ÖSTERREICHS



Fig. A.1

DIE BEZIRKE DER STEIERMARK

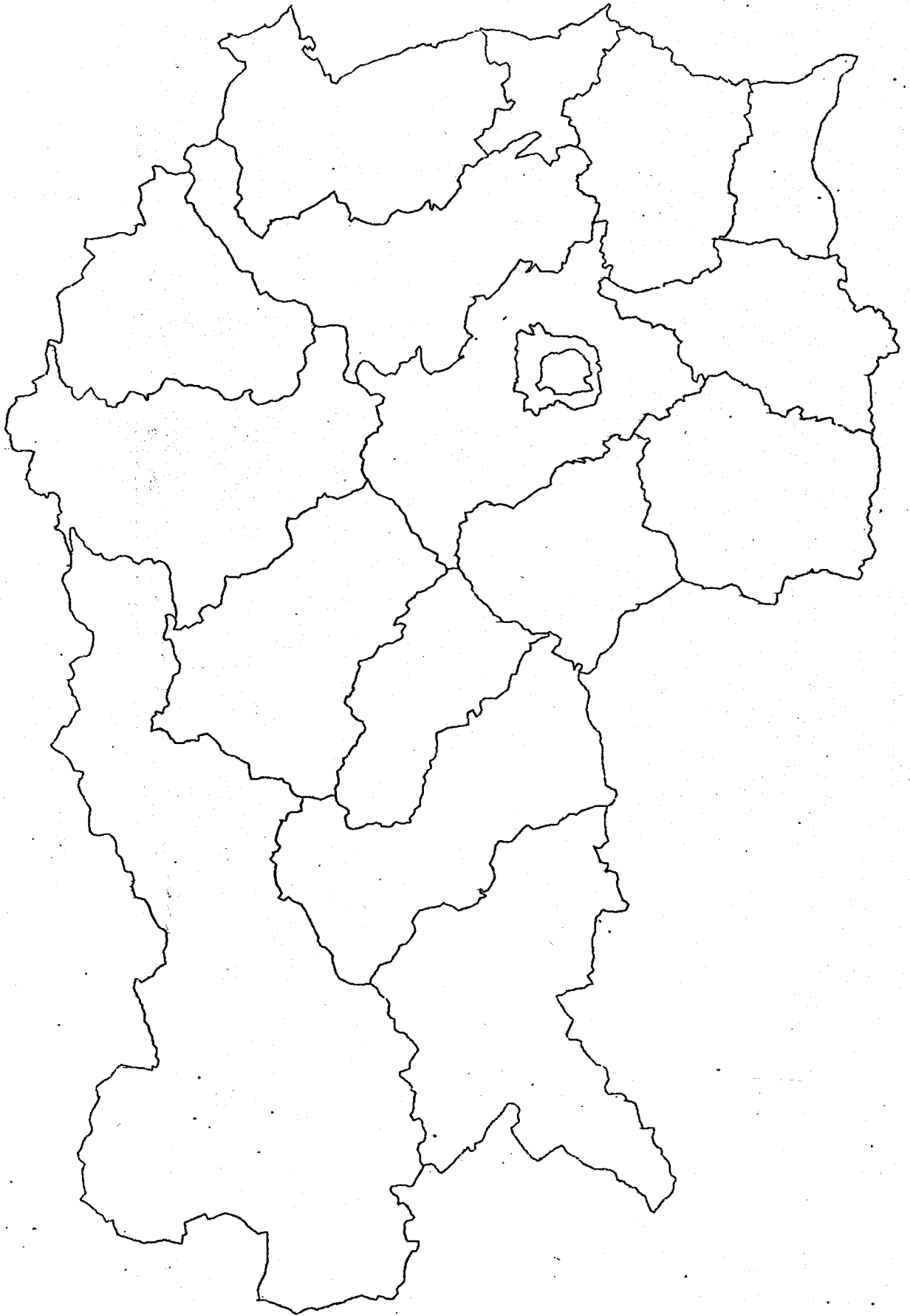


Fig. A.2

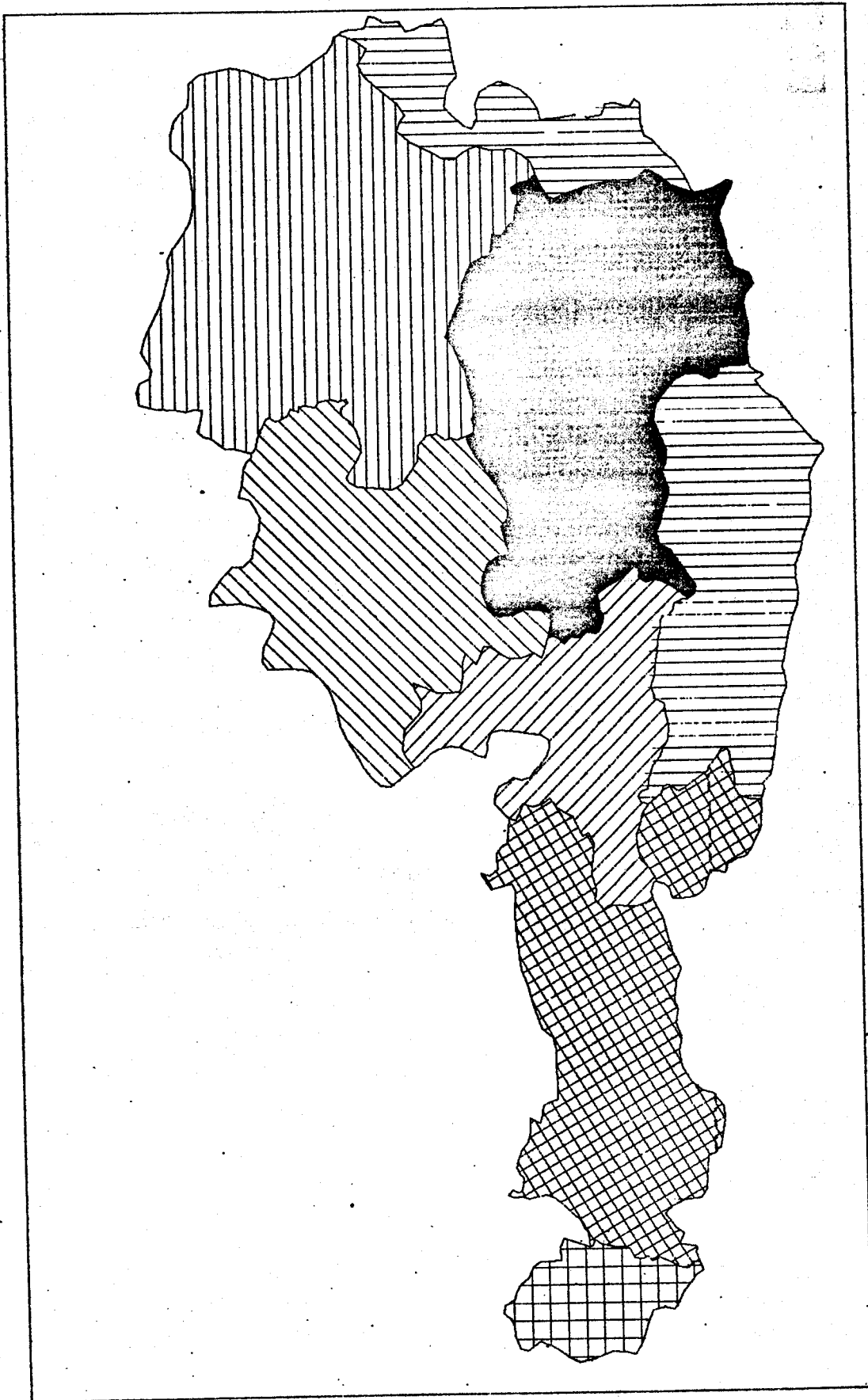


Fig. A.3

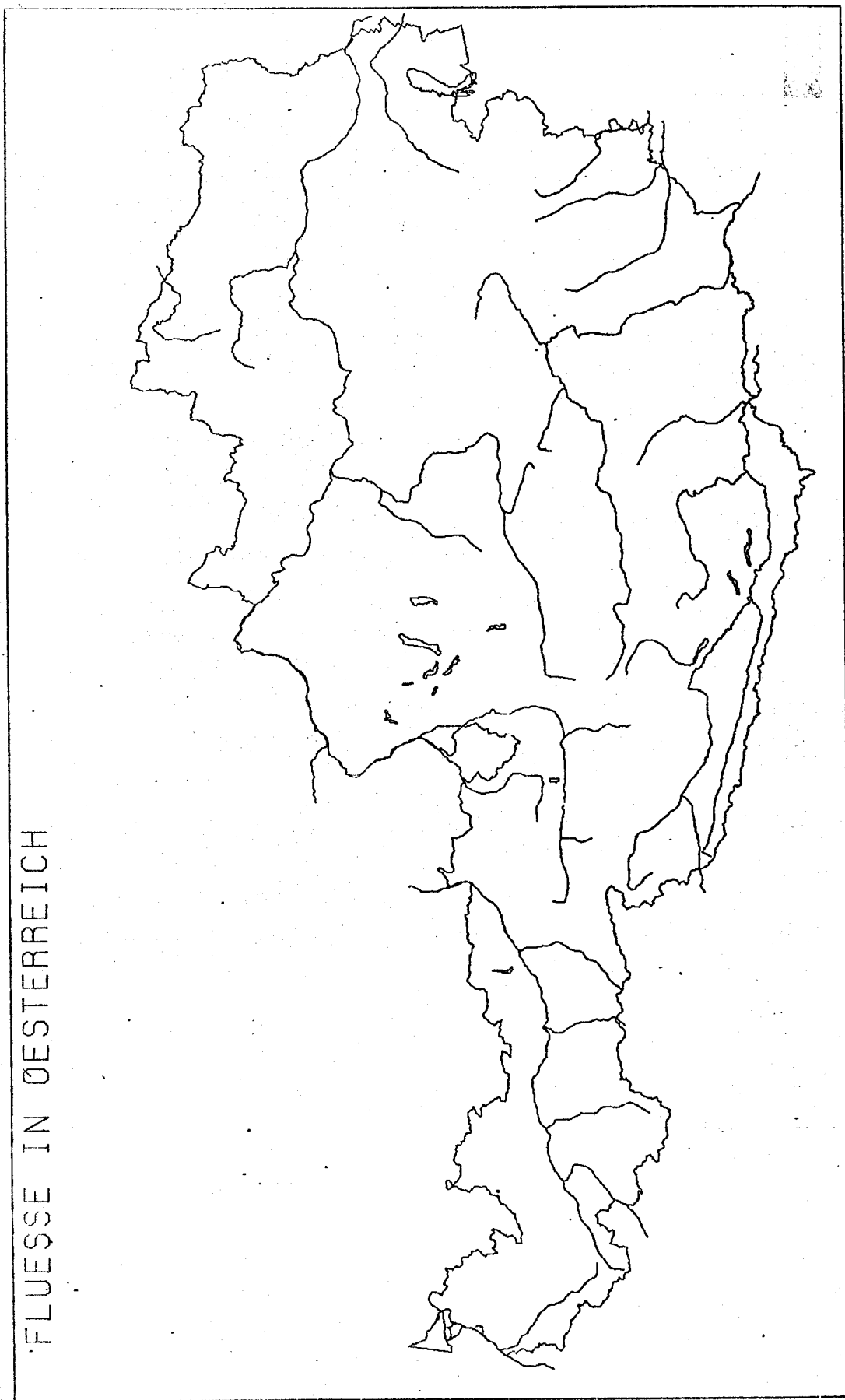


Fig. A.4

Anhang B

Die Mathematik der Karten

B.1 Einleitung

In diesem Abschnitt soll eine kurze Beschreibung der Anwendungen