Heft 10
München, Januar 1984

10 Jahre
Hochschule der Bundeswehr München

Beiträge aus dem Institut für Geodäsie

Herausgegeben von
W. Caspary, A. Schödlbauer und W. Welsch

SCHRIFTENREIHE

Wissenschaftlicher Studiengang Vermessungswesen
Hochschule der Bundeswehr München

HSBw

Heft 10
München, Januar 1984

10 Jahre
Hochschule der Bundeswehr München

Beiträge aus dem Institut für Geodäsie

Herausgegeben von
W. Caspary, A. Schödlbauer und W. Welsch

SCHRIFTENREIHE

Wissenschaftlicher Studiengang Vermessungswesen
Hochschule der Bundeswehr München

HSBw

Der Druck dieses Heftes wurde aus Haushaltsmitteln der Hochschule der Bundeswehr München gefördert.

Verantwortlich für die Herausgabe der Schriftenreihe: Prof. Dr. G. Neugebauer
Dipl.Ing. D. Beineke

Bezugsnachweis:

Studiengang Vermessungswesen
im Fachbereich Bauingenieur- und Vermessungswesen
Hochschule der Bundeswehr München
Werner-Heisenberg-Weg 39
8014 Neubiberg

ISSN 0173-1009

VORWORT

Vor 10 Jahren, im Oktober 1973, nahmen die ersten Studenten das Studium an der Hochschule der Bundeswehr München auf. Vorausgegangen war eine äußerst knapp bemessene Gründungsphase, und es folgte eine Aufbauphase, die durch Provisorien, durch Experimente mit Studien- und Prüfungsordnungen, durch das Ringen um öffentliche Anerkennung und durch Auseinandersetzungen um eine sich langsam festigende Struktur der Hochschule gekennzeichnet war. Diese Aufbauphase ist heute weitgehend abgeschlossen.

Der bayerische Kultusminister hat der Hochschule mittlerweile das Promotions- und das Habilitationsrecht verliehen und alle Examina als gleichwertig mit denen der öffentlichen Universitäten anerkannt.

Die Forschung ist in Gang gekommen und hängt wie überall in starkem Maße von der Drittmittelförderung ab. Die personelle und instrumentelle Ausstattung für Forschung und Lehre, die zunächst zügig nach Plan aufgebaut wurde, unterliegt freilich inzwischen den an allen Universitäten zu beklagenden Kürzungen und Einschränkungen. Die neuen Gebäude in Neubiberg erweisen sich schon als zu sparsam geplant. Und die Zahl der Studenten übersteigt das vorgesehene Maß. Die Hochschule der Bundeswehr hat sich damit zu einer ganz normalen deutschen Universität entwickelt.

Zentren der Lehre und Forschung sind die Institute, in denen fachlich verwandte Lehrstühle zusammengefaßt sind. Das Institut für Geodäsie, das wie die Hochschule nunmehr auf eine zehnjährige, arbeitsreiche Aufbauphase zurückblicken kann, vereint in sich die Lehrstühle für

- Allgemeine Geodäsie
- Ingenieurgeodäsie
- Vermessungskunde

und verfügt über ein modern ausgestattetes Geodätisches Labor.

Die Professoren und Wissenschaftlichen Mitarbeiter des Instituts für Geodäsie haben das Jubiläum zum Anlaß genommen, ein Heft der Schriftenreihe des Wissenschaftlichen Studiengangs Vermessungswesen zu gestalten. Das Heft enthält Einzelbeiträge aus allen drei Lehrstühlen des Instituts. Die Beiträge lassen sich in die Themenkreise

- Mathematische Grundlagen
- Netzanalyse und -optimierung
- Landesvermessung
- Instrumentenkunde

gliedern und mögen dem geschätzten Leser in der Art einer Momentaufnahme einen Einblick in zur Zeit laufende Arbeiten des Instituts vermitteln.

Neubiberg, im Oktober 1983

Wilhelm Caspary
Albert Schödlbauer
Walter Welsch

INHALTSVERZEICHNIS

Seite

MATHEMATISCHE GRUNDLAGEN

Der Rang einer Summe positiv semidefiniter Matrizen und ein Beispiel zur Anwendung in der Geodäsie
K.A.W. Bode, E.H. Knickmeyer 5

Parameterschätzung in linearen Modellen mit Hilfe von Projektoren
Wilhelm Caspary 25

Umkehrung von vollständigen Potenzreihen mit zwei Veränderlichen
Hans Glasmacher und Klaus Krack 49

Zur Bestimmung der signifikanten Parameter in Approximationsfunktionen
Peter Schwintzer 71

NETZANALYSE UND -OPTIMIERUNG

Das Programmsystem NOPTI II zum Entwurf des optimalen Beobachtungsplanes geodätischer Lagenetze nach invarianten Kriterien
R. Baumer und H. Heister 93

Ein Konzept zur interaktiven Planung, Ausgleichung und Analyse geodätischer Netze
Rolf König 119

Invarianzeigenschaften von Strainparametern
Walter Welsch 129

Seite

LANDESVERMESSUNG

Bivariate Polynome zur genäherten Bestimmung von UTM-Koordinaten (Europa-Datum 1950) aus Gauß-Krüger-Koordinaten (Potsdam-Datum)
Klaus Krack und Albert Schödlbauer 151

Strenge, nicht differentielle Lösung für den Ellipsoid-übergang und die Koordinaten-Transformation nach Helmert im System der geographischen Koordinaten
Albert Schödlbauer 165

INSTRUMENTENKUNDE

Untersuchung und Einsatz von potentiometrischen Wegaufnehmern
Wilfried Ellmer 183

Aufbau und Funktionsweise von Kompensatoren in optischen Sekundentheodoliten und elektronischen Theodoliten
Axel Geiger 203

Einige moderne Verfahren zur Kalibrierung von Nivellierlatten
Roland Tille 221

DER RANG EINER SUMME POSITIV SEMIDEFINITER MATRIZEN UND EIN BEISPIEL ZUR ANWENDUNG IN DER GEODÄSIE

von K.A.W. BODE, E.H. KNICKMEYER

SUMMARY

In order to determine the same unknowns according to the Gauß-Markoff-model there exist various groups of observables. If these groups are uncorrelated with one another, then the matrix of normal equations for a total adjustment is the sum of the normal equation matrices of the individual groups. The theorem is proved, that such a sum of positive semi-definite matrices essentially has the rank of the union of the columns or of the rows of the individual terms of the sum. Different versions of factorization of the terms of the sum are included. The theorem finds an application in an analytically represented example of a three point-levelling-network, observed at two epochs.

ZUSAMMENFASSUNG

Liegen zur Bestimmung von gleichen Unbekannten entsprechend dem Gauß-Markoff-Modell verschiedene Beobachtungsgruppen vor, die untereinander unkorreliert sind, dann ist die Normalgleichungsmatrix einer Gesamtausgleichung die Summe der Normalgleichungsmatrizen der einzelnen Gruppen. Es wird bewiesen, daß der Rang einer solchen Summe positiv semidefiniter Matrizen im wesentlichen gleich dem Rang der Vereinigung der Spalten oder der Zeilen der einzelnen Summanden ist. Verschiedene Möglichkeiten der Faktorisierung der Summanden sind mit bearbeitet. Eine Anwendung mit analytischer Darstelllung findet die mathematische Aussage im Beispiel eines zu zwei Epochen beobachteten Dreipunkt-Höhennetzes.

1. EINFÜHRUNG

Bei der Deformationsanalyse geodätischer Netze und im Prinzip auch bei anderen geodätischen sowie nichtgeodätischen Aufgaben hat man folgende Situation. Zu zeitlich getrennten Epochen liegen Beobachtungsgruppen mit nicht notwendig gleichartigen Beobachtungen bei gleichen oder formal gleichen Unbekannten vor. Die Unbekannten sind einmal für die einzelnen Beobachtungsgruppen zu bestimmen, andererseits für ein Gesamtmodell, bestehend aus der Zusammenfassung aller Beobachtungsgruppen.
Grundlage der Auswertung ist jeweils ein Gauß-Markoff-Modell, bestehend aus einem funktionalen und einem stochastischen Modell. Die Unbekannten sind somit durch die Methode der kleinsten Quadrate bestimmt und ergeben sich aus den Normalgleichungen. Bei der Deformationsanalyse schließen sich statistische Hypothesentests an.

Sowohl in den funktionalen als auch in den stochastischen Modellen sind im allgemeinen Rangdefekte zu erwarten. Diese übertragen sich auf die Normalgleichungsmatrizen. Bei der Deformationsanalyse gehen sie zudem in die Prüfgrößen der Hypothesentests ein.

Um aus den Normalgleichungen der einzelnen Epochen und des Gesamtmodells eindeutige Lösungen zu erhalten, sind die Rangdefekte durch Bedingungen zu beheben. Dazu ist der Rang der Normalgleichungsmatrizen festzustellen und die Rangdefekte sind auf Translations-, Konfigurations- und andere Defekte hin zu analysieren. Erst dann ist es möglich, einer eindeutigen Lösung bestimmte, frei festlegbare Eigenschaften zuzuordnen.
Zur Aufstellung geeigneter Bedingungen ist die Normalgleichungsmatrix des Gesamtmodells oft von besonderem Interesse. Denn in ihr verbleiben die gleichartigen Rangdefekte der verschiedenen Epochen. Sie sind für jede Epoche und für das Gesamtmodell durch gleiche Bedingungen zu beheben, um die Vergleichbarkeit der Lösungen der verschiedenen Epochen untereinander und mit der des Gesamtmodells zu gewährleisten. Diese gemeinsamen Bedingungen lassen sich oft erst aufstellen, nachdem der Rang der Normalgleichungsmatrix des Gesamtmodells bestimmt wurde und die verbleibenden Rangdefekte analysiert wurden.
Die Normalgleichungsmatrix des Gesamtmodells ist, bei Voraussetzung der Unkorreliertheit von Beobachtungen verschiedener Epochen, die Summe der Normalgleichungsmatrizen der einzelnen Epochen.

Gegenstand der vorliegenden Arbeit
ist die Rangbestimmung für eine Summe von Normalgleichungsmatrizen bzw. von positiv semidefiniten Matrizen. Aussagen und Beweise werden für positiv semidefinite Matrizen formuliert, was nur formal allgemeiner als für Normalgleichungsmatrizen ist, da sich positiv semidefinite Matrizen als Normalgleichungsmatzrizen darstellen lassen.
Bewiesen wird eine mathematische (algebraische) Grundlage der Rangbestimmung, die im wesentlichen besagt, daß der Spalten- bzw. Zeilenrang einer Summe positiv semidefiniter Matrizen gleich dem Rang der Vereinigung der Spalten bzw. Zeilen der einzelnen Summanden ist. Diese Aussage wird in verschiedenen Versionen bearbeitet, insbesondere für den Fall, daß die Summanden in der Faktorisierung von Normalgleichungsmatrizen vorliegen.
Im weiteren folgt ein Beispiel zur Anwendung, für das ein in zwei Epochen beobachtetes Höhennetz, bestehend aus drei Punkten, mit verschiedenartigen Beobachtungen zu jeder Epoche, ausgewählt wurde. Das Beispiel ist einer rein analytischen Darstellung noch zugänglich und wird in dieser Weise ausgeführt.

2. GRUNDLAGEN AUS DER LINEAREN ALGEBRA UND NOTATIONEN

Grundmenge ist jeweils ein reeller Euklidischer Vektorraum E^n der endlichen Dimension n. Als ***Skalarprodukt*** wird $(\bar{x},\bar{y}) := \bar{x}^T\bar{y}$, $\bar{x},\bar{y} \in E^n$ ($\bar{x},\bar{y}$ aus E^n) verwendet. Mit $...^T$ wird die Transposition von Vektoren und Matrizen bezeichnet. ***Lineare Abbildungen*** von einem E^n in einen E^m sind durch Matrizen repräsentiert.

Für eine Teilmenge $M \subset E^n$ von Vektoren sei $\operatorname{lin} M$ die ***lineare Hülle*** (Menge aller Linearkombinationen der Vektoren aus M). $\operatorname{lin} M$ ist Unterraum und somit Vektorraum. Die ***Dimension*** eines Vektorraums U, $\dim U$, ist die Maximalanzahl der Vektoren in linear unabhängigen Teilmengen oder die Anzahl der Vektoren einer ***Basis*** in U.
Orthogonales Komplement einer Teilmenge $M \subset E^n$ ist $M^\perp := \{\bar{y} \in E^n : \bar{y}^T\bar{x} = 0 \text{ für alle } \bar{x} \in M\} = \bigcap_{\bar{x}\in M} \{\bar{y} \in E^n : \bar{y}^T\bar{x} = 0\}$. $M^{\perp\perp}$ ist das orthogonale Komplement $(M^\perp)^\perp$ von $M^\perp$ oder die ***biorthogonale Hülle*** von M. $M^\perp$ ist ein Unterraum, der ***Orthogonalraum*** zu $\operatorname{lin} M$, und im E^n gilt $M^{\perp\perp} = \operatorname{lin} M$ [***KOCH*** 1980 Kap. 1.2.4, ***KOWALSKY*** 1979 § 20].

Für eine Matrix $\mathcal{A}$ bezeichne $s(\mathcal{A})$ bzw. $z(\mathcal{A})$ die Menge der Spaltenvektoren bzw. Zeilenvektoren von $\mathcal{A}$. Der ***Kern*** einer Matrix $\mathcal{A}$ ist das orthogonale Komplement der Zeilen $\mathcal{A}$, Kern $\mathcal{A} := \{\bar{x} : \mathcal{A}\bar{x} = \bar{0}\} = (z(\mathcal{A}))^{\perp}$.
Der ***Spaltenrang*** dim lin $s(\mathcal{A})$ und der ***Zeilenrang*** dim lin $z(\mathcal{A})$ einer Matrix $\mathcal{A}$ sind gleich, so daß der ***Rang*** einer Matrix $\mathcal{A}$ durch $rg\,\mathcal{A}$:= dim lin $s(\mathcal{A})$ = dim lin $z(\mathcal{A})$ definiert wird. Für jede $(m \times n)$-Matrix $\mathcal{A}$ gilt $rg\,\mathcal{A} \le \min\{m,n\}$. $\mathcal{A}$ hat ***vollen Rang***, wenn $rg\,\mathcal{A} \le \min\{m,n\}$ gilt. Eine quadratische Matrix mit vollem Rang heiße ***nichtsingulär***.

Der Rang einer Matrix ändert sich nicht unter ***elementaren Umformungen***, das sind:

(1) Vertauschen zweier Zeilen oder Spalten.
(2) Multiplikation einer Zeile oder Spalte mit einer reellen Zahl $\neq 0$.
(3) Addition eines reellen Vielfachen einer Zeile oder Spalte auf eine andere Zeile bzw. Spalte.

Eine Zusammenfassung elementarer Umformungen, die den Rang einer Matrix ebenfalls nicht ändert, ist die Multiplikation mit einer nichtsingulären (quadratischen) Matrix. Allgemein gilt für den Rang des Produkts zweier Matrizen $\mathcal{A}_1, \mathcal{A}_2$: $rg(\mathcal{A}_i) \ge rg(\mathcal{A}_1\mathcal{A}_2)$.
[***Koch*** 1980 Kap. 1.3.2, ***KOWALSKY*** 1979 § 7.3, ***ZURMÜHL*** 1950 § 7.2, 7.3]

Eine symmetrische $(n \times n)$-Matrix $\mathcal{A}$, $\mathcal{A}^T = \mathcal{A}$, heißt ***positiv semidefinit*** bzw. ***positiv definit***, wenn für alle $\bar{x} \in E^n$ gilt $\bar{x}^T\mathcal{A}\bar{x} \ge 0$ bzw. $\bar{x}^T\mathcal{A}\bar{x} > 0$.
Die Summe zweier positiv semidefiniter Matrizen $\mathcal{A}_1, \mathcal{A}_2$ ist wieder positiv semidefinit, $\bar{x}^T(\mathcal{A}_1+\mathcal{A}_2)\bar{x} = \bar{x}^T\mathcal{A}_1\bar{x} + \bar{x}^T\mathcal{A}_2\bar{x} \ge 0+0$ für alle $\bar{x} \in E^n$. Eine positiv semidefinite Matrix ist genau dann positiv definit, wenn sie nichtsingulär ist. Normalgleichungsmatrizen $\mathcal{A}^T\mathcal{A}$ oder $\mathcal{A}^T\mathcal{P}\mathcal{A}$, wobei die Jacobimatrix $\mathcal{A}$ reel und die Gewichtsmatrix $\mathcal{P}$ positiv semidefinit ist, sind positiv semidefinit.

Die ***Pseudoinverse*** oder ***Moore-Penrose-Inverse*** $\mathcal{A}^+$ einer Matrix $\mathcal{A}$ ist durch die 4 ***Moore-Penrose-Axiome*** eindeutig bestimmt:

(1) $\mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A}$ (Rechtsinverse), (2) $\mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+$ (Linksinverse),

(3) $(\mathcal{A}^+\mathcal{A})^T = \mathcal{A}^+\mathcal{A}$, (4) $(\mathcal{A}\mathcal{A}^+)^T = \mathcal{A}\mathcal{A}^+$ (Symmetrien).

Liegt für eine $(m \times n)$-Matrix $\mathcal{A}$ eine 'Rangfaktorisierung' $\mathcal{A} = \mathcal{F}\mathcal{G}^T$ mit

einer $(m \times r)$-Matrix $\mathcal{F}$ und einer $(n \times r)$-Matrix $\mathcal{G}$, $r = rg\,\mathcal{A}$, vor, so gilt $\mathcal{A}^{+} = \mathcal{G}\left(\mathcal{G}^{T}\mathcal{G}\right)^{-1}\left(\mathcal{F}^{T}\mathcal{F}\right)^{-1}\mathcal{F}^{T}$.
[***BJERHAMMAR*** 1973 chap. 9, ***KOCH*** 1980 Kap. 1.5.3, ***RAO-MITRA*** 1971 chap. 3.3].

3. EINE ALGEBRAISCHE AUSSAGE ZUR BESTIMMUNG DES RANGES EINER SUMME VON POSITIV SEMIDEFINITEN MATRIZEN

Die Aussage, daß der Rang einer Summe positiv semidefiniter Matrizen gleich dem Rang der Vereinigung der Spalten bzw. der Zeilen der Summanden ist, wird in verschiedenen Versionen dargestellt. Die Hauptversion bezieht sich auf die Faktorisierbarkeit von positiv semidefiniten Summanden $\mathcal{B}$ in Form von Normalgleichungsmatrizen, $\mathcal{B} = \mathcal{A}^{T}\mathcal{P}\mathcal{A}$, mit reellem $\mathcal{A}$ und positiv semidefinitem $\mathcal{P}$. Die Existenz solcher und ähnlicher ***Faktorisierungen*** ist für positiv semidefinite Matrizen durch den folgenden, aus der linearen Algebra bekannten Satz gesichert.

1. Satz

Jede reelle, symmetrische $(n \times n)$-Matrix $\mathcal{B}$ besitzt ein System von n orthogonalen Eigenvektoren $\bar{x}_i$,

$$\mathcal{B}\,\bar{x}_i = \lambda_i\,\bar{x}_i \quad \text{und} \quad \bar{x}_i^{T}\,\bar{x}_j = \delta_{ij}\,, \tag{3-1}$$

δ_{ij} das Kronecker-Symbol. Dabei können bei verschiedenen Eigenvektoren gleiche Eigenwerte auftreten, d.h. Eigenwerte können 'mehrfach' auftreten.

Beweis siehe z.B. [***KOCH*** 1980 Kap. 1.4.2, ***ZURMÜHL*** 1950 § 13.4, 15.1, 15.3].