mathematischer Theorien bei der Darstellung und Bearbeitung von Karteninhalten in einem Computer gegeben werden. Es zeigt sich, daß eine Behandlung nichttrivialer Beziehungen zwischen Kartenobjekten, ohne eine Anwendung von Ergebnissen dieser Theorien kaum möglich ist.

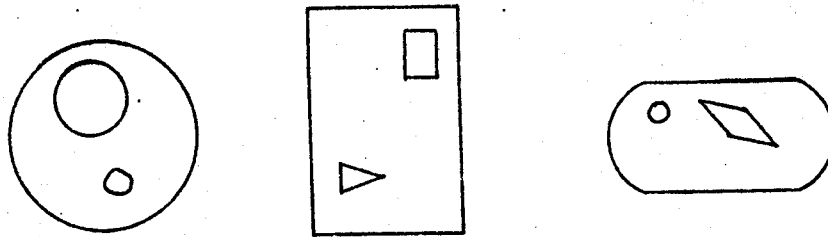
Die folgenden Ausführungen stellen eine Grundlage für weitere Überlegungen dar, wie weit diese Theorien für die Behandlung der Daten in DESBOD (z.B. für die Verifikationsphase beim Erfassen der Grundrißelemente) herangezogen werden müssen und ob manche Probleme (z.B. die Abspeicherung von Grundrißelementen in der Datenbank) damit effizienter gelöst werden können.