Faßt man die Eigenvektoren zur ***Modalmatrix*** $\mathcal{X} := [\bar{x}_1, \ldots, \bar{x}_n]$ und die zugehörigen Eigenwerte in der gleichen Reihenfolge zur Diagonalmatrix $\Lambda := \mathrm{diag}(\lambda_i)$ zusammen, so erhält die Aussage von Satz 1. die Form:

$$\mathcal{B}\,\mathcal{X} = \mathcal{X}\,\Lambda\,, \text{ mit der } \textbf{\textit{orthogonalen Matrix}}\ \mathcal{X} : \mathcal{X}^{T}\mathcal{X} = \mathcal{X}\,\mathcal{X}^{T} = \mathcal{I}\,.$$

Damit besitzt eine reelle symmetrische Matrix $\mathcal{B}$ die Faktorisierung $\mathcal{B} = \mathcal{X}\,\Lambda\,\mathcal{X}^{T}$ [***KOWALSKY*** 1979 § 23]. Weiter gilt nach (3-1) für eine symmetrische Matrix $\mathcal{B}$:

$$\bar{x}_i^{T}\mathcal{B}\,\bar{x}_i = \lambda_i\bar{x}_i^{T}\bar{x}_i = \lambda_i\delta_{ii} = \lambda_i$$

Ist nun $\mathcal{B}$ positiv semidefinit, so folgt $\lambda_i \geq 0$. Also ist die Diagonalmatrix Λ positiv semidefinit und die Diagonalmatrix $\Lambda^{1/2} = \mathrm{diag}\left(\lambda_i^{1/2}\right)$ ist reell. Folglich hat $\mathcal{B}$ die Faktosiriserungen:

$$\mathcal{B} = \mathcal{X}\,\Lambda\,\mathcal{X}^T = \mathcal{A}^T\mathcal{P}\mathcal{A} \text{ , mit } \mathcal{A} := \mathcal{X}^T \text{ , } \mathcal{P} := \Lambda \text{ .}$$

$$\mathcal{B} = \mathcal{X}\,\Lambda\,\mathcal{X}^T = \mathcal{X}\,\Lambda^{1/2}\,\Lambda^{1/2}\,\mathcal{X}^T = \mathcal{X}\,\Lambda^{1/2}\left(\mathcal{X}\Lambda^{1/2}\right)^T = \mathcal{C}^T\mathcal{C} \text{ ,}$$

mit reellem $\mathcal{C} = \left(\mathcal{X}\,\Lambda^{1/2}\right)^T$. Damit ist gezeigt:

1.a Korollar

Jede positiv semidefinite Matrix $\mathcal{B}$ läßt sich in Form einer 'Normalgleichungsmatrix' faktorisieren:

a. $\mathcal{B} = \mathcal{A}^T\mathcal{P}\mathcal{A}$, wobei $\mathcal{A} := \mathcal{X}^T$, $\mathcal{P} := \Lambda$ gewählt werden kann.

b. $\mathcal{B} = \mathcal{C}^T\mathcal{C}$, wobei $\mathcal{C} := \left(\mathcal{X}\,\Lambda^{1/2}\right)^T$ gewählt werden kann.

Entsprechend der Problemstellung, die die vorliegende Arbeit motiviert, beschränken wir uns auf den Fall symmetrischer Summanden $\mathcal{B}$ mit reellem $\mathcal{A}$ und positiv semidefinitem $\mathcal{P}$ bzw. mit reellem $\mathcal{C}$ als Faktoren, der genau für positiv semidefinite Matrizen $\mathcal{B}$ gegeben ist. Der im weiteren ausgeführte Beweis stützt sich zunächst auf die Faktorisierung der Summanden $\mathcal{B}$ in der Form $\mathcal{B} = \mathcal{C}^T\mathcal{C}$.

Beide Formen der Faktorisierung sind nicht eindeutig. Das soll für die Faktorisierung $\mathcal{B} = \mathcal{C}^T\mathcal{C}$ verdeutlicht werden.
Die Nichteindeutigkeit von $\mathcal{C}$ ist bereits daraus ersichtlich, daß $\mathcal{C} = \left(\mathcal{X}\,\Lambda^{1/2}\right)^T$ nach dem Streichen von (n-r) Nullzeilen $(r = rg\,\mathcal{B})$ im allgemeinen eine vollbesetzte $(r \times n)$-Matrix ist. Dagegen liefert die Cholesky-Zerlegung $\mathcal{B}$ eine $(r \times n)$-Matrix $\mathcal{C}$ von oberer Dreiecksform.
Selbst die Anzahl der Zeilen von $\mathcal{C}$ ist nicht eindeutig, sie kann jede ganze Zahl $m \geq r$ annehmen. Zur Erläuterung seien einige Faktoren $\mathcal{C}$ der um den Faktor $p_{\Delta H1}$ gekürzten Normalgleichungsmatrix $\mathcal{N}_1$ aus dem Beispiel in Abschnitt 4. angegeben.

$$\mathcal{B} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

a. Zweizeiliger Faktor $\mathcal{C}_2$ aus der abbrechenden Cholesky-Zerlegung.

$$\mathcal{C}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -1 & -1 \\ 0 & \sqrt{3} & -\sqrt{3} \end{bmatrix}$$

b. Dreizeiliger Faktor $\mathcal{C}_3$ entsprechend einer nichtorthogonalisierten Eigenvektorzerlegung.

$$\mathcal{Y}_3 := \frac{1}{\sqrt{6}} \begin{bmatrix} 2 & 0 \\ -1 & \sqrt{3} \\ -1 & -\sqrt{3} \end{bmatrix}, \quad \mathcal{C}_3 := \mathcal{Y}_3 \mathcal{C}_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Die Spalten von $\mathcal{C}_3$ sind Eigenvektoren von $\mathcal{B}$ zum Eigenwert 3 (Dimension des zugehörigen Eigenraums = 2).

c. Vierzeiliger Faktor $\mathcal{C}_4$.

$$\mathcal{Y}_4 := \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{2} & 0 \\ -1 & \sqrt{3} \\ -1 & -\sqrt{3} \\ \sqrt{2} & 0 \end{bmatrix}, \quad \mathcal{C}_4 := \mathcal{Y}_4 \mathcal{C}_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -1 & 2 & -1 \\ -1 & -1 & 2 \\ \sqrt{2} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Die Matrizen $\mathcal{Y}_i$ sind orthogonale Matrizen, $\mathcal{Y}_i^T \mathcal{Y}_i = \mathcal{I}$. Analog lassen sich Faktoren mit beliebiger Zeilenzahl i angeben.

Für Faktorisierungen positiv semidefiniter Matrizen $\mathcal{B}$ von der Form $\mathcal{B} = \mathcal{C}^T\mathcal{C}$ ist bekannt, daß der Rang von $\mathcal{B}$ durch den Rang von $\mathcal{C}$ gegeben ist [***KOCH*** 1980 S. 34].

2. Satz

Für jede reelle Matrix $\mathcal{C}$ gelten die äquivalenten Aussagen:

a. $\mathrm{lin\,z}(\mathcal{C}^T\mathcal{C}) = \mathrm{lin\,z}(\mathcal{C})$

b. $\mathrm{lin\,s}(\mathcal{C}^T\mathcal{C}) = \mathrm{lin\,s}(\mathcal{C}^T)$

c. $\mathrm{rg}(\mathcal{C}^T\mathcal{C}) = \mathrm{rg}(\mathcal{C}) = \mathrm{rg}(\mathcal{C}^T)$

Beweis:

Zunächst läßt sich zeigen: $\mathrm{Kern}(\mathcal{C}^T\mathcal{C}) = \mathrm{Kern}(\mathcal{C})$.

1. Sei $\bar{x} \in \mathrm{Kern}(\mathcal{C})$. Aus $\mathcal{C}\bar{x} = \bar{0}$ folgt sofort $\mathcal{C}^{\top}\mathcal{C}\bar{x} = \bar{0}$.
Also folgt $\bar{x} \in \mathrm{Kern}(\mathcal{C}^{\top}\mathcal{C})$.

2. Sei umgekehrt $\bar{x} \in \mathrm{Kern}(\mathcal{C}^{\top}\mathcal{C})$.
Aus $\mathcal{C}^{\top}\mathcal{C}\bar{x} = \bar{0}$ erhält man $(\mathcal{C}\bar{x})^{\top}\mathcal{C}\bar{x} = \bar{x}^{\top}\mathcal{C}^{\top}\mathcal{C}\bar{x} = 0$.
Aufgrund der Definitheit des Skalarprodukts $(\bar{y},\bar{z}) := \bar{y}^{\top}\bar{z}$ auf dem E^n
folgt $\mathcal{C}\bar{x} = \bar{0}$, also $\bar{x} \in \mathrm{Kern}(\mathcal{C})$.

Mit 1. und 2. ist $\mathrm{Kern}(\mathcal{C}^{\top}\mathcal{C}) = \mathrm{Kern}(\mathcal{C})$ gezeigt.
Damit ist zunächst das Äquivalent der Behauptung für die Orthogonalräume der Zeilen bewiesen:

$$\left(z(\mathcal{C}^{\top}\mathcal{C})\right)^{\perp} = \mathrm{Kern}(\mathcal{C}^{\top}\mathcal{C}) = \mathrm{Kern}(\mathcal{C}) = \left(z(\mathcal{C})\right)^{\perp}$$

Da im E^n die biorthogonale Hülle gleich der linearen Hülle ist, folgt a.:

$$\mathrm{lin}\, z(\mathcal{C}^{\top}\mathcal{C}) = \left(z(\mathcal{C}^{\top}\mathcal{C})\right)^{\perp\perp} = \left(z(\mathcal{C})\right)^{\perp\perp} = \mathrm{lin}\, z(\mathcal{C})$$

Beim Transponieren einer Matrix gehen Zeilen in Spalten über und umgekehrt. Also gilt b. äquivalent zu a..
Da Zeilenrang = Spaltenrang = Rang einer Matrix gilt, folgt c. aus a.. Umgekehrt folgt aus c. wegen $\mathrm{lin}\, z(\mathcal{C}^{\top}\mathcal{C}) \subset \mathrm{lin}\, z(\mathcal{C})$ wieder a.. q.e.d.

Betrachtet man jetzt eine Summe $\mathcal{B}_1 + \mathcal{B}_2$ positiv semidefiniter Matrizen $\mathcal{B}_1$, $\mathcal{B}_2$ mit Faktorisierungen $\mathcal{B}_1 = \mathcal{C}_1^{\top}\mathcal{C}_1$, $\mathcal{B}_2 = \mathcal{C}_2^{\top}\mathcal{C}_2$, so läßt sich Satz 2. einerseits auf $\mathcal{B}_1$, $\mathcal{B}_2$ anwenden und andererseits auf die wiederum positiv semidefinite Summe $\mathcal{B}_1 + \mathcal{B}_2$ mit einer zunächst beliebigen Faktorisierung $\mathcal{B}_1 + \mathcal{B}_2 = \mathcal{C}^{\top}\mathcal{C}$. Unter den Faktorisierungen $\mathcal{B}_1 + \mathcal{B}_2 = \mathcal{C}^{\top}\mathcal{C}$ gibt es aber eine, die den Zusammenhang der Ränge von $\mathcal{C}$ und $\mathcal{C}_1$, $\mathcal{C}_2$ unmittelbar erkennen läßt. Diese ergibt sich aus der Darstellbarkeit von Faktorisierungen der Form $\mathcal{B} = \mathcal{C}^{\top}\mathcal{C}$ als Summe von symmetrischen dyadischen Produkten (Hilfssatz 3.). (Dyadisches Produkt siehe z.B. [***ZURMÜHL*** § 2.5].) Satz 2. liefert dann den Zusammenhang der Ränge von $\mathcal{B}_1 + \mathcal{B}_2$ und $\mathcal{B}_1$, $\mathcal{B}_2$ (Satz 4.).

<u>3. Hilfssatz</u>

Jede Faktorisierung der Form $\mathcal{B} = \mathcal{C}^{\top}\mathcal{C}$ läßt sich als Summe symmetrischer ***dyadischer Produkte*** darstellen, $\mathcal{C}^{T}\mathcal{C} = \sum \bar{c}_k \bar{c}_k^{\top}$, $\bar{c}_k^{\top}$ die Zeilen von $\mathcal{C}$. Umkehrt ergibt jede solche Summe eine derartige Faktorisierung, $\sum \bar{c}_k \bar{c}_k^{\top} = \mathcal{C}^{\top}\mathcal{C}$, wobei die Zeilen von $\mathcal{C}$ bis auf Permutationen der Indizes 1, ..., k, ... durch die Zeilenvektoren $\bar{c}_k^{\top}$ festgelegt sind.

Beweis:

Seien $\bar{e}_i, \bar{e}_j$ bzw. $\bar{e}_k, \bar{e}_l$ Einheitsvektoren des E^n bzw. E^m.
Mit der Darstellung $\mathcal{C} = \sum_{j,k} c_{kj}\, \bar{e}_k \bar{e}_j^T = \sum_{j,l} c_{lj}\, \bar{e}_l \bar{e}_j^T$ der $(m \times n)$-Matrix $\mathcal{C}$, entsprechend $\mathcal{C}^T = \sum_{j,k} c_{kj}\, \bar{e}_j \bar{e}_k^T = \sum_{i,k} c_{ki}\, \bar{e}_i \bar{e}_k^T$, erhält man in direkter Rechnung:

$$\mathcal{C}^T\mathcal{C} = \left(\sum_{i,k} c_{ki}\, \bar{e}_i \bar{e}_k^T\right)\left(\sum_{j,l} c_{lj}\, \bar{e}_l \bar{e}_j^T\right) = \sum_{k,l}\sum_{i,j} c_{ki}c_{lj}\, \bar{e}_i \bar{e}_k^T \bar{e}_l \bar{e}_j^T$$

$$= \sum_{k,l}\sum_{i,j} c_{ki}c_{lj}\,\delta_{kl}\, \bar{e}_i \bar{e}_j^T = \sum_{k}\sum_{i,j} c_{ki}c_{kj}\, \bar{e}_i \bar{e}_j^T$$

$$= \sum_{k}\sum_{i} c_{ki}\, \bar{e}_i \sum_{j} c_{kj}\, \bar{e}_j^T = \sum_{k} \bar{c}_k \bar{c}_k^T \ ,$$

wobei $\bar{c}_k := \sum_i c_{ki}\, \bar{e}_i$ die k-te Spalte von $\mathcal{C}^T$ und entsprechend $\bar{c}_k^T =$
$= \sum_i c_{ki}\, \bar{e}_i^T = \sum_j c_{kj}\, \bar{e}_j^T$ die k-te Zeile von $\mathcal{C}$ ist.
Die Rechnung ist umkehrbar zu $\sum_k \bar{c}_k \bar{c}_k^T = \mathcal{C}^T\mathcal{C}$, wobei die Matrix $\mathcal{C}^T\mathcal{C}$ unabhängig von Permutationen der Zeilenindizes 1, ..., k, ... von $\mathcal{C}$ eindeutig ist. q.e.d.

4. Satz

Seien $\mathcal{B}_1$, $\mathcal{B}_2$ positiv semidefinite $(n \times n)$-Matrizen, die in einer echten oder nicht echten Faktorisierung der folgenden Form vorliegen:

$$\mathcal{B}_1 = \mathcal{A}_1^T\mathcal{P}_1\mathcal{A}_1 \ , \quad \mathcal{B}_2 = \mathcal{A}_2^T\mathcal{P}_2\mathcal{A}_2 \ ,$$

mit positiv semidefiniten $(m_1 \times m_1)$- bzw. $(m_2 \times m_2)$-Matrizen $\mathcal{P}_1$, $\mathcal{P}_2$ und (notwendig) reellen $(m_1 \times n)$- bzw. $(m_2 \times n)$-Matrizen $\mathcal{A}_1$, $\mathcal{A}_2$ als Faktoren.
Für den Rang der Summe $\mathcal{B}_1 + \mathcal{B}_2$ gilt:

a. $$\mathrm{rg}\,(\mathcal{B}_1+\mathcal{B}_2) = \mathrm{rg}\left[\mathcal{A}_1^T\mathcal{P}_1 \mid \mathcal{A}_2^T\mathcal{P}_2\right] = \mathrm{rg}\left[\begin{array}{cc}\mathcal{P}_1 & \mathcal{A}_1 \\ \mathcal{P}_2 & \mathcal{A}_2\end{array}\right]$$

Spezialfälle dieser Aussage sind:

b. $$\mathrm{rg}\,(\mathcal{B}_1+\mathcal{B}_2) = \mathrm{rg}\left[\mathcal{B}_1 \mid \mathcal{B}_2\right] = \mathrm{rg}\left[\begin{array}{c}\mathcal{B}_1 \\ \mathcal{B}_2\end{array}\right]$$

c. $\operatorname{rg}(\mathcal{B}_1+\mathcal{B}_2) = \operatorname{rg}\left[\mathcal{C}_1^\top \mid \mathcal{C}_2^\top\right] = \operatorname{rg}\begin{bmatrix}\mathcal{C}_1\\ \mathcal{C}_2\end{bmatrix}$,

wenn $\mathcal{B}_1$, $\mathcal{B}_2$ in Faktorisierungen der Form $\mathcal{B}_1 = \mathcal{C}_1^\top\mathcal{C}_1$, $\mathcal{B}_2 = \mathcal{C}_2^\top\mathcal{C}_2$ mit (notwendig) reellen $(m_1\times n)$- bzw. $(m_2\times n)$-Matrizen $\mathcal{C}_1$, $\mathcal{C}_2$ vorliegen.

Beweis:

Zunächst Beweis der Spezialfälle b. und c.:

Seien $\mathcal{B}_1$, $\mathcal{B}_2$ positiv semidefinite Matrizen und $\mathcal{B}_1 = \mathcal{C}_1^\top\mathcal{C}_1$, $\mathcal{B}_2 = \mathcal{C}_2^\top\mathcal{C}_2$ seien Faktorisierungen, die nach Korollar 1.a immer existieren (auch wenn sie nicht vorgegeben sind). Damit läßt sich die Summe $\mathcal{B}_1+\mathcal{B}_2$ mit Hilfe von Hilfssatz 3. in folgender Weise als 'Normalgleichungsmatrix' darstellen:

$$\mathcal{B}_1+\mathcal{B}_2 = \mathcal{C}_1^\top\mathcal{C}_1+\mathcal{C}_2^\top\mathcal{C}_2 = \sum_{k\in k_1}\bar{c}_k\bar{c}_k^\top + \sum_{k\in k_2}\bar{c}_k\bar{c}_k^\top$$

$$= \sum_{k\in k_1\cup k_2}\bar{c}_k\bar{c}_k^\top = \mathcal{C}^\top\mathcal{C} \ ,$$

mit $\mathcal{C} := \begin{bmatrix}\mathcal{C}_1\\ \mathcal{C}_2\end{bmatrix}$ (Hilfssatz 3.Umkehrung) oder $\mathcal{C}^\top := \left[\mathcal{C}_1^\top \mid \mathcal{C}_2^\top\right]$.

Diese Darstellung zeigt unmittelbar: $s(\mathcal{C}^\top) = s(\mathcal{C}_1^\top) \cup s(\mathcal{C}_2^\top)$ (3-2)

Berücksichtigt man für Matrizen $\mathcal{M}_1$, $\mathcal{M}_2$ noch
$s(\mathcal{M}_1) \cup s(\mathcal{M}_2) \subset \operatorname{lin} s(\mathcal{M}_1) \cup \operatorname{lin} s(\mathcal{M}_2) \subset \operatorname{lin}\big(s(\mathcal{M}_1) \cup s(\mathcal{M}_2)\big)$,
wobei Anwendung der Operation lin auf diese Inklusionen zu

$$\operatorname{lin}\big(s(\mathcal{M}_1) \cup s(\mathcal{M}_2)\big) = \operatorname{lin}\big(\operatorname{lin} s(\mathcal{M}_1) \cup \operatorname{lin} s(\mathcal{M}_2)\big) \qquad (3\text{-}3)$$

führt, so erhält man schließlich mit Satz 2.:

$\operatorname{rg}(\mathcal{B}_1+\mathcal{B}_2) = \operatorname{rg}(\mathcal{C}^\top\mathcal{C})$ - Satz 2.c

$= \operatorname{rg}(\mathcal{C}^\top) = \operatorname{rg}\left(s(\mathcal{C}^\top)\right) = \operatorname{rg}\left(s(\mathcal{C}_1^\top) \cup s(\mathcal{C}_2^\top)\right)^{(*)}$ - (3-2) u. Def. d. Ranges

$= \dim \operatorname{lin}\left(s(\mathcal{C}_1^\top) \cup s(\mathcal{C}_2^\top)\right)$ - (3-3)

$= \dim \operatorname{lin}\left(\operatorname{lin} s(\mathcal{C}_1^\top) \cup \operatorname{lin} s(\mathcal{C}_2^\top)\right)$ - Satz 2.b

$= \dim \operatorname{lin}\left(\operatorname{lin} s(\mathcal{C}_1^\top\mathcal{C}_1) \cup \operatorname{lin} s(\mathcal{C}_2^\top\mathcal{C}_2)\right)$ - $\mathcal{C}_1^\top\mathcal{C}_1 = \mathcal{B}_1$, $\mathcal{C}_2^\top\mathcal{C}_2 = \mathcal{B}_2$

$= \dim \operatorname{lin}\left(\operatorname{lin} s(\mathcal{B}_1) \cup \operatorname{lin} s(\mathcal{B}_2)\right)$ - (3-3)

$= \dim \operatorname{lin}\left(s(\mathcal{B}_1) \cup s(\mathcal{B}_2)\right) = \operatorname{rg}\left(s(\mathcal{B}_1) \cup s(\mathcal{B}_2)\right)^{(**)}$ - Definition d. Ranges

Damit ist gezeigt:

$$\mathrm{rg}(\mathcal{B}_1+\mathcal{B}_2) = \mathrm{rg}\left[\mathcal{C}_1^\top \mid \mathcal{C}_2^\top\right]^{(*)} = \mathrm{rg}[\mathcal{B}_1 \mid \mathcal{B}_2]^{(**)}$$

Die restlichen Behauptungen von b., c. sind bewiesen durch die Entsprechung von Spalten und Zeilen bei Transposition einer Matrix und durch die Symmetrie von $\mathcal{B}_1$, $\mathcal{B}_2$.

Beweis von a.

Da $\mathcal{P}_1$, $\mathcal{P}_2$ positiv semidefinit sind, gibt es nach Korollar 1.a Faktorisierungen $\mathcal{P}_1 = \mathcal{F}_1^\top\mathcal{F}_1$, $\mathcal{P}_2 = \mathcal{F}_2^\top\mathcal{F}_2$.
Mit $\mathcal{C}_1 := \mathcal{F}_1\mathcal{A}_1$, $\mathcal{C}_2 := \mathcal{F}_2\mathcal{A}_2$ erhält man:

$$\mathcal{B}_1 = \mathcal{A}_1^\top\mathcal{P}_1\mathcal{A}_1 = \mathcal{A}_1^\top\mathcal{F}_1^\top\mathcal{F}_1\mathcal{A}_1 = \mathcal{C}_1^\top\mathcal{C}_1 \ ,$$

$$\mathcal{B}_2 = \mathcal{A}_2^\top\mathcal{P}_2\mathcal{A}_2 = \mathcal{A}_2^\top\mathcal{F}_2^\top\mathcal{F}_2\mathcal{A}_2 = \mathcal{C}_2^\top\mathcal{C}_2 \ ,$$

Damit folgt:

$$\mathrm{rg}\begin{bmatrix}\mathcal{C}_1\\ --\\ \mathcal{C}_2\end{bmatrix} = \mathrm{rg}\begin{bmatrix}\mathcal{F}_1\mathcal{A}_1\\ ---\\ \mathcal{F}_2\mathcal{A}_2\end{bmatrix} \geq \mathrm{rg}\left(\begin{bmatrix}\mathcal{F}_1^\top & \mid & \mathcal{O}\\ --&+&--\\ \mathcal{O} & \mid & \mathcal{F}_2^\top\end{bmatrix}\begin{bmatrix}\mathcal{F}_1\mathcal{A}_1\\ ---\\ \mathcal{F}_2\mathcal{A}_2\end{bmatrix}\right) = \mathrm{rg}\begin{bmatrix}\mathcal{F}_1^\top\mathcal{F}_1\mathcal{A}_1\\ -----\\ \mathcal{F}_2^\top\mathcal{F}_2\mathcal{A}_2\end{bmatrix} =$$

$$= \mathrm{rg}\begin{bmatrix}\mathcal{P}_1\mathcal{A}_1\\ ---\\ \mathcal{P}_2\mathcal{A}_2\end{bmatrix} \geq \mathrm{rg}\left(\begin{bmatrix}\mathcal{A}_1^\top & \mid & \mathcal{O}\\ --&+&--\\ \mathcal{O} & \mid & \mathcal{A}_2^\top\end{bmatrix}\begin{bmatrix}\mathcal{P}_1\mathcal{A}_1\\ ---\\ \mathcal{P}_2\mathcal{A}_2\end{bmatrix}\right) = \mathrm{rg}\begin{bmatrix}\mathcal{A}_1^\top\mathcal{P}_1\mathcal{A}_1\\ -----\\ \mathcal{A}_2^\top\mathcal{P}_2\mathcal{A}_2\end{bmatrix} = \mathrm{rg}\begin{bmatrix}\mathcal{B}_1\\ --\\ \mathcal{B}_2\end{bmatrix}$$

Mit $\mathrm{rg}\begin{bmatrix}\mathcal{C}_1\\ --\\ \mathcal{C}_2\end{bmatrix} = \mathrm{rg}\begin{bmatrix}\mathcal{B}_1\\ --\\ \mathcal{B}_2\end{bmatrix}$ nach b. und c. kann in diesen Ungleichungen nur das Gleichheitszeichen stehen, so daß folgt:

$$\mathrm{rg}\begin{bmatrix}\mathcal{C}_1\\ --\\ \mathcal{C}_2\end{bmatrix} = \mathrm{rg}\begin{bmatrix}\mathcal{P}_1\mathcal{A}_1\\ ---\\ \mathcal{P}_2\mathcal{A}_2\end{bmatrix} = \mathrm{rg}\begin{bmatrix}\mathcal{B}_1\\ --\\ \mathcal{B}_2\end{bmatrix}$$

Damit ist auch a. gezeigt.
Setzt man $\mathcal{A}_1 := \mathcal{I}$, $\mathcal{A}_2 := \mathcal{I}$, $\mathcal{P}_1 := \mathcal{B}_1$, $\mathcal{P}_2 := \mathcal{B}_2$ bzw. $\mathcal{A}_1 := \mathcal{C}_1$, $\mathcal{A}_2 := \mathcal{C}_2$, $\mathcal{P}_1 := \mathcal{I}$, $\mathcal{P}_2 := \mathcal{I}$ so erhält man b. bzw. c. als Spezialfälle von a. q.e.d.

Eine analoge Anwendung von Satz 4. auf beliebige endliche Summen von n positiv semidefiniten Matrizen ist unmittelbar ersichtlich. Die zeilenweise bzw. spaltenweise Vereinigung von 2 Matrizen geht lediglich über in die von

n Matrizen. Dabei dürfen die den Aussagen a. bis c. zugrundeliegenden echten oder nicht echten Faktorisierungen gemischt auftreten. Auch können vorliegende Faktorisierungen umgeformt werden, wenn das die Rangbestimmung einfacher gestaltet. Dabei sind folgende Umformungen denkbar:

Faktorisierung $\mathcal{P} = \mathcal{F}^\top \mathcal{P} \mathcal{F}$ (auch mit $\tilde{\mathcal{P}} := \mathcal{I}$) :

$$\mathcal{B} = \mathcal{A}^\top \mathcal{P} \mathcal{A} = \mathcal{A}^\top \mathcal{F}^\top \tilde{\mathcal{P}} \mathcal{F} \mathcal{A} = \tilde{A}^\top \tilde{P} \tilde{A} \quad , \quad \tilde{\mathcal{A}} := \mathcal{F} \mathcal{A}$$

Faktorisierung $\mathcal{A} = \mathcal{F} \tilde{A}$ (auch mit $\mathcal{P} = \mathcal{I}$) :

$$\mathcal{B} = \mathcal{A}^\top \mathcal{P} \mathcal{A} = \tilde{\mathcal{A}}^\top \mathcal{F}^\top \tilde{\mathcal{P}} \mathcal{F} \tilde{\mathcal{A}} = \tilde{A}^\top \tilde{P} \tilde{A} \quad , \quad \tilde{\mathcal{P}} := \mathcal{F}^\top \mathcal{P} \mathcal{F}$$

Satz 4. ist ebenso auf eine positiv semidefinite Matrix $\mathcal{B}$ allein anwendbar, wenn man $\mathcal{B}$ als '1-gliedrige Summe' betrachtet. Damit stellt Satz 4. eine Verallgemeinerung von Satz 2. dar.

Im folgenden Beispiel wird die Anwendung von Satz 4. an der Summe zweier Normalgleichungsmatrizen demonstriert, wobei zur Bestimmung des Ranges der Normalgleichungsmatrizen selbst ebenfalls Satz 4. angewendet werden kann. Satz 4. besagt, daß 'nur' der Rang der gewichteten Jacobimatrizen bzw. deren zeilenweiser Vereinigung zu bestimmen ist. Die Normalgleichungsmatrizen müssen also zur Ranganalyse von der mathematischen Grundaussage her gar nicht vollständig aufgestellt werden. Dieser Vorteil wird auch beim folgenden Beispiel genützt.

4. BEISPIEL

Ein in 2 Epochen beobachtetes Höhennetz, bestehend aus 3 Punkten.

Liegen Beobachtungen aus 2 Epochen vor, so können für den Rang der Normalgleichungsmatrizen (mit $\mathcal{N}$ bezeichnet) folgende Fälle auftreten:

1. $\operatorname{rg}(\mathcal{N}_1 + \mathcal{N}_2) = \operatorname{rg}(\mathcal{N}_1) + \operatorname{rg}(\mathcal{N}_2)$, $\operatorname{lin s}(\mathcal{N}_1) \cap \operatorname{lin s}(\mathcal{N}_2) = \bar{\mathcal{O}}$
2. $\operatorname{rg}(\mathcal{N}_1 + \mathcal{N}_2) < \operatorname{rg}(\mathcal{N}_1) + \operatorname{rg}(\mathcal{N}_2)$, $\operatorname{lin s}(\mathcal{N}_1) \cap \operatorname{lin s}(\mathcal{N}_2) \neq \bar{\mathcal{O}}$
3. $\operatorname{rg}(\mathcal{N}_1 + \mathcal{N}_2) = \operatorname{rg}(\mathcal{N}_1)$, $\operatorname{lin s}(\mathcal{N}_2) \subset \operatorname{lin s}(\mathcal{N}_1)$

In der geodätischen Praxis tritt häufig der Fall 2. auf, der somit für eine Ranganalyse am interessantesten ist. Ein solcher Fall wird im vorliegenden Beispiel bearbeitet.

Zunächst werden beide Epochen getrennt behandelt. Dabei werden die Normalgleichungsmatrizen aufgestellt, jedoch in expliziter Form nur insoweit, als es zur Ranganalyse erforderlich ist, und die Ranganalyse wird durchgeführt. Dann wird die Normalgleichungsmatrix für das Gesamtmodell, unter Zusammenfassung der Beobachtungen aus beiden Epochen, aufgestellt. Unter der getroffenen Annahme der Unkorreliertheit der Beobachtungen aus verschiedenen Epochen ist sie die Summe der Normalgleichungsmatrizen der einzelnen Epochen. Auch ihr Rang wird bestimmt. Eine vergleichende geometrische Betrachtung beschließt die Ranganalysen.

Unbekannte sind jeweils die Höhen H_i , genauer die Höhenverbesserungen $\Delta H_i := H_i - H_{i0}$, H_{i0} Näherungen, da das Beispiel zu Epoche 2 in linearisierter Form vorgegeben ist.
Für jede Epoche nehmen wir an, daß die beobachteten Höhendifferenzen $\Delta H_{ij} := H_j - H_i$ unkorreliert und gleichgewichtig sind. Weiter gehen wir davon aus, daß die Beobachtungsgleichungen durch Division mit dem mittleren Fehler homogenisiert wurden. Diese Annahmen erklären die Einführung der Diagonalmatrizen $\mathcal{Q}_{ll} = p_{\Delta H_1}^{-1} \mathcal{I}$ (Epoche 1), $\mathcal{Q}_{\Delta H_2} = p_{\Delta H_1}^{-1} \mathcal{I}$ (Epoche 2) als Kofaktormatrizen für die beobachteten Höhendifferenzen sowie der Blockdiagonalmatrix $\mathcal{Q}_{ll_{20}}$ als Kofaktorenmatrix der für Epoche 2 als Beobachtungen eingeführten Höhendifferenzverhältnisse. Dazu sind $p_{\Delta H_1}$, $p_{\Delta H_2}$ die reziproken Varianzen der jeweils beobachteten Höhendifferenzen und $\mathcal{I}$ die Einheitsmatrix. Durch die Homogenisierung sämtlicher Beobachtungen sind die unterschiedlichen Gewichte der Beobachtungen aus Epoche 1 und Epoche 2 bereits in Rechnung gestellt. Dabei nehmen wir an, daß keine Korrelation zwischen Beobachtungen aus den verschiedenen Epochen 1 und 2 besteht.

<u>Epoche 1.</u> <u>Höhendifferenzen beobachtet.</u>

<u>a. Funktionales Modell.</u>
Beobachtete Größen:

$$\Delta H_{12} := H_2 - H_1 \quad , \quad \Delta H_{23} := H_3 - H_2 \quad , \quad \Delta H_{31} := H_1 - H_3$$

Jacobimatrix der beobachteten Größen:

$$\mathcal{A}_1 := \left[\frac{\partial \Delta H_{ij}}{\partial H_k} \right] = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

b. Stochastisches Modell.

$$\mathcal{Q}_{l_1} = p_{\Delta H_1}^{-1}\,\mathcal{J}$$

c. Normalgleichungsmatrix.

$$\mathcal{N}_1 = \mathcal{A}_1^T \mathcal{Q}_{l_1}^{-1} \mathcal{A}_1 = p_{\Delta H_1} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

d. Ranganalyse.

Die beiden ersten Zeilen von $\mathcal{N}_1$ sind linear unabhängig. Die letzte Zeile ist die negative Summe der beiden ersten und somit linear abhängig. Folglich gilt $rg\mathcal{N}_1 = 2$.

Epoche 2. Höhendifferenzverhältnisse und eine Höhe beobachtet.

a. Funktionales Modell.

Die beobachteten Höhendifferenzen sind gegenüber den Beobachtungen in Epoche 1 mit einem unbekannten Maßstabsfaktor λ behaftet. Das würde zunächst bedeuten, daß eine zusätzliche Unbekannte λ einzuführen ist. Diese zusätzliche Unbekannte wird jedoch vermieden, indem wir die Höhendifferenzverhältnisse als beobachtete Größen einführen.

$$V_1 = \Delta H_{31}/\Delta H_{12}\ ,\ V_2 = \Delta H_{12}/\Delta H_{23}\ ,\ V_3 = \Delta H_{23}/\Delta H_{31} \qquad (4\text{-}1)$$

Damit die Höhendifferenzverhältnisse V_i definiert sind, beschränken wir uns auf den fast immer auftretenden Fall nicht verschwindender Höhendifferenzen, $\Delta H_{ij} \neq 0$.
Es bestehen folgende Beziehungen:

$$\Delta H_{12} + \Delta H_{23} + \Delta H_{31} = 0 \qquad (4\text{-}2)$$

$$V_1 \cdot V_2 \cdot V_3 = 1 \qquad (4\text{-}3)$$

$$V_1 V_3 = V_2^{-1} = -(1 + V_1) \qquad (4\text{-}4a)$$

$$V_1 V_2 = V_3^{-1} = -(1 + V_2) \qquad (4\text{-}4b)$$

$$V_2 V_3 = V_1^{-1} = -(1 + V_3) \qquad (4\text{-}4c)$$

Diese Beziehungen gehen in einige einfache Umformungen ein, die nicht näher kommentiert sind.
Als weitere Beobachtung liege die Höhe H_1 vor (z.B. barometrisch bestimmt). Als Jacobimatrix der beobachteten Größen erhält man somit:

$$\mathcal{A}_2 = \begin{bmatrix} \mathcal{A}_{20} \\ \bar{e}_1^T \end{bmatrix}, \qquad \bar{e}_1^T = \begin{bmatrix} \frac{\partial H_1}{\partial H_j} \end{bmatrix},$$

$$\mathcal{A}_{20} = \begin{bmatrix} \frac{\partial V_i}{\partial H_j} \end{bmatrix} = -\begin{bmatrix} (\Delta H_{12} \cdot V_2)^{-1} & \Delta H_{12}^{-1} V_1 & \Delta H_{12}^{-1} \\ \Delta H_{23}^{-1} & (\Delta H_{23} \cdot V_3)^{-1} & \Delta H_{23}^{-1} V_2 \\ \Delta H_{31}^{-1} V_3 & \Delta H_{31}^{-1} & (\Delta H_{31} \cdot V_1)^{-1} \end{bmatrix} =$$

$$= -\begin{bmatrix} \Delta H_{12}^{-1} & 0 & 0 \\ 0 & \Delta H_{23}^{-1} & 0 \\ 0 & 0 & \Delta H_{31}^{-1} \end{bmatrix} \begin{bmatrix} V_1 V_3 & V_1 & 1 \\ 1 & V_1 V_2 & V_2 \\ V_3 & 1 & V_2 V_3 \end{bmatrix} = \mathcal{D}_{\Delta H}^{-1} \begin{bmatrix} V_1 \\ V_1 V_2 \\ 1 \end{bmatrix} [V_3, 1, V_2 V_3] =$$

$$= \mathcal{D}_{\Delta H}^{-1} \bar{v}_1 \bar{v}_2^T, \quad \mathcal{D}_{\Delta H}^{-1} := -\begin{bmatrix} \Delta H_{12}^{-1} & 0 & 0 \\ 0 & \Delta H_{23}^{-1} & 0 \\ 0 & 0 & \Delta H_{31}^{-1} \end{bmatrix}, \quad \bar{v}_1 := \begin{bmatrix} V_1 \\ V_1 V_2 \\ 1 \end{bmatrix}, \quad \bar{v}_2 := \begin{bmatrix} V_3 \\ 1 \\ V_2 V_3 \end{bmatrix}$$

(4-5)

Die Zeilen von $\mathcal{A}_{20}$ sind mit der Voraussetzung $\Delta H_{ij} \neq 0$ definierte sowie nicht verschwindende Vielfache der Zeile $\bar{v}_2^T = [V_3, 1, V_2 V_3]$. Die Zeile $\bar{e}_1^T = [1,0,0]$ von $\mathcal{A}_2$ ist linear unabhängig davon. Also gilt $\mathrm{rg}\mathcal{A}_2 = 2$.

b. Stochastisches Modell.

Da die Höhendifferenzen ΔH_{ij} als nicht korreliert vorausgesetzt sind, ist die Kofaktorenmatrix für die Höhendifferenzverhältnisse V_i :

$$Q_{l_{20}} = \mathcal{M} Q_{\Delta H_2} \mathcal{M}^T \quad , \quad \text{mit der Jacobimatrix}$$

$$\mathcal{M} := \begin{bmatrix} \frac{\partial V_k}{\partial \Delta H_{ij}} \end{bmatrix} = \begin{bmatrix} -\Delta H_{31} \Delta H_{12}^{-2} & 0 & \Delta H_{12}^{-1} \\ \Delta H_{23}^{-1} & -\Delta H_{12} \Delta H_{23}^{-2} & 0 \\ 0 & \Delta H_{31}^{-1} & -\Delta H_{23} \Delta H_{31}^{-2} \end{bmatrix} =$$

$$= \begin{bmatrix} -\Delta H_{12}^{-1} V_1 & 0 & \Delta H_{12}^{-1} \\ \Delta H_{23}^{-1} & -\Delta H_{23}^{-1} V_2 & 0 \\ 0 & \Delta H_{31}^{-1} & -\Delta H_{31}^{-1} V_3 \end{bmatrix} = \mathcal{F}\widetilde{\mathcal{M}}^{\top} \quad ,$$

$$\mathcal{F} := \begin{bmatrix} -\Delta H_{12}^{-1} & 0 \\ 0 & -\Delta H_{23}^{-1} \\ V_3 \Delta H_{31}^{-1} & V_1 V_3 \Delta H_{31}^{-1} \end{bmatrix} \quad , \quad \widetilde{\mathcal{M}} := \begin{bmatrix} V_1 & -1 \\ 0 & V_2 \\ -1 & 0 \end{bmatrix}$$

Mit $\mathcal{Q}_{\Delta H_2} = p_{\Delta H_2}^{-1} \mathcal{I}$ folgt:

$$\mathcal{Q}_{l_{20}} = p_{\Delta H_2}^{-1} \mathcal{F} \widetilde{\mathcal{M}}^{\top} \widetilde{\mathcal{M}} \mathcal{F}^{T} = p_{\Delta H_2}^{-1} \mathcal{F} \mathcal{P} \mathcal{F}^{\top} \quad ,$$

$$\mathcal{P} := \widetilde{\mathcal{M}}^{T} \widetilde{\mathcal{M}} = \begin{bmatrix} 1+V_1^2 & -V_1 \\ -V_1 & 1+V_2^2 \end{bmatrix}$$

Die Pseudoinverse von $\mathcal{Q}_{l_{20}}$ hat damit die Faktorisierung:

$$\mathcal{Q}_{l_{20}}^{+} = p_{\Delta H_2} \mathcal{F} \left(\mathcal{F}^{\top}\mathcal{F}\right)^{-1} \left((\mathcal{F}\mathcal{P})^{\top} \mathcal{F}\mathcal{P}\right)^{-1} (\mathcal{F}\mathcal{P})^{\top}$$

$$= p_{\Delta H_2} \mathcal{F} \left(\mathcal{F}^{\top}\mathcal{F}\right)^{-1} \mathcal{P}^{-1} \left(\mathcal{F}^{\top}\mathcal{F}\right)^{-1} \mathcal{F}^{\top}$$

$$= p_{\Delta H_2} \mathcal{F} \left(\left(\mathcal{F}^{\top}\mathcal{F}\right)^{\top} \mathcal{P} \left(\mathcal{F}^{\top}\mathcal{F}\right) \right)^{-1} \mathcal{F}^{\top}$$

$$= \mathcal{F} \tilde{\mathcal{P}}_{20} \mathcal{F}^{\top} \quad , \quad \tilde{\mathcal{P}}_{20} := p_{\Delta H_2} \left(\left(\mathcal{F}^{\top}\mathcal{F}\right)^{\top} \mathcal{P} \left(\mathcal{F}^{\top}\mathcal{F}\right) \right)^{-1} \quad .$$

$\mathcal{P}$ ist mit $\det\mathcal{P} = 1 + V_2^2 \left(1 + V_1^2\right)$ nichtsingulär (Determinante $\neq 0$) und somit positiv definit. Aufgrund der Voraussetzung $\Delta H_{ij} \neq 0$ gilt $\operatorname{rg}\mathcal{F} = 2$, so daß $\mathcal{F}^{\top}\mathcal{F}$ nach Satz 2. nichtsingulär ist. Also ist $\left(\mathcal{F}^{\top}\mathcal{F}\right)\mathcal{P}\left(\mathcal{F}^{\top}\mathcal{F}\right)$ positiv definit, somit nichtsingulär und invertierbar. $\tilde{\mathcal{P}}_{20}$ ist dann als Inverse einer positiv definiten Matrix ebenfalls positiv definit.