Bei der Erfassung von Grundrißelementen gilt es, die topologischen Beziehungen von der Karte in eine digitale (zahlenmäßige) Form überzuführen. Diese digitalen Daten werden auf einem Speichermedium (Magnetplatte, Magnetband) abgelegt und dienen in der Folge zur Analyse und Darstellung der gespeicherten Information.

Um alle wesentlichen Beziehungen, die zwischen den Grundrißelementen bestehen, computergerecht darstellen zu können, bedient man sich der Methoden verschiedener mathematischer Theorien (White, 1979).

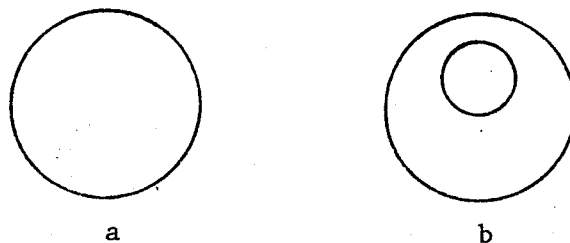
B.2 Topologie

Betrachten wir zuerst die Topologie, weil sie die Grundlage für eine automatisierte Kartographie darstellt, genauso wie unsere Fähigkeit des räumlichen Inbeziehungssetzens die Grundlage der manuellen Kartographie ist. Die drei Objekte in Fig. B.1 sind topologisch äquivalent, d.h. sie können durch stetige Deformation ineinander übergeführt werden.



topologisch äquivalent

Fig. B.1



topologisch verschieden

Fig. B.2

Eine solche stetige Deformation nennt man topologische Abbildung oder Homöomorphismus. Bei einer solchen Abbildung sind z.B. keine Schnitte (Unstetigkeiten) erlaubt.

Die drei Objekte in Fig. B.1 sind homöomorph, die zwei in Fig. B.2 sind es nicht, weil im Inneren von a ein Schnitt gemacht werden muß, um sie in b überzuführen.

Mit Hilfe der topologischen Abbildungen kann man einfache Objekte deformieren und diese zu komplexen Strukturen kombinieren. Das führt zur kombinatorischen Topologie.

Die Bausteine der kombinatorischen Topologie sind die sogenannten Zellen.

Eine 0-Zelle ist ein Punkt; eine 1-Zelle ist eine einfache Kurve; eine 2-Zelle ist eine Kreisscheibe oder ein homöomorphes Bild derselben. Im allgemeinen ist eine n-Zelle ein homöomorphes Bild einer offenen Umgebung des \mathbb{R}^n , d.h.

$$\{x \in \mathbb{R}^n \mid d(x,0) < 1\}.$$

Mit diesen Hilfsmitteln kann man eine Karte mit topologischen Begriffen beschreiben; auf eine Art also, die nicht von Projektionen, Maßstäben oder von einer bestimmten Gestalt abhängt.

Eine Karte ist ein zweidimensionaler Komplex bestehend aus einer Ansammlung von 0-Zellen, 1-Zellen und 2-Zellen, so daß gilt:

1. Jede 2-Zelle wird von 1-Zellen begrenzt.
2. Jede 1-Zelle wird von 0-Zellen begrenzt.
3. Jede 0-Zelle liegt an der Grenze einer 1-Zelle.

4. Jede 1-Zelle liegt an der Grenze einer 2-Zelle.

Fig. B.3 veranschaulicht die topologische Beschreibung einer Karte. Es gibt drei 2-Zellen, sechs 1-Zellen und vier 0-Zellen.

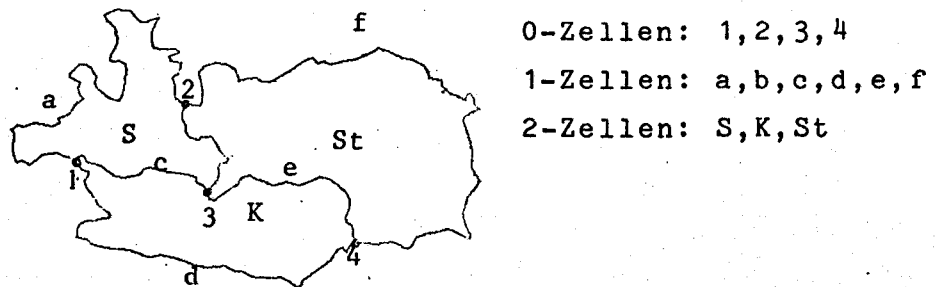


Fig. B.3

Durch diese Darstellung wird es möglich, die Struktur der Karte, z.B. die Beziehungen von Gebieten zueinander, in einem Computer darzustellen, wenn bestimmte Anforderungen erfüllt sind. Diese seien hier nur vage angedeutet:

- F1. Jede 1-Zelle inzidiert mit genau zwei 2-Zellen.
- F2. Bei jeder 0-Zelle gibt es einen eindeutigen "Schirm", d.h. eine zyklisch abwechselnde Kette von 1-Zellen und 2-Zellen.

Sind diese zwei Punkte erfüllt, dann spricht man von einer Oberfläche (Fig. B.4).

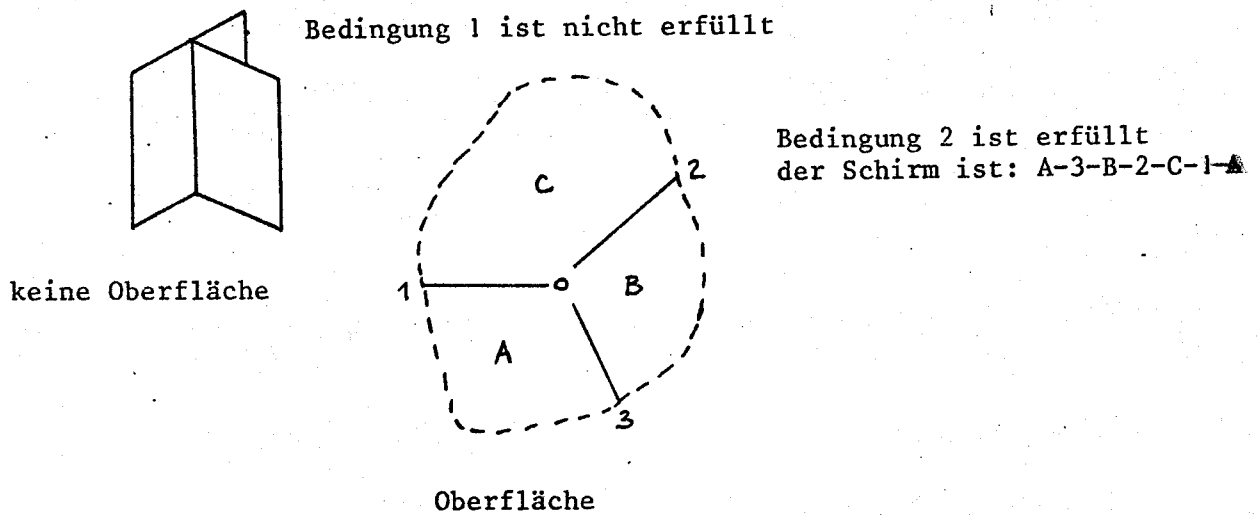


Fig. B.4

Corbett (1979) erweiterte die oben beschriebene topologische Theorie, um Singularitäten miteinschließen zu können. Er verwendet dazu Singulartransformationen anstelle von Homöomorphismen (Fig. B.5).

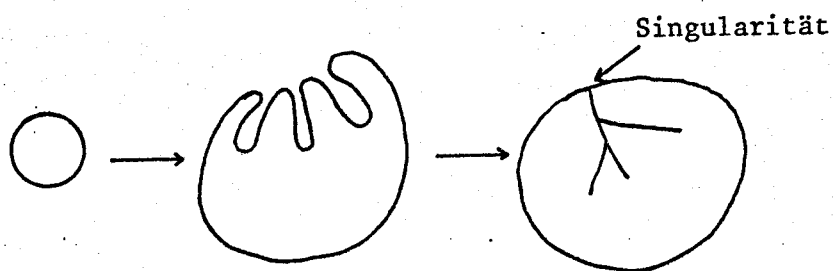


Fig B.5

Erzeugung eines Baumes durch eine
Singulartransformation

Eine Singularität tritt immer dann auf, wenn getrennte Punkte zu einem verschmelzen und so die Bedingungen für einen Homöomorphismus (Stetigkeit) verletzen.