Weiter nehmen wir an, daß die Höhendifferenzverhältnisse V_i und die Höhe H_1 unkorreliert sind. Zusammen mit der Beobachtungsgleichung von H_1, homogenisiert mit der reziproken Varianz p_{H_2} von H_1, ist dann die Kofaktormatrix für Epoche 2 gegeben durch:

$$\mathcal{Q}_{1_2} = \left[\begin{array}{c|c} \mathcal{Q}_{1_{20}} & \bar{\mathcal{O}} \\ \hline \bar{\mathcal{O}}^{\top} & p_{H_2}^{-1} \end{array}\right] \quad , \quad \text{womit gilt:} \quad \mathcal{Q}_{1_2}^{+} = \left[\begin{array}{c|c} \mathcal{Q}_{1_{20}}^{+} & \bar{\mathcal{O}} \\ \hline \bar{\mathcal{O}}^{\top} & p_{H_2} \end{array}\right]$$

c. Normalgleichungsmatrix.

Sie läßt sich in folgender Form faktorisieren:

$$\mathcal{N}_2 = \mathcal{A}_2^{\top}\,\mathcal{Q}_{1_2}^{+}\,\mathcal{A}_2 = \left[\mathcal{A}_{20}^{\top} \mid \bar{e}_1\right] \left[\begin{array}{c|c} \mathcal{Q}_{1_{20}}^{+} & \bar{\mathcal{O}} \\ \hline \bar{\mathcal{O}}^{\top} & p_{H_2} \end{array}\right] \left[\begin{array}{c} \mathcal{A}_{20} \\ \hline \bar{e}_1^{\top} \end{array}\right] =$$

$$= \left[\mathcal{A}_{20}^{\top} \mid \bar{e}_1\right] \left[\begin{array}{c|c} \mathcal{F}\,\tilde{\mathcal{P}}_{20}\,\mathcal{F}^{\top} & \bar{\mathcal{O}} \\ \hline \bar{\mathcal{O}}^{\top} & p_{H_2} \end{array}\right] \left[\begin{array}{c} \mathcal{A}_{20} \\ \hline \bar{e}_1^{\top} \end{array}\right] = \left[\left(\mathcal{F}^{\top}\mathcal{A}_{20}\right)^{\top} \mid \bar{e}_1\right] \left[\begin{array}{c|c} \tilde{\mathcal{P}}_{20} & \bar{\mathcal{O}} \\ \hline \bar{\mathcal{O}}^{\top} & p_{H_2} \end{array}\right] \left[\begin{array}{c} \mathcal{F}^{\top}\mathcal{A}_{20} \\ \hline \bar{e}_1^{\top} \end{array}\right] \quad ,$$

$$\mathcal{F}^{\top}\mathcal{A}_{20} = \mathcal{F}^{\top}\,\mathcal{D}_{\Delta H}^{-1}\,\bar{v}_1\,\bar{v}_2^{\top} = \left[\begin{array}{c} \Delta H_{12}^{-2} - V_3\,\Delta H_{31}^{-2} \\ \\ V_1\left(\Delta H_{23}^{-2}\,V_2 - V_3\,\Delta H_{31}^{-2}\right) \end{array}\right] \bar{v}_2^{\top} =$$

$$= \bar{x}\,\bar{v}_2^{\top} \;, \quad \bar{x} := \mathcal{F}^{\top}\,\mathcal{D}_{\Delta H}^{-1}\,\bar{v}_1\,\bar{v}_2^{\top} = \left[\begin{array}{c} x_1 \\ \\ x_2 \end{array}\right] = \Delta H_{12}^{-2} \left[\begin{array}{c} V_1 - V_2^2\,V_3^3 \\ \\ V_1\,V_2^2\left(V_2 - V_3^3\right) \end{array}\right] \qquad (4\text{-}6)$$

Damit erhält man die zur Ranganalyse geeignete Faktorisierung:

$$\mathcal{N}_2 = \tilde{\mathcal{A}}_2^{\top}\,\tilde{\mathcal{P}}_2\,\tilde{\mathcal{A}}_2 \;, \qquad \tilde{\mathcal{P}}_2 := \left[\begin{array}{c|c} \tilde{\mathcal{P}}_{20} & \bar{\mathcal{O}} \\ \hline \bar{\mathcal{O}}^{\top} & p_{H_2} \end{array}\right], \qquad \tilde{\mathcal{A}}_2 := \left[\begin{array}{c} \bar{x}\,\bar{v}_2^{\top} \\ \hline \bar{e}_1^{\top} \end{array}\right] \quad .$$

$\tilde{\mathcal{P}}_2$ ist aufgrund seiner Struktur wie $\tilde{\mathcal{P}}_{20}$ positiv definit. $\bar{x}, \bar{v}_2$ entsprechend (4-6), (4-5).
Unter der getroffenen Voraussetzung $\Delta H_{ij} \neq 0$ läßt sich $\bar{x} \neq 0$ zeigen, so daß das dyadische Produkt $\bar{x}\,\bar{v}_2^{\top}$ den Rang 1 und nicht den Rang 0 hat.
Wäre $\bar{x} = 0$, dann müßte nach (4-6) gelten:

$$V_1 = V_2^2\,V_3^3 \quad \text{und} \quad V_2 = V_3^3$$

Daraus folgt nacheinander:

$1 = V_1\,V_2\,V_3 = V_2^3\,V_3^4 = V_3^{13} \implies V_3 = +1 \implies$
$\implies V_2 = V_3^3 = +1 \implies V_1 = V_2^2\,V_3^3 = +1$, also $V_1 = V_2 = V_3 = +1 \implies$
$\implies \Delta H_{12} = \Delta H_{23} = \Delta H_{31} = 0$, im Widerspruch zu $\Delta H_{ij} \neq 0$.

Damit ist $\bar{x} \neq 0$ gezeigt.

d. Ranganalyse.

Nach Satz 4.a gilt für $\mathcal{N}_2 = \tilde{\mathcal{A}}_2^\top \tilde{\mathcal{P}}_2 \tilde{\mathcal{A}}_2$ als '1-gliedrige' Summe positiv semidefiniter Matrizen, nachdem $\tilde{\mathcal{P}}_2$ positiv semidefinit ist: $\mathrm{rg}\mathcal{N}_2 = \mathrm{rg}\tilde{\mathcal{P}}_2\tilde{\mathcal{A}}_2$. $\tilde{\mathcal{P}}_2$ ist als positiv definite Matrix nichtsingulär und ändert daher den Rang nicht, so daß gilt: $\mathrm{rg}\tilde{\mathcal{P}}_2\tilde{\mathcal{A}}_2 = \mathrm{rg}\tilde{\mathcal{A}}_2$.
Die ersten beiden Zeilen $\bar{x}\,\bar{v}_2^\top$ von $\tilde{\mathcal{A}}_2$ sind nichtverschwindende Vielfache ein und derselben Zeile $\bar{v}_2^\top = [V_3, 1, V_2V_3]$.
Linear unabhängig davon ist die dritte Zeile $\bar{e}_1^\top = [1,0,0]$, so daß $\mathrm{rg}\tilde{\mathcal{A}}_2 = 2$ gilt. Es folgt: $\mathrm{rg}\mathcal{N}_2 = \mathrm{rg}\tilde{\mathcal{P}}_2\tilde{\mathcal{A}}_2 = \mathrm{rg}\tilde{\mathcal{A}}_2 = 2$.

Gesamtmodell.

Zusammenfassung der Beobachtungen aus Epoche 1 und 2.

a. Funktionales Modell.

Jacobimatrix der beobachteten Größen: $\mathcal{A}_{12} := \begin{bmatrix} \mathcal{A}_1 \\ -- \\ \mathcal{A}_2 \end{bmatrix}$

b. Stochastisches Modell.

Da keine Korrelationen zwischen Beobachtungen aus den verschiedenen Epochen 1 und 2 bestehen sollen, erhält man als Kofaktormatrix:

$$\mathcal{Q}_{l_{12}} = \left[\begin{array}{c|c} \mathcal{Q}_{l_1} & \mathcal{O} \\ \hline \mathcal{O} & \mathcal{Q}_{l_2} \end{array}\right]$$

c. Normalgleichungsmatrix.

$$\mathcal{N}_{12} = \mathcal{A}_{12}^\top \mathcal{Q}_{l_{12}}^{+} \mathcal{A}_{12} = \left[\mathcal{A}_1^\top \mid \mathcal{A}_2^\top\right] \left[\begin{array}{c|c} \mathcal{Q}_{l_1}^{-1} & \mathcal{O} \\ \hline \mathcal{O} & \mathcal{Q}_{l_2}^{-1} \end{array}\right] \begin{bmatrix} \mathcal{A}_1 \\ -- \\ \mathcal{A}_2 \end{bmatrix} =$$

$$= \mathcal{A}_1^\top \mathcal{Q}_{l_1}^{-1} \mathcal{A}_1 + \mathcal{A}_2^\top \mathcal{Q}_{l_2}^{-1} \mathcal{A}_2 = \mathcal{N}_1 + \mathcal{N}_2 \ .$$

Also gilt $\mathcal{N}_{12} = \mathcal{N}_1 + \mathcal{N}_2$, entsprechend der Unkorreliertheit der Beobachtungen aus verschiedenen Epochen.

d. Ranganalyse.

Satz 4.a besagt, daß der Rang der Summe $\mathcal{N}_{12}$ der positiv semidefiniten Matrizen $\mathcal{N}_1$, $\mathcal{N}_2$ mit den Faktorisierungen $\mathcal{N}_1 = \mathcal{A}_1^T \mathcal{Q}_{l_1}^{-1} \mathcal{A}_1$, $\mathcal{N}_2 = \mathcal{A}_2^T \mathcal{Q}_{l_2}^{-1} \mathcal{A}_2$, wobei $\mathcal{Q}_{l_1}^{-1}$, $\tilde{\mathcal{P}}_2$ positiv definit sind, gleich dem Rang der zu einer Matrix $\mathcal{F}_{12}$ zusammengefaßten Faktoren $\mathcal{Q}_{l_1}^{-1}\mathcal{A}_1$, $\tilde{\mathcal{P}}_2\tilde{\mathcal{A}}_2$ ist,

$$\mathrm{rg}\mathcal{N}_{12} = \mathrm{rg}\mathcal{F}_{12} \quad , \qquad \mathcal{F}_{12} := \begin{bmatrix} \mathcal{Q}_{l_1}^{-1}\mathcal{A}_1 \\ \hline \tilde{\mathcal{P}}_2\tilde{\mathcal{A}}_2 \end{bmatrix} = \left[\begin{array}{c|c} \mathcal{Q}_{l_1}^{-1} & \mathcal{O} \\ \hline \mathcal{O} & \tilde{\mathcal{P}}_2 \end{array}\right] \cdot \begin{bmatrix} \mathcal{A}_1 \\ \hline \tilde{\mathcal{A}}_2 \end{bmatrix} = \mathcal{P}_{12}\tilde{\mathcal{A}}_{12}$$

Mit $\mathcal{Q}_{l_1}^{-1}$, $\tilde{\mathcal{P}}_2$ ist auch die Blockdiagonalmatrix $\mathcal{P}_{12}$ nichtsingulär und ändert daher den Rang nicht, so daß gilt: $\mathrm{rg}\mathcal{F}_{12} = \mathrm{rg}\tilde{\mathcal{A}}_{12}$. Elementare Zeilenumformungen vom Type (3), zu einer unteren Dreiecksmatrix $\mathcal{L}_e$ zusammengefaßt und auf $\tilde{\mathcal{A}}_{12}$ angewandt, ergeben:

$$\mathcal{L}_e \cdot \tilde{\mathcal{A}}_{12} =$$

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & & & \\ 0 & 1 & 0 & & \mathcal{O} & \\ 1 & 1 & 1 & & & \\ \hline x_1V_3 & -x_1V_2V_3 & 0 & 1 & 0 & 0 \\ x_2V_3 & -x_2V_2V_3 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \cdot \left[\begin{array}{ccc} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \\ \hline x_1[V_3, & 1 & ,V_2V_3] \\ x_2[V_3, & 1 & ,V_2V_3] \\ \hline 1 & 0 & 0 \end{array}\right] =$$

$$= \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 1 & 0 & 0 \end{bmatrix}$$

Die Determinante der (3x3)-Untermatrix mit den nicht verschwindenden Zeilen ist $\neq 0$, so daß gilt: $\mathrm{rg}\mathcal{L}_e\tilde{\mathcal{A}}_{12} = 3$.
Damit ist gezeigt: $\mathrm{rg}\mathcal{N}_{12} = \mathrm{rg}\mathcal{F}_{12} = \mathrm{rg}\,\tilde{\mathcal{A}}_{12} = \mathrm{rg}\mathcal{L}_e\tilde{\mathcal{A}}_{12} = 3$.
Die 3 unbekannten Höhen lassen sich also durch die zusammengefaßten Beobachtungen der beiden Epochen vollständig bestimmen.

Ihre Bestimmung ist äquivalent mit der Festlegung der folgenden 3 Größen:

T Translation

M Maßstab

K Konfiguration (Ähnlichkeit)

[***GRAFAREND and SCHAFFRIN*** 1976].

$T \cup M$ (T und M) bestimmen das Datum des Höhennetzes, $M \cup K$ ergeben die vollständige Konfiguration (Kongruenz). Unter der Voraussetzung, daß das stochastische Modell den Rang des funktionalen Modells nicht vermindert, bilden sich die Beobachtungen wie folgt ab:

Epoche 1 : $\{\Delta H_{ij}\} \rightarrow M \cup K$

Epoche 2 : $H_1 \rightarrow T$, $\{V_i\} \rightarrow K$

Gesamtmodell: $\{\Delta H_{ij}\} \cup \{H_1, \{V_i\}\} \rightarrow (M \cup K) \cup T \cup K = M \cup K \cup T$

Zu den beiden Epochen werden also je 2 Größen derart bestimmt, daß im Gesamtmodell alle Größen bestimmt sind. Diese geometrisch anschaulichen Überlegungen sind durch die Ranganalyse mit Hilfe von Satz 4. bestätigt.

LITERATUR

BJERHAMMAR, A.: ***Theory of Errors and Generalized Matrix Inverses***. Elsevier, Amsterdam 1973

GRAFAREND, E.W., and SCHAFFRIN, B.: ***Equivalence of Estimable Quantities and Invariants in Geodetic Networks***. Zeitschrift für Vermessungswesen (ZfV) 1976, S. 485-491

KOCH, K.R.: ***Parameterschätzung und Hypothesentests in linearen Modellen.*** Dümmler, Bonn 1980

KOWALSKY, H.J.: ***Lineare Algebra***. De Gruyter, Berlin 1979.

NIEMEIER, W.: ***Kongruenzprüfung in Geodätischen Netzen***, aus ***'Geodätische Netze in Landes- und Ingenieurvermessung'***. Editor: Pelzer, H., Konrad Wittwer, Stuttgart 1980

PELZER, H.: ***Zur Analyse geodätischer Deformationsmessungen***. Deutsche Geodätische Kommission (DGK), Reihe C: Dissertationen - Heft Nr. 164, München 1971

RAO, C.R., MITRA, S.K.: ***Generalized Inverse of Matrices and its Applications***. John Wiley & Sons, 1971

ZURMÜHL, R.: ***Matrizen***. Springer, 1964

PARAMETERSCHÄTZUNG IN LINEAREN MODELLEN MIT HILFE VON PROJEKTOREN

von Wilhelm CASPARY

SUMMARY

Under the criteria of unbiasedness and minimum variance the basic equations for estimation in linear models are developed. Starting from the standard Gauß-Markoff model the case of a singular coefficient matrix and of a singular dispersion structure is examined. The result is that for the adjusted observation vector, for certain linear functions of the parameters and for the variance factor general estimators exist, which are valid for all variations of the model. During the derivation of the equations the properties of projectors are used to simply and geometrically interpret the outcome.

ZUSAMMENFASSUNG

Unter den Kriterien Erwartungstreue und minimale Varianz werden die wichtigsten Schätzfunktionen in linearen Modellen entwickelt. Ausgehend von dem gewöhnlichen Gauß-Markoff-Modell werden Modelle mit singulärer Koeffizientenmatrix und mit singulärer Varianz-Kovarianz-Struktur betrachtet. Es zeigt sich dabei, daß für die ausgeglichenen Beobachtungen, für gewisse lineare Funktionen der Parameter und für die Varianz allgemeine Schätzer angegeben werden können, die für alle Modellvarianten gültig sind. Bei den Formelableitungen werden die Eigenschaften von Projektionsmatrizen ausgenutzt, um die Ergebnisse einfach und anschaulich zu interpretieren.

1. EINLEITUNG

In den letzten Jahrzehnten haben sich der Umfang und die Darstellungsweise der Methoden der Ausgleichungsrechnung beträchtlich gewandelt. Die in der mathematischen Statistik entwickelten Verfahren zum Testen von Hypothesen haben Eingang gefunden, und moderne Methoden der linearen Algebra erweisen sich als wirkungsvolles Hilfsmittel für Problemanalysen und zur Beschreibung des mathematischen Hintergrunds. Auch die Vielfalt der heute verwendeten mathematischen Modelle und die Art ihrer Formulierung sind überwiegend anderen Zweigen der Wissenschaft entlehnt.

Mit dieser "Modernisierung" der Ausgleichungsrechnung geht ein Wandel der Bezeichnungen einher, der es erleichtert, Parallelen zu verwandten Verfahren zu erkennen, der aber auch gelegentlich Unmut erzeugt, da die Aufgabe traditioneller Begriffe manchen zu eilfertig erfolgt. So könnte dieser Beitrag auch überschrieben sein: "Die Formeln der Ausgleichung nach vermittelnden Beobachtungen, abgeleitet mit Hilfe von Projektoren".

Nach der einleitenden Definition der betrachteten Modelle und der Festlegung der zu schätzenden Größen und Funktionen werden im dritten Abschnitt die wichtigsten Sätze über Projektionsmatrizen zusammengestellt und diskutiert. In den folgenden Abschnitten werden sodann Projektionsmatrizen benutzt, um in linearen Modellen, auf verschiedenen Stufen der Verallgemeinerung, Schätzfunktionen abzuleiten. Da sie auf einfache Weise eine geometrische Interpretation zulassen, sind Projektoren besonders geeignet, das Verständnis des mathematischen Hintergrunds der Methode der kleinsten Quadrate zu fördern. Die wesentlichen Entwicklungen, die bei dieser Betrachtungsweise der Ausgleichungsrechnung zum Tragen kommen, stammen von ***S.K. MITRA***, ***R.A. RAO*** und ***G. ZYSKIND***.

2. DAS GAUSS-MARKOFF-MODELL

Mit $\boldsymbol{\ell}$ als m-Vektor von Zufallsgrößen (Beobachtungen), $\boldsymbol{x}$ als n-Vektor von nicht stochastischen Parametern (Unbekannten), $m \geq n$, und $\boldsymbol{\mathcal{A}}$ als bekannter $m \times n$-Matrix von Koeffizienten lautet das Gauß-Markoff-Modell

$$\begin{aligned} \boldsymbol{\ell} &= \boldsymbol{\mathcal{A}x} + \varepsilon \;,\; E(\boldsymbol{\ell}) = \boldsymbol{\mathcal{A}x} \;, \\ V(\boldsymbol{\ell}) &= E\left(\varepsilon\varepsilon^t\right) = \sigma_0^2 \boldsymbol{Q}_1 \;. \end{aligned} \tag{2-1}$$

Hierin ist ε ein Vektor wahrer Fehler und $E(\cdot)$ bzw. $V(\cdot)$ stehen für den Erwartungswert bzw. die Varianz des in Klammern stehenden Arguments.