Corbett verallgemeinerte auch 2-Zellen zu 2-Zellen mit Löchern, um Gebilde wie Seen oder Inseln in Seen behandeln zu können, ohne künstliche Schnitte vom Rand des Gebildes zum Loch machen zu müssen. Diese Erweiterung gelang mit Hilfe der Homologietheorie.

B.3 Graphentheorie

Neben der Topologie spielt auch die Graphentheorie eine wesentliche Rolle in der Kartographie. Diese Theorie findet viele Anwendungen bei Flußproblemen (z.B. bestimmen eines kürzesten Weges).

Die Dualität zwischen Graphentheorie und Topologie findet ihre Entsprechung in der Tatsache, daß die Bedingungen F1, F2 der Topologie ziemlich genau den Bedingungen für einen planaren Graphen entsprechen (Planarität ist etwas stärker). Das führt dazu, daß viele topologische Tests mit Hilfe graphentheoretischer Methoden durchgeführt werden können.

B.4 Verbandstheorie

Eine weitere Theorie, die in der computerunterstützten Bearbeitung von Karten Anwendung findet, ist die Verbandstheorie. Damit lassen sich Strukturen, in denen die Hierarchie (z.B. das Enthaltensein einer Gemeinde in einem Bezirk) gestört ist, mittels partieller Ordnungsstrukturen (Halbordnungen) beschreiben.

B.5 Schluß

Der Grund für die Entwicklung und Anwendung dieser Theorien ist ein Informationssystem zur Unterstützung der automatisierten Kartographie zu erhalten. Ein solches Informationssystem sollte alle Fragen beantworten können, die man selbst beim Studium einer Karte oder eines Kartensatzes beantworten könnte.

Anhang C

Polynomische Mehrparametertransformation

Eine Methode, um die beim Digitalisieren erhaltenen Tischkoordinaten in Benutzerkoordinaten (z.B. geographische Länge und Breite) umzurechnen, ist die mehrparametrische Polynomtransformation. Wir besprechen hier eine 12-Parametertransformation über mindestens 6 Passpunkte.

Sei $\tilde{p}_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix}$ ein Punkt in Tischkoordinaten, so soll

der transformierte Punkt die Benutzerkoordinaten

$$\tilde{p}_b = \begin{pmatrix} x_b \\ y_b \end{pmatrix} = A \tilde{p}_t = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 \\ b_0 & b_1 & b_2 & b_3 & b_4 & b_5 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x_t \\ y_t \\ x_t^2 \\ y_t^2 \end{pmatrix}$$

haben.

Die Matrix A der a_i und b_i , $i = 0, \dots, 5$, nennt man Transformationsmatrix.

Die Elemente der Matrix werden mittels linearer Regression über die gemessenen x_t - und y_t - Werte als Faktoren und die gegebenen x_b - und y_b -Werte errechnet.

Sei $\tilde{w}_b = (x_1, x_2, \dots, x_i)^T$, $i > 5$, der Vektor aller x_b -Werte (x-Koordinaten der gegebenen Passpunkte), und sei $\tilde{w}_t = (x_1, x_2, \dots, x_i)$ der Vektor aller x_t -Werte (x-Koordinaten der gemessenen Passpunkte), dann hat man das lineare Modell

$$\tilde{w}_b = C \tilde{a} + \tilde{\varepsilon},$$

wobei C die Designmatrix genannt wird,

$$C = \begin{pmatrix} 1 & x_1 & y_1 & x_1 y_1 & x_1^2 & y_1^2 \\ 1 & x_2 & y_2 & x_2 y_2 & x_2^2 & y_2^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_i & y_i & x_i y_i & x_i^2 & y_i^2 \end{pmatrix}, \quad i > 5, \quad x_i \text{ aus } \tilde{w}_t$$

$\tilde{a} = (a_0, a_1, a_2, a_3, a_4, a_5)^T$ sind die gesuchten Parameter und $\tilde{\varepsilon}$ ist ein (normalverteilter) Fehlervektor.

Die a_i berechnen sich aus

$$\tilde{a} = (C^T C)^{-1} C^T \tilde{w}_b$$

Analog dazu bestimmt man die Parameter \tilde{b} aus den y-Koordinaten.