Definition 1 : Das oben eingeführte Modell sei bezeichnet als:

1.1: reguläres Gauß-Markoff-Modell, wenn $r(\mathcal{A}) = n$ und $r(\mathcal{Q}_l) = m$ gilt,

1.2: $\mathcal{A}$-singulräes Gauß-Markoff-Modell, wenn $r(\mathcal{A}) = p < n$ und $r(\mathcal{Q}_l) = m$ gilt,

1.3: $\mathcal{Q}$-singuläres Gauß-Markoff-Modell, wenn $r(\mathcal{A}) = n$ und $r(\mathcal{Q}_l) = q < m$ gilt,

1.4: allgemeines Gauß-Markoff-Modell, wenn $r(\mathcal{A}) = p < n$ und $r(\mathcal{Q}_l) = q < m$ gilt.

In den so spezifizierten Modellen werden die zu einem beobachteten Vektor $\boldsymbol{\ell}$ gehörenden "bestmöglichen" Schätzwerte für den Vektor $\boldsymbol{x}$, für lineare Funktionen von $\boldsymbol{x}$ und σ_0^2 gesucht, wobei $\mathcal{Q}_l$ als fest und a priori gegeben betrachtet wird. Zur Konkretisierung des Ausdrucks "bestmöglich" dienen die folgenden Definitionen.

Definition 2 : Als bester linearer Schätzer für eine Funktion
$y = \boldsymbol{b}^t\boldsymbol{x}$ der Parameter des Modells (2-1) sei die Schätzfunktion
$\hat{y} = \boldsymbol{c}^t\boldsymbol{\ell}$ bezeichnet, wenn sie

a) erwartungstreu ist
$E(\hat{y}) = E(\boldsymbol{c}^t\boldsymbol{\ell}) = \boldsymbol{b}^t\boldsymbol{x} \quad , \quad \forall\, \boldsymbol{x} \in \boldsymbol{E}^n$
und

b) minimale Varianz besitzt
$V(\hat{y}) = \sigma_0^2\, \boldsymbol{c}^t\mathcal{Q}_l\boldsymbol{c} = \min$.

Es wird sich im folgenden zeigen, daß die Bedingungen für beste Schätzbarkeit und die Menge der Funktionen, für die beste Schätzer existieren, von der Art des Modells gemäß Definition 1 abhängen.

Definition 3 : Als bester quadratischer Schätzer für die Varianz σ_0^2 sei eine Schätzfunktion der als normalverteilt angenommenen Beobachtungen $s_0^2 = \boldsymbol{\ell}^t\mathcal{B}\,\boldsymbol{\ell}$ bezeichnet, wenn sie

a) erwartungstreu ist
$E(s_0^2) = E(\ell^t \mathcal{B}\, \ell) = sp\,[\mathcal{B}\, V(\ell)] + x^t \mathcal{A}^t \mathcal{B}\, \mathcal{A}\, x = \sigma_0^2$

b) nicht von dem Parametervektor x abhängt (invariant ist) $\mathcal{A}^t \mathcal{B} = \mathcal{O}$ und

c) minimale Varianz besitzt
$V(\ell^t \mathcal{B}\, \ell) = 2\sigma_0^4\, sp[\mathcal{B}\, Q_1 \mathcal{B}\, Q_1] = \min$.

Hier handelt es sich um eine Spezialisierung der allgemeineren Varianzkomponentenschätzung, die z.B. in (***KOCH*** 1980, S. 205 ff.) behandelt ist. Die Kriterien a) bis c) sind erfüllt, wenn $s_0^2 = (sp[\mathcal{B}\, Q_1 \mathcal{B}\, Q_1])^{-1} \cdot \ell^t (\mathcal{B}\, Q_1 \mathcal{B})\, \ell$ gesetzt wird.

3. PROJEKTOREN

Jeder Vektorraum U kann in disjunkte Unterräume V und W zerlegt werden. Es gilt dann die direkte Summe $U = V \oplus W$, und jeder Vektor $u \in U$ ist eindeutig als Summe $u = v + w$ darstellbar, wobei $v \in V$ und $w \in W$ ist.

Definition 4 : Eine Projektion ist eine lineare Transformation, die den Vektorraum U auf den Unterraum V abbildet. Die Matrix $\mathcal{P}[V]$, die diese Abbildung erzeugt, ist genau dann ein Projektor, wenn

$$\mathcal{P}u = \mathcal{P}(v + w) = v \quad , \quad \forall\, u \in U \quad . \tag{3-1}$$

gilt.

Die Projektion ist eng mit der Zerlegung des Vektorraumes $U = V \oplus W$ verbunden. Aus der Definition folgt, daß der Projektor $\mathcal{P}$ quadratisch und idempotent ist

$$\mathcal{P}u = v \quad , \quad \mathcal{P}\mathcal{P}u = \mathcal{P}v = v \quad , \quad \forall\, u \in U \quad , \quad v \in V \quad , \tag{3-2}$$

außerdem gilt

$$\mathcal{P}w = \mathcal{O} \quad . \tag{3-3}$$

Der zu der Zerlegung $U = V \oplus W$ gemäß Def. 4 gehörende Projektor sei mit $\mathcal{P}_W[V]$ bezeichnet. Er projiziert den Vektorraum U auf den Unterraum V in Richtung des Unterraumes W. Es gilt daher $S(\mathcal{P}) = V$ und $N(\mathcal{P}) = W$, wobei $S(\cdot)$ den Spaltenraum und $N(\cdot)$ den Nullraum bezeichnet, der durch die Spaltenvektoren der in der Klammer stehenden Matrix aufgespannt wird.

Satz 1: Sei $\mathcal{P}_W[V]$ ein Projektor gemäß Def. 4, dann ist

$$\mathcal{I} - \mathcal{P}_W[V] = P_V[W] \tag{3-4}$$

der Projektor von U auf W in Richtung V.

Der Beweis folgt unmittelbar aus (3-1)

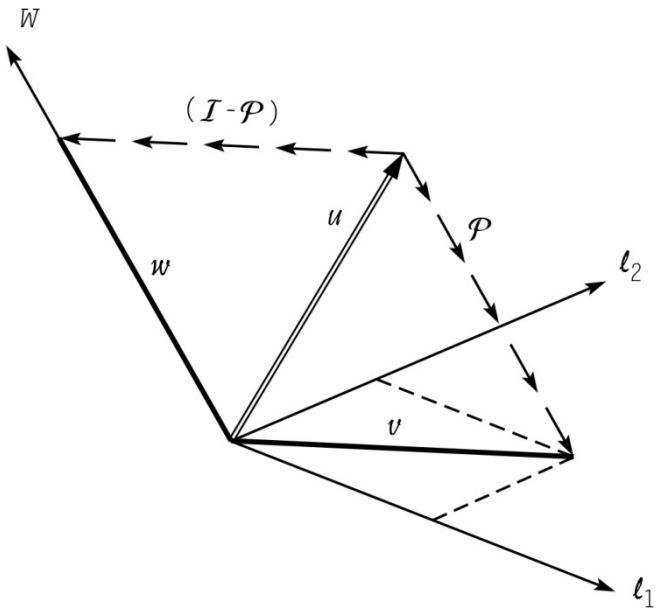
$$(\mathcal{I}-\mathcal{P})\,u = u - v = w \quad , \quad \forall\, u \in U \quad . \tag{3-5}$$

Außerdem gelten, wie leicht verifiziert werden kann, die Beziehungen

$$\mathcal{P}_W[V] + \mathcal{P}_V[W] = \mathcal{I} \tag{3-6}$$

$$\mathcal{P}_W[V] \cdot \mathcal{P}_V[W] = \mathcal{O} \tag{3-7}$$

In Abbildung 1 werden die Zusammenhänge für den E^3 geometrisch veranschaulicht.



$U = E^3 = V \oplus W$

$V \subset E^3$ (Ebene ℓ_1, ℓ_2)

$V \subset E^3$ (Ebene ℓ_1, ℓ_2)

$W \subset E^3$ (Raumgerade)

$u = v + w$

$\mathcal{P}u = v$

$\mathcal{P}w = \mathcal{O}$

$(\mathcal{I}-\mathcal{P})\,u = w$

Abbildung 1 Projektion in E^3

Wenn mit $V = S(\mathcal{A})$ und $W = S(\mathcal{B})$ zwei Unterräume gegeben sind, für die $U = V \oplus W$ gilt, so kann der damit verknüpfte Projektor explizit angegeben werden.

Satz 2: Sei $\mathcal{R}$ eine Matrix mit maximalem Rang, für die $\mathcal{R}\mathcal{B} = \mathcal{O}$ gilt, und sei $r(\mathcal{R}) = r(\mathcal{A})$, so lautet eine Darstellung des Projektors

$$P_B[\mathcal{A}] = \mathcal{A}(\mathcal{R}\mathcal{A})^-\mathcal{R} \tag{3-8}$$

Die generalisierte Inverse in (3-8) hat, wie in allen folgenden Beziehungen, lediglich die Gleichung

$$(\mathcal{R}\mathcal{A})\,(\mathcal{R}\mathcal{A})^-\,(\mathcal{R}\mathcal{A}) = (\mathcal{R}\mathcal{A}) \tag{3-9}$$

zu erfüllen. Wegen $\mathcal{P}\mathcal{B} = \mathcal{O}$ muß $\mathcal{P} = \mathcal{L}\mathcal{R}$ gelten. Wird dies in $\mathcal{P}\mathcal{A} = \mathcal{A}$ eingesetzt, so folgt $\mathcal{L}\mathcal{R}\mathcal{A} = \mathcal{A}$. Und mit $\mathcal{L} = \mathcal{A}\,(\mathcal{R}\mathcal{A})^-$ erhält man sofort (3-8).

Eine Wahl für die Matrix $\mathcal{R}$ ist zum Beispiel

$$\mathcal{R} = \left(\mathcal{I} - \mathcal{B}\,\mathcal{B}^-\right) . \tag{3-10}$$

Wie schon in (3-8) wird im folgenden vereinfachend die Schreibweise $P_B[\mathcal{A}]$ für $P_{S(\mathcal{B})}[S(\mathcal{A})]$ gewählt. Wenn die Vektorraumzerlegung aus dem Text hervorgeht, wird einfach $\mathcal{P}$ geschrieben.

Wenn $U = V \oplus W$ kein Vektorraum sondern ein Vektorunterraum des E^m ist, $U \subset E^m$, so gibt es Vektoren $x \in E^m$, für die $x^tU = x^tV = x^tW = \mathcal{O}$ gilt. Seien diese Vektoren in der Matrix $\mathcal{X}$ zusammengefaßt, so gilt

$$\left(P_W[V] + \mathcal{X}\,\mathcal{Y}\,\mathcal{X}^t\right)u = v \tag{3-11}$$

für jede Matrix $\mathcal{Y}$ passender Ordnung. $P_W[V] + \mathcal{X}\,\mathcal{Y}\,\mathcal{X}^t$ ist daher ein Projektor nach Definition 4, der jedoch keinesfalls eindeutig und nur für spezielle Matrizen $\mathcal{X}$ und $\mathcal{Y}$ idempotent ist.

Eine für die Schätzung in linearen Modellen besonders wichtige Gruppe sind die orthogonalen Projektoren. Sie sind bezüglich der orthogonalen Zerlegung des Vektorraumes U definiert: $U = V \oplus V^\perp$. Die durch $V^\perp$ gegebene Projektionsrichtung steht senkrecht auf V. Daraus folgt, daß für jede Vektorzerlegung $u = v + w$ mit $u \in U$, $v \in V$ und $w \in V^\perp$ das Produkt $v^tw = 0$ wird.

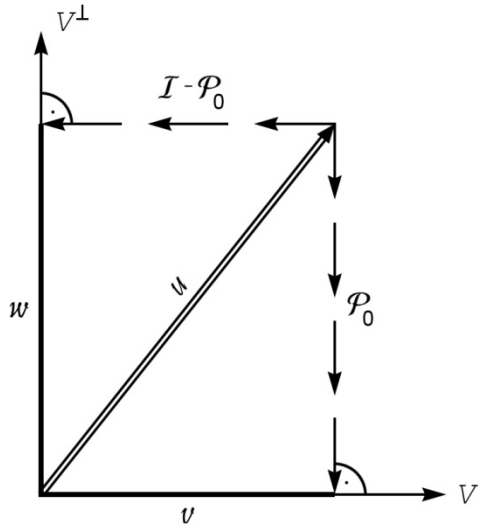


Abbildung 2 Orthogonale Projektion in E^2

Satz 3: Sei mit $V = S(\mathcal{A})$ und $V^{\perp} = S^{\perp}(\mathcal{A}) = S(\mathcal{B})$ eine orthogonale Zerlegung des E^m gegeben, so ist

$$P_0[\mathcal{A}] = \mathcal{A}\left(\mathcal{A}^t\mathcal{A}\right)^{-}\mathcal{A}^t \tag{3-12}$$

der orthogonale Projektor auf $S(\mathcal{A})$.

Wegen $S^{\perp}(\mathcal{A}) = S(\mathcal{B})$ gilt $\mathcal{A}^t\mathcal{B} = \mathcal{O}$. Ferner ist $r(\mathcal{A}) = r\left(\mathcal{A}^t\mathcal{A}\right)$. Daher kann in (3-8) für $\mathcal{R}$ die Matrix $\mathcal{A}^t$ eingesetzt werden. Da $\mathcal{A}\left(\mathcal{A}^t\mathcal{A}\right)^{-}\mathcal{A}^t$ unabhängig von der Wahl der g-Inversen ist, ist $P_0[\mathcal{A}]$ durch (3-12) eindeutig bestimmt.

Satz 4: P_0 ist genau dann ein orthogonaler Projektor, wenn P_0 symmetrisch und idempotent ist, so daß

$$P_0 = P_0^t = P_0^2 \ . \tag{3-13}$$

Aus (3-2) folgt $P_0 u = v \in S(\mathcal{A})\ \forall\ u \in E^m$ und aus (3-5) erhält man $\left(I - P_0\right)u = w \in S(\mathcal{B})\ \forall\ u \in E^m$. Da nach Voraussetzung $S(\mathcal{B}) = S^{\perp}(\mathcal{A})$ gilt, ist $v^t w = 0\ \forall\ v \in S(\mathcal{A})$, $w \in S(\mathcal{B})$. Setzt man nun die Projektionen ein, so ist $v^t w = u^t P_0^t (I - P_0)\, u = 0\ \forall\ u \in E^m$. Daher muß $P_0^t = P_0^t P_0$ und $P_0 = \left(P_0^t P_0\right)^t = P_0^t P_0$ gelten, d.h. P_0 muß symmetrisch und idempotent sein. Wenn andererseits $P_0 = P_0^t = P_0^2$ gilt, so folgt $P_0^t = P_0^t P_0$ und damit $u^t P_0^t \left(I - P_0\right) u = 0\ \forall\ u \in E^m$.

Eine Erweiterung des Projektionsbegriffs für Vektorräume, in denen die Länge durch die Norm $\|u\|_M = \left(u^t \mathcal{M} u\right)^{1/2}$, $\mathcal{M}$ symmetrisch, gemessen wird, ist einfach möglich. Die Analogie wird besonders deutlich, wenn folgende Bezeichnungen eingeführt werden.

Definition 5 : Eine quadratische Matrix $\mathcal{P}$ sei bezeichnet als

(a) $\mathcal{M}$-symmetrisch, wenn $\mathcal{MP} = (\mathcal{MP})^t$,

(b) $\mathcal{M}$-idempotent, wenn $\mathcal{MP} = \mathcal{MP}^2$ gilt.

Die Zerlegung des Vektorraumes $U = V \oplus W$ in eine direkte Summe sei bezeichnet als

(c) $\mathcal{M}$-orthogonal, wenn $v^t \mathcal{M} w = 0 \; \forall \; v \in V, \; w \in W$.

Wenn U ein Vektorraum ist, so kann in Anwendung von (c) $W = V_M^\perp$ als $\mathcal{M}$-orthogonales Komplement von V bezeichnet werden.

Satz 5: $\mathcal{P}_{M0}[V]$ ist genau dann ein $\mathcal{M}$-orthogonaler Projektor auf V , wenn er $\mathcal{M}$-symmetrisch und $\mathcal{M}$-idempotent ist:

$$\mathcal{MP} = (\mathcal{MP})^t \quad ; \quad \mathcal{MP} = \mathcal{MP}^2$$

In Analogie zur Beweisführung für Satz 4 muß

$$v^t \mathcal{M} w = u^t \mathcal{P}^t \mathcal{M} \left(\mathcal{I} - \mathcal{P}\right) u = 0 \; \forall \; u \in E^m \tag{3-14}$$

gelten. Daraus folgt unmittelbar $\mathcal{P}^t \mathcal{M} = \mathcal{P}^t \mathcal{MP} = \left(\mathcal{P}^t \mathcal{M}\right)^t = \mathcal{MP}$ und $\mathcal{P}^t \mathcal{M} = \mathcal{P}^t \mathcal{MPP} = \mathcal{MP} = \mathcal{MPP}$, also die Notwendigkeit von $\mathcal{M}$-Symmetrie und $\mathcal{M}$-Idempotenz. Daß diese Eigenschaften auch hinreichend sind, ergibt sich durch Einsetzen in (3-14).

Wird der Satz 2 auf die $\mathcal{M}$-orthogonale Zerlegung des Vektorraumes angewandt, so erhält man mit $V = S(\mathcal{A})$ und $V_M^\perp = S(\mathcal{B})$ die Beziehung

$$\mathcal{A}^t \mathcal{M} \mathcal{B} = \mathcal{O} \; , \tag{3-15}$$

die zeigt, daß in (3-8) $\mathcal{R} = \mathcal{A}^t \mathcal{M}$ zu setzen ist. Daraus folgt

$$\mathcal{P}_{M0}[\mathcal{A}] = \mathcal{A}\left(\mathcal{A}^t \mathcal{M} \mathcal{A}\right)^- \mathcal{A}^t \mathcal{M} \tag{3-16}$$

Satz 6: Unter der $\mathcal{M}$-orthogonalen Projektion des E^m auf den $S(\mathcal{A})$ gilt

$$\left\|u - \mathcal{P}u\right\|_M \leq \left\|u - a\right\|_M \; \forall \; u \in E^m, \; a \in S(\mathcal{A}) \; . \tag{3-17}$$

Wegen $(u-\mathcal{P}u) \in S_M^{\perp}(\mathcal{A})$ und $a \in S(\mathcal{A})$ gilt für alle u und a entsprechend (3-15) $a^t\mathcal{M}(u-\mathcal{P}u) = 0$. Schreibt man nun $u - a = (u - \mathcal{P}u) - (a - \mathcal{P}u)$ so folgt

$$\|u-a\|_M^2 = \|u-\mathcal{P}u\|_M^2 + \|a-\mathcal{P}u\|_M^2 - 2(u-\mathcal{P}u)^t\mathcal{M}(a-\mathcal{P}u) \quad , \tag{3-18}$$

d.h. (3-17) ist genau dann erfüllt, wenn die Bilinearform auf der rechten Seite von (3-18) verschwindet. Dies ist aber unter der $\mathcal{M}$-orthogonalen Projektion gesichert, da $(a - \mathcal{P}u) \in S(\mathcal{A})$.