Mit Hilfe statistischer Kenngrößen (mittlerer quadratischer Fehler der Residuen, oder Bestimmtheitsgrad) entscheidet man über die Güte der Anpassung.

LITETRATURVERZEICHNIS

- ANDERSON K.D. (1980): A Review of Digitizer Technology,
in Havard Library of Computer Graphics,
1980 Mapping Collection, Vol. 9.
- AUTOMETRIC (1978): MOSS Map Overlay and Statistical System
Autometric, Inc., Falls Church, Virginia
- AUTOMETRIC (1979): The Wetlands Analytical Mapping System
(WAMS) User's Manual.
Autometric, Inc., Falls Church, Virginia
- BAHRENBERG G., GIESE E. (1975): Statistische Methoden und
ihre Anwendung in der Geographie
Teubner Studienbücher Geographie, Stuttgart, 1975
- CODD E.F. (1970): A Relational Model of Data for Large
Storage Data Banks.
Communications of the ACM, Vol. 13, June 1970,
S. 377 - 387.
- CORBETT James P. (1979): Topological Principles in Carto-
graphy.
U.S. Department of Commerce, BUREAU OF THE CENSUS,
Technical Paper 48. 1979
- DANGERMOND J., HARDISON L., SMITH L.K. (1980): Some Trends
in the Evolution of GIS Technology.
Environmental Systems Research Institute, Redlands,
California
- DANGERMOND J. (1980): Selecting New Town Sites in the
United States Using Regional Data Bases.
Environmental Systems Research Institute, Redlands,
California

- DAVIS J.C., Mc CULLAGH M.J. (Eds.) (1975): Display and Analysis of Spatial Data. Wiley, New York, 1975.
- DIME (1974): Geographic Base (DIME) File System: A Forward Look. Conference Proceedings, April 16-17, Boston, Massachusetts, zu beziehen von U.S. Gov. Printing Office, Washington, D.C., 20402. Ref. Series GE 60 No.5
- HARVARD (1979): Harvard Library of Computer Graphics, 1979 Mapping Collection, Vol. 1 - Vol. 6. Harvard University. 1979
- HARVARD (1980): Lab-LOG 1980. Laboratory for Computer Graphics and Spatial Analysis, Harvard University.
- HYDROLOGICAL ENGINEERING CENTER (1978): Guide for the Creation of Grid Cell Data Bases. Davis, California, 1978
- JOURNEL A.G., HUIJBREGTS C.J. (1978): Mining Geostatistics Academic Press, London, 1978
- KNAPP E. (1978): LANDSAT and Ancillary Data Inputs to an Automated Geographic Information System: Applications for Urbanized Area Delineation. Computer Science Corporation. 1978
- KNUTH D.E. (1973): The Art of Computer Programming. Vol. III, Sorting and Searching. Reading, Mass., 1973

- KROPATSCH, W. und LEBERL, F. (1981): Organisation karto-
graphischer Daten zur kenntnisgestützten Bildanal~~y~~
Informatik - Fachberichte 49, Modelle und Struktur~~e~~
DAGM-Symposium, Hamburg, Oktober 1981;
Springer Verlag, Berlin - Heidelberg, New York
- LEBERL, F. und OLSON, D. (1981): Raster Scanning for
Operational Digitizing of Graphical Data
- MAURER H: (1974): Datenstrukturen und Programmierverfahren
Teubner Studienbücher Informatik, Stuttgart, 1974
- MINNESOTA (1979): Minnesota Cropland Resources.
North Shore Data Atlas, Minnesota Coastal Zone.
Minnesota State Planning Agency, Environmental
Planning, St. Paul, Minnesota. 1979
- MITCHELL W.B. (1979): GIRAS: A Geographic Information
Retrieval and Analysis System for Handling Land
Use and Land Cover Data.
U.S. Geological Survey Professional Paper 1059,
USA. 1979
- MOLDS (1975): Proceedings of the North American Conference
on Modernization of Land Data Systems
(A Multipurpose Approach).
Washington, D.C. 1975
- MOLDS (1978): Land Data Systems Now.
Proceedings of the Second MOLDS Conference (Imple-
mentation of a Modern Multipurpose Land Data System).
Washington, D.C. 1978