4. SCHÄTZUNG IM REGULÄREN GAUSS-MARKOFF-MODELL

In dem in (2-1) formulierten Modell mit den Spezifikationen nach Definition 1.1

$$\ell = \mathcal{A}x + \varepsilon \quad , \quad V(\ell) = \sigma_0^2 Q_1$$

werden die Modellparameter x nach der Methode der kleinsten Quadrate geschätzt. Das als bekannt vorausgesetzte Ergebnis lautet

$$\begin{aligned} \hat{x} &= \left(\mathcal{A}^t Q_1^{-1}\mathcal{A}\right)^{-1}\mathcal{A}^t Q_1^{-1}\ell \\ V(\hat{x}) &= \sigma_0^2\left(\mathcal{A}^t Q_1^{-1}\mathcal{A}\right)^{-1} = \sigma_0^2 Q_{\hat{x}} \quad . \end{aligned} \tag{4-1}$$

Wird die Gleichung für $\hat{x}$ links mit $\mathcal{A}$ multipliziert, so folgt

$$\mathcal{A}\hat{x} = \mathcal{A}\left(\mathcal{A}^t Q_1^{-1}\mathcal{A}\right)^{-1}\mathcal{A}^t Q_1^{-1}\ell = \ell + v = \hat{\ell} \tag{4-2}$$

oder im Hinblick auf (3-16)

$$\hat{\ell} = \mathcal{P}_{Q^{-1}0}[\mathcal{A}]\,\ell \quad . \tag{4-3}$$

Der Vektor der ausgeglichenen Beobachtungen $\hat{\ell}$ wird durch Q_1^{-1}-orthogonale Projektion des Beobachtungsvektors ℓ auf den Spaltenraum von $\mathcal{A}$ erhalten. Die Methode der kleinsten Quadrate bewirkt also eine Q_1^{-1}-orthogonale Zerlegung des Vektors $\ell \in E^m$ in die Summe der Vektoren $\hat{\ell} \in S(\mathcal{A})$ und $\ell - \hat{\ell} = -v \in S_{Q^{-1}}^{\perp}(\mathcal{A})$. Aus den Eigenschaften der Q_1^{-1}-orthogonalen Projektion folgt sofort, daß die Zerlegung eindeutig ist und folgende Beziehungen gelten:

$$\hat{\ell}^t Q_l^{-1} v = 0 \quad \forall \ \hat{\ell} \in S(A)$$

$$A^t Q_l^{-1} v = 0 \tag{4-4}$$

$$(\ell - \hat{\ell})^t Q_l^{-1} (\ell - \hat{\ell}) = \min .$$

Die Suche nach besten linearen Schätzern gemäß Definition 2 erfolgt in zwei Schritten. Die Forderung der Erwartungstreue führt für $y = b^t x = c^t \ell$ zu

$$E(c^t \ell) = c^t E(\ell) = c^t A x = b^t x \tag{4-5}$$

und damit zu der Bedingung

$$c^t A = b^t \quad \text{bzw.} \quad b \in S(A^t) . \tag{4-6}$$

Da nach Modelldefinition $r(A) = n$ ist, liegen alle n-Vektoren b im $S(A^t)$. Daraus folgt, daß jede lineare Funktion $y = b^t x$ durch $c^t \ell$ erwartungstreu schätzbar ist. Aus (4-6) erhält man für c die allgemeine Lösung

$$c = (A^t)^- b + (I - (A^t)^- A^t) z \quad , \quad z \text{ beliebig.} \tag{4-7}$$

Um unter allen erwartungstreuen Schätzern $c^t \ell$ jenen mit minimaler Varianz zu erhalten, ist $V(c^t \ell) = \sigma_0^2 c^t Q_l c$ zu minimieren, und zwar durch geeignete Wahl von $(A^t)^-$ und z. Wegen der in den nächsten Abschnitten auftretenden Besonderheiten ist es zweckmäßig, den Vektor c mit Hilfe der Lagrangeschen Funktion

$$L = c^t Q_l c + 2\lambda^t (A^t c - b) \tag{4-8}$$

direkt zu bestimmen. Die Ableitungen von L nach c und λ werden null gesetzt und führen auf das konsistente Gleichungssystem

$$\begin{bmatrix} Q_l & A \\ A^t & 0 \end{bmatrix} \begin{bmatrix} c \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \tag{4-9}$$

mit der Lösung

$$\begin{bmatrix} c \\ \lambda \end{bmatrix} = \begin{bmatrix} Q_l^{-1} - G A Q_l^{-1} & G \\ G^t & -N^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ b \end{bmatrix} , \tag{4-10}$$

in der zur Abkürzung

$$\mathcal{N} = \mathcal{A}^t \mathcal{Q}_l^{-1} \mathcal{A} \quad \text{und} \quad \mathcal{G} = \mathcal{Q}_l^{-1} \mathcal{A} \mathcal{N}^{-1} \tag{4-11}$$

gesetzt wurde. Es erweist sich also

$$\boldsymbol{c} = \mathcal{G} \boldsymbol{b} = \mathcal{Q}_l^{-1} \mathcal{A} \left(\mathcal{A}^t \mathcal{Q}_l^{-1} \mathcal{A} \right)^{-1} \boldsymbol{b} \tag{4-12}$$

als der gesuchte Vektor für die beste lineare Schätzung von $\boldsymbol{b}^t \boldsymbol{x}$

$$\hat{y} = \boldsymbol{b}^t \mathcal{G}^t \boldsymbol{\ell} = \boldsymbol{b}^t \hat{\boldsymbol{x}} = \boldsymbol{c}^t \hat{\boldsymbol{\ell}} \ . \tag{4-13}$$

Für die Varianz erhält man

$$V(\hat{y}) = \sigma_0^2 \, \boldsymbol{c}^t \mathcal{Q}_l \boldsymbol{c} = \sigma_0^2 \, \boldsymbol{b}^t \mathcal{Q}_{\hat{x}} \boldsymbol{b} \ . \tag{4-14}$$

Gleichung (4-13) ist eine Verallgemeinerung von (4-1) und (4-2). Soll nämlich $\boldsymbol{y} = \mathcal{I} \boldsymbol{x} = \boldsymbol{x}$ nach (4-13) geschätzt werden, so ist $\boldsymbol{b}$ durch die Einheitsmatrix $\mathcal{I}$ zu ersetzen, womit sofort (4-1) folgt. Ist dagegen $\boldsymbol{y} = \mathcal{A} \boldsymbol{x} = \boldsymbol{\ell}$ zu schätzen, so findet man mit $\mathcal{A}$ für $\boldsymbol{b}^t$ in (4-13) das Ergebnis (4-2). Diese Identitäten zeigen die Äquivalenz der Methode der kleinsten Quadrate und der besten linearen Schätzung.

Als Schätzwert für σ_0^2 wird bei der Methode der kleinsten Quadrate der Schätzer

$$s_0^2 = \boldsymbol{v}^t \mathcal{Q}_l^{-1} \boldsymbol{v} \, / \, (m - n) \tag{4-15}$$

benutzt. Es soll nun noch gezeigt werden, daß s_0^2 ein bester quadratischer Schätzer gemäß Definition 3 ist.

Aus $\boldsymbol{v} = \hat{\boldsymbol{\ell}} - \boldsymbol{\ell} = (\mathcal{P} - \mathcal{I}) \boldsymbol{\ell}$ mit $\mathcal{P}$ nach (4-3) folgt

$$s_0^2 = \boldsymbol{\ell}^t \mathcal{Q}_l^{-1} (\mathcal{I} - \mathcal{P}) \boldsymbol{\ell} \, / \, (m - n) = \boldsymbol{\ell}^t \mathcal{B} \boldsymbol{\ell} \tag{4-16}$$

$$E(s_0^2) = E(\boldsymbol{\ell}^t \mathcal{B} \boldsymbol{\ell}) = \sigma_0^2 \, sp[\mathcal{B} \mathcal{Q}_l] + \boldsymbol{x}^t \mathcal{A}^t \mathcal{B} \mathcal{A} \boldsymbol{x} \ . \tag{4-17}$$

Wird nun in (4-17)

$$\mathcal{B} = \mathcal{Q}_l^{-1} (\mathcal{I} - \mathcal{P}) \, / \, (m - n) \tag{4-18}$$

eingesetzt, so findet man nach Ausmultiplizieren, daß $E(s_0^2) = \sigma_0^2$ erfüllt

ist. Fernerzeigt sich, daß wegen $\mathcal{A}^t\mathcal{B} = \mathcal{O}$ der Schätzer invariant ist. Daß s_0^2 nach (4-16) auch minimale Varianz besitzt, läßt sich ebenfalls durch Einsetzen zeigen,

$$\mathrm{sp}(\mathcal{B}Q_1\mathcal{B}Q_1) = \frac{1}{(m-n)}$$

$$\ell^t\mathcal{B}\,Q_1\mathcal{B}\,\ell = \ell^t\mathcal{B}\,\ell\,/\,(m-n)$$

$$s_0^2 = (\mathrm{sp}[\mathcal{B}Q_1\mathcal{B}Q_1])^{-1}\ell^t\mathcal{B}\,Q_1\mathcal{B}\,\ell = \ell^t\mathcal{B}\,\ell \qquad (4\text{-}19)$$

5. SCHÄTZUNG IM $\mathcal{A}$-SINGULÄREN GAUSS-MARKOFF-MODELL

Im ersten Schritt der Verallgemeinerung des in (2-1) formulierten Modells

$$\ell = \mathcal{A}x + \varepsilon\ ,\ V(\ell) = \sigma_0^2 Q_1$$

soll gemäß Definition 1.2 angenommen werden, daß der Rang der $m \times n$-Matrix $\mathcal{A}$ kleiner als n ist,

$$r(\mathcal{A}) = p < n\ .$$

Bildet man nun nach der Methode der kleinsten Quadrate die Normalgleichungen

$$\mathcal{A}^tQ_1^{-1}\mathcal{A}\,\hat{x} = \mathcal{A}^tQ_1^{-1}\ell\ , \qquad (5\text{-}1)$$

so können diese bekanntlich nicht in gewohnter Weise aufgelöst werden, da wegen $r\left(\mathcal{A}^tQ_1^{-1}\mathcal{A}\right) = r(\mathcal{A}) < n$ die Inverse von $\mathcal{A}^tQ_1^{-1}\mathcal{A}$ nicht existiert. Die Gleichung (5-1) hat keinen eindeutigen Lösungsvektor $\hat{x}$ sondern eine Lösungsmenge

$$L = \left\{\hat{x}\,,\ \hat{x} = \mathcal{N}^-\mathcal{A}^tQ_1^{-1}\ell + \left(\mathcal{I} - \mathcal{N}^-\mathcal{N}\right)z\right\} \qquad (5\text{-}2)$$

mit $\mathcal{N} = \mathcal{A}^tQ_1^{-1}\mathcal{A}$, $\mathcal{N}\mathcal{N}^-\mathcal{N} = \mathcal{N}$ und $z \in E^n$ beliebig.

Im singulären Gauß-Markoff-Modell ist die Frage nach dem Parametervektor x offenbar nicht sinnvoll. Es kann nur nach der Lösungsmenge L gefragt werden, die durch alle Vektoren, die (5-1) befriedigen, gegeben ist. L enthält genau $n-p+1$ linear unabhängige Vektoren $\hat{x}$. Für praktische Zwecke kann es sinnvoll sein, ein Element $\hat{x} \in L$ unter ganz bestimmten Gesichtspunkten auszuwählen, die dann am einfachsten in $n-p$ geeignete

Bedingungsgleichungen $\mathcal{R}\hat{x} - g = 0$ gekleidet werden. Einzelheiten dazu sind in ***CASPARY*** 1978 behandelt.

Wird (5-2) linksseitig mit $\mathcal{A}$ multipliziert, so erhält man den eindeutigen Vektor

$$\mathcal{A}\hat{x} = \mathcal{A}\mathcal{N}^{-}\mathcal{A}^{t}Q_{l}^{-1}\ell = \ell + v = \hat{\ell} \quad . \tag{5-3}$$

Ein Vergleich mit (3-16) zeigt, daß $\mathcal{A}\mathcal{N}^{-}\mathcal{A}^{t}Q_{l}^{-1} = \mathcal{P}_{Q^{-1}0}[\mathcal{A}]$ der Q_{l}^{-1}-orthogonale Projektor auf $S(\mathcal{A})$ ist. Auch im $\mathcal{A}$-singulären Modell bewirkt also die Methode der kleinsten Quadrate eine eindeutige Q_{l}^{-1}-orthogonale Zerlegung des Beobachtungsvektors ℓ . Die in (4-4) formulierten Beziehungen gelten daher unverändert im $\mathcal{A}$-singulären Modell.

Die Überlegungen zur Erwartungstreue einer linearen Funktion $y = b^{t}x$ in Abschnitt 4 sind unabhängig von $r(\mathcal{A})$ durchgeführt worden, so daß das Ergebnis $b^{t} \in S(\mathcal{A}^{t})$ seine Gültigkeit behält. Da $r(\mathcal{A}) = p < n$ vorausgesetzt ist, liegen allerdings nicht mehr alle n-Vektoren im $S(\mathcal{A}^{t})$, so daß sich eine Beschränkung der Zahl der zulässigen Vektoren b ergibt. Im $\mathcal{A}$-singulären Modell (2-1) sind nur die Funktionen $y = b^{t}x$ erwartungstreu schätzbar, für die $\mathcal{N}\mathcal{N}^{-}b = b$ bzw. $\mathcal{A}^{t}(\mathcal{A}^{t})^{-}b = b$ gilt.

Die Forderung nach minimaler Varianz führt genau wie im 4. Kapitel auf das Gleichungssystem für c

$$\begin{bmatrix} Q_{l} & \mathcal{A} \\ \mathcal{A}^{t} & 0 \end{bmatrix}\begin{bmatrix} c \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \quad , \tag{5-4}$$

das nun jedoch einen Rangdefekt hat. Nach ***PRINGLE/RAYNER*** 1971, S. 50 ff., hat dieses System die Lösung

$$\begin{bmatrix} c \\ \lambda \end{bmatrix} = \begin{bmatrix} Q_{l}^{-1} - \mathcal{G}\mathcal{A}Q_{l}^{-1} & \mathcal{G} \\ \mathcal{G}^{t} & -\mathcal{N}^{-} \end{bmatrix}\begin{bmatrix} 0 \\ b \end{bmatrix} , \tag{5-5}$$

wobei $\mathcal{N}^{-} = (\mathcal{A}^{t}Q_{l}^{-1}\mathcal{A})^{-}$ eine beliebige g-Inverse von $\mathcal{N}$ ist und $\mathcal{G} = Q_{l}^{-1}\mathcal{A}\mathcal{N}^{-}$ gesetzt wurde. Für den gesuchten Vektor erhält man analog zu (4-12) die Lösung

$$c = \mathcal{G}b = Q_{l}^{-1}\mathcal{A}(\mathcal{A}^{t}Q_{l}^{-1}\mathcal{A})^{-}b \quad , \tag{5-6}$$

mit der der beste lineare Schätzer lautet

$$\hat{y} = \mathcal{b}^t \mathcal{G}^t \ell = \mathcal{b}^t (\mathcal{N}^-)^t \mathcal{A}^t \mathcal{Q}_l^{-1} \ell \quad . \tag{5-7}$$

Da nach (4-6) $\mathcal{b}^t = c^t \mathcal{A}$ gilt, folgt aus (5-7)

$$\hat{y} = c^t \mathcal{A} \mathcal{G}^t \ell = \mathcal{b}^t \hat{x} = c^t \hat{\ell} \quad . \tag{5-8}$$

Das Produkt $\mathcal{A} \mathcal{G}^t$ ist $\mathcal{Q}_l^{-1}$-symmetrisch. Dies folgt aus der Invarianz des Ausdrucks $\mathcal{A} \mathcal{N}^- \mathcal{A}^t$ bezüglich der benutzten g-Inversen. Wegen $\mathcal{N}\mathcal{N}^-\mathcal{N} = \mathcal{N} = \mathcal{N}(\mathcal{N}^-)^t\mathcal{N}$ ist auch $(\mathcal{N}^-)^t$ g-Inverse von $\mathcal{N}$, so daß ebenfalls in (5-7) $(\mathcal{N}^-)^t$ durch $\mathcal{N}^-$ ersetzt werden darf. Die Eindeutigkeit des Schätzers $\hat{y}$ kann in (5-8) abgelesen werden, da $\hat{\ell}$ nach (5-3) und ebenfalls c eindeutig sind. Die Varianz von $\hat{y}$ lautet

$$V(\hat{y}) = \sigma_0^2 \, c^t \mathcal{Q}_l c = \sigma_0^2 \, \mathcal{b}^t \mathcal{N}^- \mathcal{b} \quad .$$

Gleichung (5-7) ist eine Verallgemeinerung von (5-3), die man erhält, indem man $\mathcal{b}$ durch $\mathcal{A}^t$ ersetzt.

Der beste quadratische Schätzer für σ_0^2 wird ganz entsprechend der Vorgehensweise in Kapitel 4 abgeleitet. Mit $\mathcal{B} = \mathcal{Q}_l^{-1}(\mathcal{I} - \mathcal{P}) / (m - p)$, wobei $\mathcal{P} = \mathcal{A}\mathcal{N}^-\mathcal{A}^t\mathcal{Q}_l^{-1}$ der $\mathcal{Q}_l^{-1}$-orthogonale Projektor nach (5-3) ist, erhält man

$$\mathrm{sp}[\mathcal{B}\mathcal{Q}_l\mathcal{B}\mathcal{Q}_l] = \frac{1}{(m - p)}$$

$$\ell^t \mathcal{B} \mathcal{Q}_l \mathcal{B} \ell \qquad \ell^t \mathcal{B} \ell \, / \, (m - p)$$

und schließlich

$$s_0^2 = \ell^t \mathcal{B} \ell = \frac{\ell^t \mathcal{Q}_l^{-1} (\mathcal{I} - \mathcal{P}) \ell}{m - p} \quad . \tag{5-9}$$

6. DAS ALLGEMEINE GAUSS-MARKOFF-MODELL

Die allgemeine Form des in (2-1) formulierten Gauß-Markoff-Modells

$$\ell = \mathcal{A}x + \varepsilon \;, \quad V(\ell) = \sigma_0^2 \mathcal{Q}_l = \Sigma_l$$

ist in Definition 1.4 spezifiziert. Es wird angenommen, daß sowohl die Koeffizientenmatrix $\mathcal{A}$ als auch die Kofaktorenmatrix $\mathcal{Q}_l$ einen Rangdefekt besitzen:

$$r(\mathcal{A}) = p < n \quad , \quad r(\mathcal{Q}_l) = q < m$$

Eine gesonderte Behandlung des $\mathcal{Q}$-singulären Falles gemäß Definition 1.3 ist entbehrlich, da die dort benötigten Schätzfunktionen sofort aus den Ergebnissen dieses Kapitels abgelesen werden können.

Will man im allgemeinen Modell zur Ableitung von Schätzern die Methode der kleinsten Quadrate anwenden, so stößt man auf die Schwierigkeit, daß die Berechnung der Gewichtsmatrix $\mathcal{G}$ aus der singulären Kofaktorenmatrix $\mathcal{Q}_l$ nicht so ohne weiteres möglich ist. Es ist also zu untersuchen, welche Matrix $\mathcal{G}$ zur Bildung der zu minimierenden quadratischen Form zu verwenden ist.

Ehe auf Lösungsmöglichkeiten für dieses Problem eingegangen wird, soll eine Analyse des Modells Klarheit darüber verschaffen, welche Vektorraumbeziehungen vorliegen.

Aus $\ell = \mathcal{A}x + \varepsilon$ folgt, daß $(\ell - \varepsilon) \in S(\mathcal{A})$ gilt. Da nur mit dem Modell verträgliche Schätzverfahren betrachtet werden sollen, muß auch für die ausgeglichene Beobachtung gelten: $\ell + v = \hat{\ell} \in S(\mathcal{A})$. Für den wahren Fehler ε gilt definitionsgemäß $\varepsilon \in S(\Sigma_l)$, so daß man für den Beobachtungsvektor $\ell \in S(\mathcal{A} : \mathcal{Q}_l)$ erhält. Das Modell (2-1) belegt also den Vektorraum $S(\mathcal{A} : \mathcal{Q}_l) = S(\mathcal{T})$, der im allgemeinen Fall ein Unterraum des E^m ist. Der Modellraum $S(\mathcal{T})$ wird durch die Spalten der Matrizen $\mathcal{A}$ und $\mathcal{Q}_l$ aufgespannt

$$S(\mathcal{A} : \mathcal{Q}_l) = S(\mathcal{A}) + S(\mathcal{Q}_l) = S(\mathcal{T}) \quad . \tag{6-1}$$

Weitere Klarheit bringt eine orthogonale Zerlegung des E^m entsprechend *CASPARY* 1983 mit den im folgenden definierten Matrizen bzw. Vektorunterräumen

$$S(\mathcal{K}) = S(\mathcal{A}) \cap S^{\perp}(\mathcal{Q}_l) \quad , \quad \mathcal{K}^t \mathcal{Q}_l = \mathcal{O} \tag{6-2}$$

$$S(\mathcal{L}) = S(\mathcal{A}) \cap S(\mathcal{Q}_l) \quad , \quad \mathcal{L}^t \mathcal{K} = \mathcal{O} \tag{6-3}$$

$$S(\mathcal{M}) = S(\mathcal{Q}_l) \cap S^{\perp}(\mathcal{A}) \quad , \quad \mathcal{L}^t \mathcal{M} = \mathcal{O} \quad , \quad \mathcal{M}^t \mathcal{A} = \mathcal{O} \tag{6-4}$$

$$S(\mathcal{N}) = S^{\perp}(\mathcal{Q}_l) \cap S^{\perp}(\mathcal{A}) \quad , \quad \mathcal{N}^t \mathcal{A} = \mathcal{O} \quad , \quad \mathcal{N}^t \mathcal{Q}_l = \mathcal{O} \tag{6-5}$$

$$S(\mathcal{K}) \oplus S(\mathcal{L}) \oplus S(\mathcal{M}) \oplus S(\mathcal{N}) = E^m \tag{6-6}$$

$$S(\mathcal{K}) \oplus S(\mathcal{L}) = S(\mathcal{A}) \quad , \quad S(\mathcal{L}) \oplus S(\mathcal{M}) = S(\mathcal{Q}_l) \tag{6-7}$$

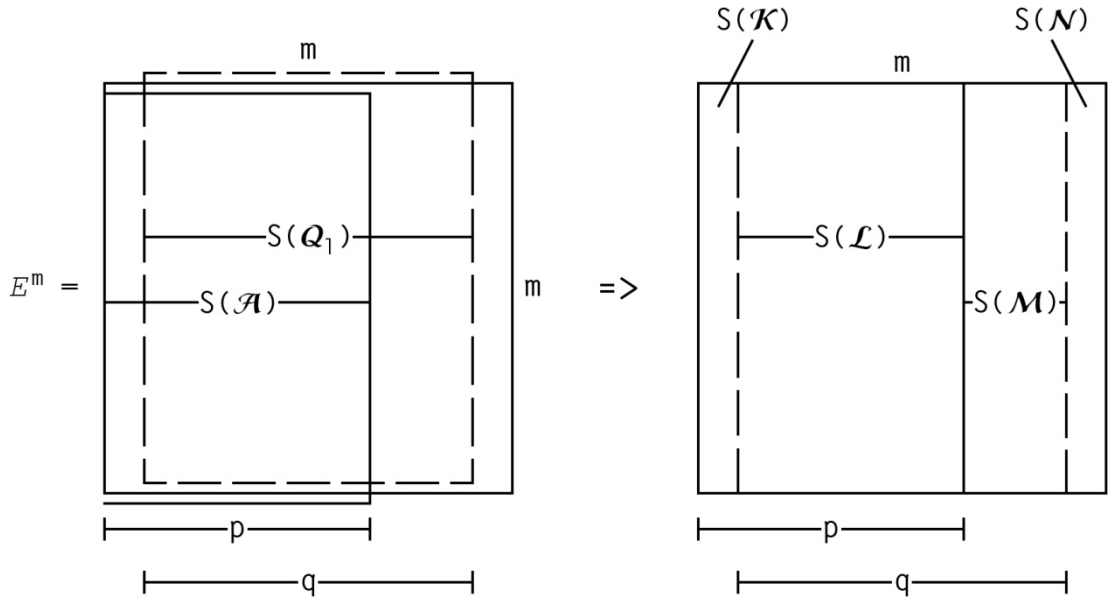


Abbildung 3 Vektorraumbelegung im allgemeinen Gauß-Markoff-Modell

Aus $\mathcal{K}^t\ell = \mathcal{K}^t\mathcal{A}x + \mathcal{K}^t\varepsilon$ folgt wegen $\mathcal{K}^t\varepsilon = \mathcal{O}$, daß die Parameter die Bedingung $\mathcal{K}^t\mathcal{A}x = \mathcal{K}^t\ell$ erfüllen müssen, deren rechte Seite erst nach Ausführung der Beobachtungen angegeben werden kann. Diese Bedingung verschwindet nur in dem Sonderfall $S(\mathcal{A}) \subset S(\mathcal{Q}_1)$, da dann $\mathcal{K} \equiv \mathcal{O}$ ist. Ferner müssen die Beobachtungen die Gleichung $\mathcal{N}^t\ell = \mathcal{O}$ erfüllen, mit der geprüft werden kann, ob die Beobachtungen modellkonform sind.

Nach ***MITRA/RAO*** 1968 erhält man eine Darstellung der Matrix $\mathcal{X}$ für $S(\mathcal{X}) = S(\mathcal{Y}) \cap S(\mathcal{Z})$ durch die Gleichung

$$\mathcal{X} = \mathcal{Y}\left(\mathcal{Y}^t\mathcal{Z}^\perp\right)^\perp \quad \text{mit} \quad r(\mathcal{X}) = r(\mathcal{Y}) + r(\mathcal{Z}) - r(\mathcal{Y} : \mathcal{Z}) \quad , \tag{6-8}$$

wobei unter $\mathcal{A}^\perp$ eine Matrix verstanden wird, deren Spalten das orthogonale Komplement zum $S(\mathcal{A})$ bilden.

Die Anwendung dieser Beziehung ergibt zum Beispiel mit

$$\mathcal{M} = \mathcal{Q}_1(\mathcal{Q}_1\mathcal{A})^\perp \tag{6-9}$$

eine Möglichkeit zur Bestimmung einer Matrix $\mathcal{M}$, deren Spalten den Unterraum $S(\mathcal{M})$ aufspannen. Mit der orthogonalen Zerlegung (6-2) bis (6-5) erhält man folgende Darstellungsmöglichkeiten für die in (6-1) eingeführte Modellraummatrix $\mathcal{T}$,

$$\mathcal{T} = \mathcal{Q}_1 + \mathcal{K}\mathcal{E}\mathcal{K}^t \tag{6-10}$$

$$\mathcal{T} = \mathcal{Q}_1 + \mathcal{A}\mathcal{F}\mathcal{A}^t \tag{6-11}$$

$$\mathcal{T} = \mathcal{K}\mathcal{X}\mathcal{K}^t + \mathcal{L}\mathcal{Y}\mathcal{L}^t + \mathcal{M}\mathcal{Z}\mathcal{M}^t \quad , \tag{6-12}$$

bei denen $\mathcal{E}$, $\mathcal{F}$, $\mathcal{X}$, $\mathcal{Y}$ und $\mathcal{Z}$ Matrizen passender Ordnung sind, die lediglich die offensichtlichen Rangbedingungen, wie z.B. $r(\mathcal{K}\mathcal{E}\mathcal{K}^t) = r(\mathcal{K})$, erfüllen müssen.

Die Schätzaufgabe kann nun so formuliert werden: Der Beobachtungsvektor $\ell \in S(\mathcal{T})$ ist $\mathcal{G}$-orthogonal in die Summe der Vektoren $\hat{\ell} = \ell + v \in S(\mathcal{A})$ und $-v \in S(\mathcal{M})$ zu zerlegen. Daher muß analog (4-4) gelten

$$\hat{\ell}^t \mathcal{G} v = 0 \;\forall\; \hat{\ell} \in S(\mathcal{A}) \quad \text{und} \quad -v \in S(\mathcal{M}) \quad . \tag{6-13}$$

Gleichung (6-13) kann mit $\hat{\ell} = \mathcal{A}y$ und $-v = \mathcal{M}z$ umgeformt werden in $y^t \mathcal{A}^t \mathcal{G} \mathcal{M} z = 0 \;\forall\; y \in E^n$ und $z \in E^{r(M)}$. Dies ist gleichbedeutend mit

$$\mathcal{A}^t \mathcal{G} \mathcal{M} = \mathcal{O} \quad . \tag{6-14}$$

Nach (3-16) erhält man diese Zerlegung durch die Projektion mit

$$\mathcal{P}_{G0}[\mathcal{A}] = \mathcal{A}(\mathcal{A}^t \mathcal{G} \mathcal{A})^- \mathcal{A}^t \mathcal{G} \quad . \tag{6-15}$$

Um die zunächst noch unbekannte Gewichtsmatrix $\mathcal{G}$ zu bestimmen, wird gefordert, daß die ausgeglichenen Beobachtungen $\hat{\ell} = \mathcal{P}\ell$ erwartungstreu sind und minimale Varianz aufweisen. Die Erwartungstreue ist wegen $\ell \in S(\mathcal{A})$ gewährleistet wenn

$$\mathcal{P}\mathcal{A} = \mathcal{A} \tag{6-16}$$

gilt, das heißt, $\mathcal{P}$ und $\mathcal{A}$ müssen denselben Spaltenraum aufspannen.

Damit die Varianzen der ausgeglichenen Beobachtungen minimal werden, ist

$$\mathrm{sp}\left[\mathcal{P}\, Q_l \mathcal{P}^t\right] = \min \tag{6-17}$$

zu fordern. Um (6-16) und (6-17) gleichzeitig zu erfüllen, ist die Lagrange-Funktion

$$L = \mathrm{sp}\left[\mathcal{P}\, Q_l \mathcal{P}^t\right] + 2\,\mathrm{sp}\,\Lambda\,(\mathcal{P}\mathcal{A} - \mathcal{A})$$

zu minimieren. Die Verallgemeinerung des schon in (4-8) verwendeten Verfahrens führt auf das Gleichungssystem

$$\begin{bmatrix} Q_l & \mathcal{A} \\ \mathcal{A}^t & \mathcal{O} \end{bmatrix} \begin{bmatrix} \mathcal{P}^t \\ \Lambda \end{bmatrix} = \begin{bmatrix} \mathcal{O} \\ \mathcal{A}^t \end{bmatrix} \quad . \tag{6-18}$$

Da die Koeffizientenmatrix nicht den vollen Rang hat, ist die Lösung von (6-18) nicht eindeutig. Mit einer beliebigen g-Inversen

$$\begin{bmatrix} Q_1 & A \\ A^t & 0 \end{bmatrix}^- = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} \tag{6-19}$$

und einer frei wählbaren Matrix Z erhält man

$$\begin{bmatrix} P^t \\ \Lambda \end{bmatrix} = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} \begin{bmatrix} 0 \\ A^t \end{bmatrix} + \left\{ I - \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} \begin{bmatrix} Q_1 & A \\ A^t & 0 \end{bmatrix} \right\} Z \quad ,$$

woraus für P^t folgt

$$P^t = C_2 A^t + \left\{ I - C_1 Q_1 - C_2 A^t \right\} Z_1 - C_1 A Z_2 \quad . \tag{6-20}$$

Führt man die Abkürzungen

$$T = Q_1 + A A^t \qquad N = A^t T^- A \tag{6-21}$$

ein, so findet man nach ***PRINGLE/RAYNER*** 1971, S. 48

$$\begin{aligned} C_1 &= T^- - T^- A N^- A^t T^- & C_3 &= N^- A^t T^- \\ C_2 &= T^- A N^- & C_4 &= N^- N - N^- \end{aligned} \tag{6-22}$$

als Lösungen für (6-19), in die beliebige g-Inversen eingesetzt werden können. Als eine der möglichen Formen für den Projektor kann die Sonderlösung mit $Z_i = 0$ gewählt werden.

$$P = A C_2^t = A (N^-)^t A^t (T^-)^t \tag{6-23}$$

Da T und N nach (6-21) symmetrisch sind und da Produkte der Form AB^-C mit $S(A^t) \subset S(B^t)$ und $S(C) \subset S(B)$ invariant in Bezug auf die Wahl für B^- sind, kann für (6-23) vereinfacht geschrieben werden

$$P = A \left(A^t (Q_1 + A A^t)^- A \right)^- A^t (Q_1 + A A^t)^- \quad . \tag{6-24}$$

Der Vergleich mit (6-15) und (6-11) zeigt, daß für die Gewichtsmatrix

$$G = (Q_1 + A A^t)^- = T^- \tag{6-25}$$

gesetzt werden kann. Mit Satz 5 ist leicht überprüft, daß (6-24) tatsächlich ein T^--orthogonaler Projektor ist, der alle Vektoren aus $S(T)$ in den

$S(\mathcal{A})$ abbildet.

$$\hat{\ell} = \mathcal{P}_{T^-0}\,\ell \quad , \qquad -v = \left(\mathcal{I} - \mathcal{P}_{T^-0}\right)\ell \quad , \tag{6-26}$$

wobei

$$\hat{\ell}^t \mathcal{T}^- v = \ell^t \mathcal{P}^t \mathcal{T}^- (\mathcal{I} - \mathcal{P})\,\ell = 0$$

ist.

Der Projektor (6-24) ist eine Verallgemeinerung des in Satz 5 und 6 behandelten Projektors, da er nur über dem Vektorunterraum $S(\mathcal{T})$ definiert ist und da wegen $r(\mathcal{T}^-) \leq m$ in (3-18) eine Seminorm minimiert wird. Wie die in (6-10) bis (6-12) angegebenen Möglichkeiten zur Darstellung von $\mathcal{T}$ zeigen, ist $\mathcal{P}_{T^-0}[\mathcal{A}]$ nicht eindeutig festgelegt. Außerdem ist wegen (3-11) und (6-5) $\mathcal{P}_{T^-0} + \mathcal{N}\mathcal{B}\mathcal{N}^t$ mit $\mathcal{B}$ beliebig ebenfalls ein Projektor, der alle Kriterien befriedigt.

Aus (6-26) folgt

$$\hat{\ell} = \mathcal{A}\,\hat{x} = \mathcal{A}\left(\mathcal{A}^t \mathcal{T}^- \mathcal{A}\right)^- \mathcal{A}^t \mathcal{T}^- \ell \quad . \tag{6-27}$$

Als Lösungsvektor im allgemeinen Gauß-Markoff-Modell ist offensichtlich jeder Vektor $\hat{x}$ der Lösungsmenge

$$L = \left\{\hat{x}, \hat{x} = \left(\mathcal{A}^t \mathcal{T}^- \mathcal{A}\right)^- \mathcal{A}^t \mathcal{T}^- \ell + \left[\mathcal{I} - \left(\mathcal{A}^t \mathcal{T}^- \mathcal{A}\right)^- \mathcal{A}^t \mathcal{T}^- \mathcal{A}\right] z \right\} \tag{6-28}$$

zulässig, wobei z frei gewählt werden kann.

Diese Mehrdeutigkeit für den Parametervektor legt wieder die Frage nahe, ob es außer den ausgeglichenen Beobachtungen weitere lineare Funktionen der Parameter gibt, die erwartungstreu und mit minimaler Varianz schätzbar sind.

Gemäß Definition 2 sei für $y = b^t x$ eine lineare Schätzfunktion $c^t \ell$ gesucht, die erwartungstreu ist, woraus für b wie früher folgt, daß $b \in S(\mathcal{A}^t)$ sein muß, und die ferner minimale Varianz aufweist, also $V(y) = \sigma_0^2\, c^t Q_l c$ minimiert. Den Ableitungen in Abschnitt 4 folgend führen diese beiden Forderungen auf das (4-9) entsprechende Gleichungssystem

$$\begin{bmatrix} Q_l & \mathcal{A} \\ \mathcal{A}^t & 0 \end{bmatrix} \begin{bmatrix} c \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \quad , \tag{6-29}$$

das nun aber wie in (6-18) durch einen Rangdefekt in Q_1 und A gekennzeichnet ist. Die in (6-19) und (6-22) angegebene g-Inverse kann auch hier zur Lösung benutzt werden:

$$c = C_2 b + \left[I - C_1 Q_1 + C_2 A^t\right] z_1 - C_1 A z_2 \quad . \tag{6-30}$$

Dies führt auf die Schätzfunktion

$$\hat{y} = b^t C_2^t l + z_1^t \left[I - Q_1 C_1^t + A C_2^t\right] l - z_2^t A^t C_1^t l \quad ,$$

die sich vereinfacht, wenn man berücksichtigt, daß wegen $l \in S(T)$ die Darstellung $l = Tg$ folgt, die mit C_1 und C_2 nach (6-22) auf

$$\left[I - Q_1 C_1^t + A C_2^t\right] T = O \quad ; \quad A^t C_1^t T = O \tag{6-31}$$

führt, so daß

$$\hat{y} = b^t C_2^t l \tag{6-32}$$

übrig bleibt. Nach (6-22) ist $C_2 = T^- A \left(A^t T^- A\right)^-$. Wird dies in (6-32) eingesetzt, so zeigt der Vergleich mit (6-28), daß $\hat{y} = b^t \hat{x}$ für jedes Element von L gilt. Darüber hinaus kann wegen $b^t = c^t A$ mit (6-27) auch $\hat{y} = c^t l$ geschrieben werden.

Die Varianz von $\hat{y}$ lautet mit $N = A^t T^- A$

$$\begin{aligned} V(\hat{y}) &= \sigma_0^2 \, b^t C_2^t Q_1 C_2 b \\ &= \sigma_0^2 \, b^t N^- A^t T^- Q_1 \, T^- A \, N^- b \quad . \end{aligned} \tag{6-33}$$

Setzt man in (6-33) $Q_1 = T - A A^t$ ein und berücksichtigt, daß $T T^- Q_1 = Q_1$ und $T T^- A = A$ ist, so erhält man die einfachere Form

$$V(\hat{y}) = \sigma_0^2 \, b^t \left(N^- - N^- N\right) b = -\sigma_0^2 \, b^t C_4 \, b \tag{6-34}$$

mit C_4 nach (6-22). Nun ist aber $b \in S\left(A^t\right)$, so daß $b^t N^- N = b^t$, eingesetzt in (6-34), die endgültige Varianzschätzung liefert:

$$V(\hat{y}) = \sigma_0^2 \, b^t \left[\left(A^t T^- A\right)^- - J\right] b \quad . \tag{6-35}$$

Der quadratische Schätzer für σ_0^2 gemäß Definition 3 lautet im allgemeinen Modell

$$s_0^2 = \frac{\boldsymbol{v}^t \mathcal{T}^- \boldsymbol{v}}{f} = \frac{\boldsymbol{\ell}^t \mathcal{T}^- (\mathcal{I} - \mathcal{P})\, \boldsymbol{\ell}}{f} = \boldsymbol{\ell}^t \mathcal{B}\, \boldsymbol{\ell} \qquad (6\text{-}36)$$

mit $f = r(\mathcal{T}) - r(\mathcal{A})$.

Die Erwartungstreue dieses Schätzers kann auf dem üblichen Weg gezeigt werden:

$$E\left(\boldsymbol{v}^t \mathcal{T}^- \boldsymbol{v}\right) = E\left(sp\left[\mathcal{T}^- \boldsymbol{v}\boldsymbol{v}^t\right]\right) = sp\left[\mathcal{T}^- \Sigma_v\right] .$$

Setzt man

$$\Sigma_v = \sigma_0^2\, (\mathcal{P} - \mathcal{I})\, \mathcal{Q}_l \left(\mathcal{P}^t - \mathcal{I}\right)$$

ein und multipliziert aus, wobei die Beziehungen $\mathcal{T} = \mathcal{Q}_l + \mathcal{A}\,\mathcal{A}^t$ und $sp[\mathcal{A}\mathcal{B}] = sp[\mathcal{B}\mathcal{A}]$ benutzt werden, so erhält man

$$E\left(\boldsymbol{v}^t \mathcal{T}^- \boldsymbol{v}\right) = \sigma_0^2\, sp[\mathcal{T}\mathcal{T}^- - \mathcal{N}\mathcal{N}^-]$$

$$= \sigma_0^2 \left[r(\mathcal{T}) - r(\mathcal{A})\right] = \sigma_0^2\, f .$$

Ferner liest man wegen $\mathcal{A}^t \mathcal{B} = \mathcal{O}$ mit $\mathcal{B} = \mathcal{T}^-\left(\mathcal{I} - \mathcal{P}\right)/f$ die Invarianz des Schätzers ab. Auch das Kriterium der minimalen Varianz entsprechend (4-19) erweist sich durch Einsetzen und Ausmultiplizieren als erfüllt. Wie man leicht zeigen kann, darf (6-36) noch etwas vereinfacht werden, da für beliebige g-Inverse von $\mathcal{Q}_l$ die Beziehung

$$\boldsymbol{v}^t \mathcal{T}^- \boldsymbol{v} = \boldsymbol{v}^t \mathcal{Q}_l^- \boldsymbol{v} \qquad (6\text{-}37)$$

gilt.

Nach (6-13) ist $-\boldsymbol{v} \in S(\mathcal{M})$, und wegen $\mathcal{M} = \mathcal{Q}_l(\mathcal{Q}_l\mathcal{A})^\perp$ nach (6-9) kann $\boldsymbol{v} = \mathcal{M}\boldsymbol{a} = \mathcal{Q}_l \bar{\boldsymbol{a}}$ gesetzt werden. Führt man dies in (6-37) ein, so erhält man

$$\boldsymbol{v}^t \mathcal{T}^- \boldsymbol{v} = \boldsymbol{v}^t \mathcal{T}^- \mathcal{Q}_l \bar{\boldsymbol{a}}$$

$$= \boldsymbol{v}^t \mathcal{T}^- \left[\mathcal{T} - \mathcal{A}\mathcal{A}^t\right] \bar{\boldsymbol{a}}$$

$$\boldsymbol{v}^t \mathcal{T}^- \boldsymbol{v} = \boldsymbol{v}^t \mathcal{T}^- \mathcal{T} \bar{\boldsymbol{a}} - \boldsymbol{v}^t \mathcal{T}^- \mathcal{A}\mathcal{A}^t \bar{\boldsymbol{a}} . \qquad (6\text{-}38)$$

Der zweite Ausdruck auf der rechten Seite ist null, was sich sogleich durch Einsetzen von $\boldsymbol{v}^t = \boldsymbol{\ell}^t\left(\mathcal{P}^t - \mathcal{I}\right)$ ergibt. Es bleibt damit

$$v^t T^- v = \bar{a}^t Q_1 T^- T \bar{a} = \bar{a}^t Q_1 \bar{a}$$

und nach Rückwärtseinsetzen mit $Q_1 Q_1^- Q_1 = Q_1$ folgt die behauptete Beziehung (6-37).

Im allgemeinen Gauß-Markoff-Modell spielt die Matrix T^-, wobei T mit (6-10), (6-11) oder (6-12) gegeben ist, die Rolle der Gewichtsmatrix. Die Beschränkung der vorstehenden Ableitungen auf den Sonderfall $T = Q_1 + A A^t$ diente lediglich der Vereinfachung der Formeln.

Wenn im Modell $S(A) \subset S(Q_1)$ gilt, so kann in allen Formeln die Matrix T^- durch eine beliebige g-Inverse Q_1^- ersetzt werden.