- MOORE P.A. (Ed.) (1980): Harvard Library of Computer Graphics, 1980 Mapping Collection, Vol. 7 - Vol. 11. Harvard University, 1980
- MOYER D. David (1978): Land Information Systems. An Annotated Bibliography MOLDS, 1978
- NAGY G., WAGLE S: (1979): Geographic Data Processing. Computing Surveys. Vol. 11, No. 2, June 1979
- NAGY G. (1981): Criteria for Selecting Automatic Digitizers (Optical Scanners)
- NEWMAN W.M., SPROULL R.F. (1979): Principles of Interactive Computer Graphics. Mc Graw Hill New York, 1979
- NISEN W.G. (1980): A Brief Tutorial on the Selection of Digitizing Equipment for Computer Production. In Harvard Library of Computer Graphics, 1980 Mapping Collection, Vol. 9.
- ORR J. (1978): Mini-Micro Systems. Computer Graphics, February 1978.
- REED C. (1980): Habitat Characterization: A Production Mode Usage of a Geographic Information System. Fort Collins, 1980
- RIPLEY B.D. (1981): Spatial Statistics Wiley. New York, 1981
- ROSENFELD A., DYER C.R., SAMET H. (1980): Region Representation: Boundary Codes from Quadrees. Communications of the ACM, March 1980, Vol. 23, Number 3.

- SAMET H. (1981): An Algorithm for Converting Rasters to
Quadtrees.
IEEE Transactions on Pattern Analysis and Machine
Intelligence, Vol. PAMI-3, No. 1, January 1981.
- SHEEHAN, D.E. (1979): A Discussion of the SYMAP Programm.
Harvard Library of Computer Graphics, 1979 Mapping
Collection, Vol. 2
- TOMLINSON R.F., CALKINS H.W., MARBLE D.F. (1976): Computer
Handling of Geographical Data.
The Unesco Press, 1976
- TUREK J.E., WALKER D.J. (1979): Large Format Laser Scanner/
Plotter System.
AUTOCARTO IV, Proceedings of the International
Symposium on Cartography and Computing:
Applications in Health and Environment. Volume II.
Published by AMERICAN CONGRESS ON SURVEYING AND
MAPPING, AMERICAN SOCIETY OF PHOTOGRAMMETRY. 1980
- URISA (1975): Papers from the Thirteenth Annual Conference
of the Urban and Regional Information Systems
Association:
Computers, Local Government and Productivity, Vol. I,
II. Chicago, 1976
- URISA (1977): Papers from the Fifteenth Annual Conference
of the Urban and Regional Information Systems
Association: Information System Inputs to Policies,
Plans, and Programs
Chicago, 1977

URISA (1978): Papers from the Sixteenth Annual Conference of the Urban and Regional Information Systems Association: Data Resources and Requirements: Federal and Local Perspectives. Chicago, 1978

URISA (1979): Papers from the Annual Conference of the Urban and Regional Information Systems Association: More for Less: Information Systems in an Era of Limits. Chicago, 1979

URISA (1980): Papers from the Annual Conference of the Urban and Regional Information Systems Association: Ends and Means of Urban and Regional Information Systems: Bridging the Gap in the 1980's. 1980

VOXLAND P.M. (1981): WORLD Projection and Mapping Program Reference Manual Ver. 3.43. University of Minnesota, Minneapolis, Minnesota. 1981

WEBER W. (1978): Geographische Informationssysteme - ein Überblick und Gedanken zur weiteren Entwicklung. Nachrichten aus dem Karten - und Vermessungswesen, Heft 75. Verlag des Instituts für Angewandte Geodäsie, Frankfurt am Main, 1978

WHITE Marvin (1979): A Survey of the Mathematics of Maps. AUTOCARTO IV, Proceedings of the International Symposium on Cartography and Computing: Applications in Health and Environment. Volume I. Published by AMERICAN CONGRESS ON SURVEYING AND MAPPING, AMERICAN SOCIETY OF PHOTOGRAMMETRY, 1980

WHITE Marvin S., GRIFFIN Patricia E. (1979): Coordinate
Free Cartography AUTOCARTO IV, Proceedings of the
International Symposium on Cartography and Computing:
Applications in Health and Environment. Volume I.
Published by AMERICAN CONGRESS ON SURVEYING AND
MAPPING, AMERICAN SOCIETY OF PHOTOGRAMMETRY. 1980