Falls A vollen Rang hat, also ein Q-singuläres Modell vorliegt, so ändert sich nichts an den Formeln, die für Schätzungen im allgemeinen Modell abgeleitet wurden.

7. SCHLUSSBEMERKUNGEN

Die mit Hilfe von Projektionsmatrizen mögliche direkte Lösung der Schätzaufgaben im allgemeinen Gauß-Markoff-Modell wurde unter den Kriterien Erwartungstreue und minimale Varianz entwickelt. Wegen der Äquivalenz dieser Kriterien mit der Methode der kleinsten Quadrate handelt es sich bei den Ergebnissen um Verallgemeinerungen der klassischen Formeln der Ausgleichungsrechnung.

Wie in ***CASPARY*** 1983 gezeigt wurde, ist es jedoch auch möglich, einen anderen Weg zu wählen, der auf ein reguläres Modell mit Bedingungsgleichungen führt, in dem die bekannten Schätzformeln ihre Gültigkeit behalten. Dazu ist lediglich eine Transformation des Modells nötig.

Der hier eingeschlagene Weg hat allerdings den Vorteil, daß die in (6-27), (6-32) und (6-36) gefundenen Schätzfunktionen in allen Modellen nach Definition 1 gültig sind.

LITERATUR

CASPARY, W.: *Zur Lösung singulärer Ausgleichungsmodelle durch Bedingungsgleichungen.* Allgemeine Vermessungsnachrichten 2 (1978), S. 81-87

CASPARY, W.: *Zur Singularität von Varianz-Kovarianzmatrizen.* Zeitschrift für Vermessungswesen, 1983, S. 209-215

KOCH, K.-R.: *Parameterschätzung und Hypothesentest in linearen Modellen.* Dümmlers Verlag, Bonn 1980

MITRA, S.K.: *Unified Least Squares Approach to Linear Estimation in a General Gauss-Markov Model.* SIAM J. Appl. Math. Vol. 25, 1973, S. 671-680

MITRA, S.K., RAO, C.R.: *Some results in estimation and tests of linear hypotheses under the Gauss-Markoff Model.* Sankhya, Ser. A 30, 1968, S. 281-290

MITRA, S.K., RAO, C.R.: *Projections under Seminorms and Generalized Moore-Penrose Inverses.* Linear Algebra and its Applications 9, 1974, S. 155-167

PRINGLE, R.M., RAYNER, A.A.: *Generalized Inverse Matrices with Applications to Statistics.* Griffin's Statistical Monographs & Courses, London 1971

RAO, C.R.: *Representations of Best Linear Unbiased Estimators in the Gauss-Markoff Model with a Singular Dispersion Matrix.* Journal of Multivariate Analysis 3, 1973, S. 276-292

RAO, C.R.: *Projectors, Generalized Inverses and the BLUE's.* Journal of the Royal Statistic Society, Serie B, Vol. 36, 1974, S. 442-446

ZYSKIND, G.: *On Canonical Forms, Non-negative Covariance Matrices and Best and Simple Least Squares Linear Estimations in Linear Models.* Annals of Mathematical Statistics 38, 1967, S. 1092-1109

ZYSKIND, G.: *Error Structures, Projections and Conditional Inverses in Linear Model Theory.* In J.N. Srivastava (Hrsg.): *A Survey of Statistical Design and Linear Models,* 1975, S. 647-663

ZYSKIND, G., MARTIN, F.: *On Best Linear Estimation and a General Gauss-Markov Theorem in Linear Models with Arbitrary Non-negative Covariance Structure.* SIAM Journal on Applied Mathem. 17, 1969, S. 1190-1202

UMKEHRUNG VON VOLLSTÄNDIGEN POTENZREIHEN MIT ZWEI VERÄNDERLICHEN

von Hans GLASMACHER und Klaus KRACK

SUMMARY

Two different solution methods are presented for the automation of the reversal problem of a complete power series with two variables.

The required programming input is demonstrated by program listings and the practical explanation is proved by a numerical example.

ZUSAMMENFASSUNG

Zur Automatisierung des Problems der Umkehrung einer vollständigen Potenzreihe mit zwei Variablen werden verschiedene Lösungswege aufgezeigt.

Der notwendige Programmieraufwand wird anhand von Programmausdrucken dargestellt und die praktische Ausführung mit einem Zahlenbeispiel belegt.

1. EINFÜHRUNG

Die ausgeschriebenen Formeln der Umkehrung einer vollständigen Potenzreihe mit zwei Variablen werden bei höherem Grade der Reihe sehr umfangreich. Es ist dann nicht mehr sinnvoll, diese von Hand aufzuschreiben, um später damit zu rechnen, da die Wahrscheinlichkeit eines Schreib- oder Rechenfehlers zu groß wird. Lösungen müssen daher entweder automatisch erzeugt oder direkt als numerischer Algorithmus programmiert werden. Nachfolgend werden diese beiden Vorgehensweisen dargestellt und bezüglich ihrer Vor- und Nachteile miteinander verglichen.

2. NUMERISCHE REIHENUMKEHR

Die numerische Reihenumkehr basiert auf einer allgemeinen Lösung aus ***KÖNIG-WEISE*** (1951). Für die Darstellung dieser Lösung wird eine Bezeichnungsweise verwendet, die zwischen der Schreibweise in ***KÖNIG-WEISE*** und dem beigefügten Programmausdruck vermitteln soll.

Gegeben seien die Reihen

$$u = \sum_{\substack{i=0,n \\ k=0,n-i}} A(i,k)\cdot x^i\cdot y^k \quad ; \quad v = \sum_{\substack{i=0,n \\ k=0,n-i}} B(i,k)\cdot x^i\cdot y^k \qquad (2\text{-}1)$$

mit

$$A(0,0) = 0 \quad ; \quad B(0,0) = 0 \quad .$$

Gesucht sind die Reihen

$$x = \sum_{\substack{i=0,n \\ k=0,n-i}} C(i,k)\cdot u^i\cdot v^k \quad ; \quad y = \sum_{\substack{i=0,n \\ k=0,n-i}} D(i,k)\cdot u^i\cdot v^k \qquad (2\text{-}2)$$

mit

$$C(0,0) = 0 \quad ; \quad D(0,0) = 0 \quad .$$

Der klassische Ansatz zur Lösung der Reihenumkehr besteht aus dem Rückeinsetzen der Reihen für u und v in die Reihen für x und y und anschließendem Koeffizientenvergleich. Sortiert man diese Reihen nach Potenzen von u und v, so müssen deren Koeffizienten bei der ersten Ordnung gleich 1 oder 0 sein und bei allen höheren Ordnungen gleich 0 sein. Es entsteht ein System von Gleichungen, das sukzessive aufgelöst werden kann.

Die Umkehrkoeffizienten der 1. Ordnung können ohne Schwierigkeit direkt angegeben werden.

$$C(1,0) = \frac{B(0,1)}{D1} \quad ; \quad C(0,1) = -\frac{A(0,1)}{D1}$$
$$D(1,0) = -\frac{B(1,0)}{D1} \quad ; \quad D(0,1) = \frac{A(1,0)}{D1} \qquad (2\text{-}3)$$

mit

$$D1 = A(1,0) \cdot B(0,1) - A(0,1) \cdot B(1,0) \qquad (2\text{-}4)$$

Die Umkehrreihen-Koeffizienten C(i,k), D(i,k) für $i+k \geq 2$ bestehen jeweils aus einer Summe über ein Produkt, dessen erster Faktor FK für beide Reihen gleich ist, während der zweite Faktor FC bzw. FD für C und D unterschiedlich ist. Durch diese Aufspaltung können die Koeffizientenpaare gemeinsam ohne Doppelarbeit berechnet werden. Zur Summenbildung werden zwei Hilfsvariablen *) L und M eingeführt, die Summe läuft über alle Kombinationen, die die Bedingung $1 < L+M \leq I+K$ erfüllen.

$$C(I,K) = \sum_{\substack{L=0,I+K \\ M=0,I+K-L}} IF(L+M>1) \; : \; FK(I,K,L,M) \cdot FC(L,M) \qquad (2\text{-}5)$$

$$D(I,K) = \sum_{\substack{L=0,I+K \\ M=0,I+K-L}} IF(L+M>1) \; : \; FK(I,K,L,M) \cdot FD(L,M) \qquad (2\text{-}6)$$

*) Um einen direkten Bezug zu den Programmanweisungen in Kapitel 3 zu erhalten, werden die Indizes und Laufvariablen im weiteren mit den dort verwendeten Großbuchstaben bezeichnet.

Die zweiten Faktoren FC(L,M) und FD(L,M) sind Lösungen eines Gleichungssystems mit zwei Unbekannten. Sie lauten mit dem D1 aus (2-4)

$$FC(L,M) = \frac{C(L,M) \cdot D(0,1) - D(L,M) \cdot C(0,1)}{D1} \qquad (2\text{-}7)$$

$$FD(L,M) = \frac{D(L,M) \cdot C(0,1) - C(L,M) \cdot D(0,1)}{D1} \qquad (2\text{-}8)$$

Der erste Faktor FK(I,K,L,M) besteht aus einer Summe von Produkten über alle möglichen Kombinationen, die mit den Hilfsvariablen LL, MM und SR, ST, SS und SU die Bedingungen

$$\sum_{LL=1,L} R(LL) + \sum_{MM=1,M} T(MM) = I = SR + ST \qquad (2\text{-}9)$$

und

$$\sum_{LL=1,L} S(LL) + \sum_{MM=1,M} U(MM) = K = SS + SU \qquad (2\text{-}10)$$

erfüllen.

R(LL), LL = 1 bis L ist, wie entsprechend T, S und M, ein Vektor mit L ganzzahligen Elementen zwischen 0 und SR, dessen Element-Summe SR ist, z.Bsp.

$$SR = 3 \quad , \quad L = 2 \; ; \; R = /\, 3 \, , \, 0 \, / \quad , \quad R(LL=1) = 3 \quad , \quad R(LL=2) = 0 \qquad (2\text{-}11)$$

Es müssen jeweils alle Vektoren eingesetzt werden, die diese Bedingungen erfüllen, z.Bsp. auch

$$R = /\, 2 \, , \, 1 \, / \quad , \quad R = /\, 1 \, , \, 2 \, / \quad \text{und} \quad R = /\, 0 \, , \, 3 \, / \, . \qquad (2\text{-}12)$$

Es gibt $KR=\binom{SR+L-1}{SR}$ Vektoren, die die Bedingungen Elementzahl = L und Elementsumme = SR erfüllen. Das gleiche gilt entsprechend für die Vektoren S, T und U.

Die Binominalkoeffizienten $\binom{NB}{KB}$ werden berechnet mit

$$\binom{NB}{KB} = \prod_{J=1,KB} \frac{NB-J+1}{J} \quad . \qquad (2\text{-}13)$$

Um die möglichen Vektoren nicht doppelt berechnen zu müssen, werden sie zu Beginn des Programms, geordnet nach aufsteigender Element-Summe und -Anzahl, erzeugt und gespeichert.

Dies geschieht einfach mit Programmschleifen für die Element-Summe E und die Elemente Ei und Vergleich von E mit der Summe der Ei. Da es beim Grade N der Reihe insgesamt $\binom{2\cdot N-1}{N}$ verschiedene Vektoren gibt, werden sie, um Kernspeicherplatz zu sparen, nach Zehnerpotenzen geordnet zu einer mehrstelligen Integerzahl zusammengefaßt und hintereinander abgespeichert. Dabei wird die letzte laufende Nummer jeder Element-Summe für den späteren gezielten Zugriff in SA festgehalten (siehe Programmausdruck in Kapitel 3), Subroutine "KOMBI").

Der jeweilige Vektor R läßt sich dann als mehrstellige Zahl in Abhängigkeit von der Elementsumme SR und einer laufenden Kombinations-Nummer KR abrufen.

$$R = KO(SA(SR+1)+KR) \quad ; \quad R(LL) = \frac{R}{10^{LL-1}} - \text{INTEGER}\left(\frac{R}{10^{LL}}\right)\cdot 10 \qquad (2\text{-}14)$$

Entsprechend werden S(LL), T(MM) und U(MM) bestimmt.

Zur Berechnung von FK muß also sowohl über die Bedingungen (2-9) und (2-10) summiert werden, als auch über die jeweilige Anzahl der Kombinationen KR, KS, KT und KU der Vektoren.

Zur abschließenden Produktbildung werden die Elemente der Vektoren paarweise als Indices der bereits berechneten Koeffizienten C(i,k) und D(i,k) benutzt.

Z.Bsp. ist bei

SR=3 , SS=0 , L=2 : R = / 3 , 0 / , S = / 0 , 0 /

$$\prod_{LL=1,L} C\big(R(LL),S(LL)\big) = C(3,0)\cdot C(0,0)$$

und die weiteren Produkte für die anderen zulässigen R aus (2-12) sind

$$C(2,0)\cdot C(1,0) \quad , \quad C(1,0)\cdot C(2,0) \quad \text{und} \quad C(0,0)\cdot C(3,0)$$

Der Faktor FK(I,K,L,M) läßt sich daher folgendermaßen darstellen:

$$FK(I,K,L,M) = \sum_{\substack{SR=0,I \\ ST=I-SR}} \; \sum_{\substack{KR=1,\binom{SR+L-1}{SR} \\ KT=1,\binom{ST+M-1}{ST} \\ SU=0,K \\ KU=1,\binom{SU+M-1}{SU} \\ SS=K-SU \\ KS=1,\binom{SS+L-1}{SS}}} P(L,M,SR,ST,SU,SS,KR,KT,KU,KS) \qquad (2\text{-}15)$$

mit

$$P(....) = \prod_{\substack{IF(L>0):LL=1,L \\ IF(M>0):MM=1,M}} C\big(R(LL),S(LL)\big)\cdot D\big(T(MM),U(MM)\big) \qquad (2\text{-}16)$$

Bei der Programmierung sind einige logische Sonderfälle zu berücksichtigen, die den Fall L=0 und M=0 betreffen. So muß ein Vektor mit Element-Zahl L=0 zwangsläufig auch die Element-Summe SR=SS=0 haben und damit wird ST=I und SU=K. Bei M=0 wird ST=SU=0, SR=I und SS=K. Der Vektor mit Elementzahl Null wird bei der Produktbildung nicht berücksichtigt. Das Produkt muß also nicht Null werden (siehe Gl. (2-16)). Gleichung (2-15) muß daher durch logische Abfragen wie folgt erweitert werden:

$$FK(....) = \sum_{\substack{SR=0,I \\ IF(M=0):SR=I \\ ST=I-SR \\ SU=0,K \\ IF(L=0):SU=K \\ SS=K-SU}} IF\big((L\#0 \text{ oder } SR=0) \text{ und } (M\#0 \text{ oder } SU=0)\big) : \Sigma\, P(..) \qquad (2\text{-}15a)$$

Zur Umkehrung einer Reihe wird also zunächst die 1. Ordnung nach Gleichungen (2-3) und (2-4) berechnet und für die weitere schrittweise Näherungsrechnung gespeichert. Danach werden alle möglichen Vektoren mit Subroutine "KOMBI" und in aufsteigender Reihenfolge die weiteren Koeffizientenpaare C(i,k) und D(i,k) nach Gleichungen (2-5) und (2-6) mit (2-7), (2-8), (2-15a), (2-15), (2-13), (2-16) und (2-14) berechnet.

3. ANGABEN ZUM PROGRAMM UND RECHENBEISPIEL

Das Rechenprogramm ist in FORTRAN geschrieben und hat etwa 85 wirksame Programmzeilen. Die Rechenzeiten auf dem Großrechner BURROUGHS B 7800/162 sind nachstehender Tabelle zu entnehmen. Es sind trotz der zahlreichen Summationen keine wesentlichen Rundungseffekte festzustellen. Nach zweimaliger Reiheninversion mit einfacher Genauigkeit (12 Stellen) lag die größte Abweichung bei 10^{-10}. Der Fehler durch Rundungen ist damit wesentlich kleiner als der durch den Abbruch der unendlichen Reihe verursachte Fehler.

Programmausdruck:

```
C*************************************************
C
C  PROGRAMM ZUR UMKEHRUNG ALLGEMEINER POTENZREIHEN MIT 2 VERAENDERLICHEN
C
C  METHODE NACH  KOENIG / WEISE  BD.1  S.501 FF.
C
      SUBROUTINE NINA2(N,A,B,C,D)
      INTEGER KO(65535),ISA(9)
      REAL*4 A(10,10),B(10,10),C(10,10),D(10,10),FK,FC,FD,D1
      CALL KOMBI(N,KO,ISA)
      A(1,1)=0 ; B(1,1)=0
      D1 = A(2,1)*B(1,2)-A(1,2)*B(2,1)
      C(2,1)= B(1,2)/D1 ; C(1,2)=-A(1,2)/D1
      D(2,1)=-B(2,1)/D1 ; D(1,2)= A(2,1)/D1
      DO 1 IN=1,N
      DO 1 I=0,IN ; K=IN-I
C
C  SUMMATION UEBER L+M = 1,I+K
      DO 1 L=0,I+K
      DO 1 M=0,I+K-L
      IF(L+M.LE.1) GOTO 1
      FC = A(L+1,M+1)*B(1,2)-A(1,2)*B(L+1,M+1)
      FD = A(2,1)*B(L+1,M+1)-A(L+1,M+1)*B(2,1)
      CALL GK(FK,I,K,L,M,C,D,KO,ISA)
      C(I+1,K+1)=C(I+1,K+1)-FK*FC/D1
      D(I+1,K+1)=D(I+1,K+1)-FK*FD/D1^
    1 CONTINUE
C
      END
```

```
C***********************************************
C
C  BERECHNUNG DER KOEFFIZIENTEN K (I,K,L,M)
C
      SUBROUTINE GK(FK,I,K,L,M,C,D,KO,ISA)
      INTEGER KO(65535),ISA(9)
      REAL*4 A(10,10,B(10,10),C(10,10),D(10,10),FK,SUM,PROD
C
C  SUMMATION UEBER  SR+ST=I , SS+SU=K  UND ALLE VEKTOR-KOMBINATIONEN
      DO 1 SR=0,I
      IF(M.EQ.0)SR=I
      IF(L.EQ.0.AND.SR.NE.0) GOTO 1
      ST=I-SR
      DO 1 SU=0,K
      IF(L.EQ.0)SU=K
      IF(M.EQ.0.AND.SU.NE.0) GOTO 1
      SS=K-SU
      DO 2 KR=1,BINKO(SR+L-1,SR)
      DO 2 KS=1,BINKO(SS+L-1,SS)
      DO 2 KT=1,BINKO(ST+M-1,ST)
      DO 2 KU=1,BINKO(SU+M-1,SU)
C
C  PRODUKTBILDUNG  LL , MM  = 0 BIS L , M
      PROD=1
      IF(L.EQ.0) GOTO 3
      DO 4 LL=1,L
      RL=KO(ISA(SR+1)+KR)/10**(LL-1) - KO(ISA(SR+1)+KR)/10**LL*10
      SL=KO(ISA(SS+1)+KS)/10**(LL-1) - KO(ISA(SS+1)+KS)/10**LL*10
      IF(RL+SL.EQ.0) GOTO 2
    4 PROD=PROD*C(RL+1,SL+1)
    3 IF(M.EQ.0) GOTO 5
      DO 6 MM=1,M
      TM=KO(ISA(ST+1)+KT)/10**(MM-1) - KO(ISA(ST+1)+KT)/10**MM*10
      UM=KO(ISA(SU+1)+KU)/10**(MM-1) - KO(ISA(SU+1)+KU)/10**MM*10
      IF(UM+TM.EQ.0) GOTO 2
    6 PROD=PROD*D(TM+1,UM+1)
C
    5 IF(L+M.GT.0) SUM=SUM+PROD
    2 CONTINUE
    1 FK=SUM
      END
C***************************************************
C
C  BERECHNUNG VON BINOMIALKOEFFIZIENTEN  N ueber K
C
      REAL FUNCTION BINKO(N,K)
      BINKO=1
      IF(K.LE.0) RETURN
      DO 10 J=1,K
 10   BINKO=BINKO*(N-J+1.)/J
      END
C***********************************************
C
C  BERECHNUNG ALLER KOMBINATIONEN VON 1 BIS IS ZIFFERN
C  ( IN DIESER REIHENFOLGE ) , DEREN SUMME IS IST
C
      SUBROUTINE KOMBI(N,KO,ISA)
      INTEGER KO(65535), ISA(9)
      DO 1 IS=0,N
      ISA(IS+1)=NR
      DO 1 K8=0,IS
      DO 1 K7=0,IS-K8
      DO 1 K6=0,IS-K8-K7
      DO 1 K5=0,IS=K8-K7-K6
      KZ=K8+K7+K6+K5
      DO 1 K4=0,IS-KZ
      DO 1 K3=0,IS-KZ-K4
      DO 1 K2=0,IS-KZ-K4-K3
      DO 1 K1=0,IS-KZ-K4-K3-K2
      KS=K1+K2+K3+K4+KZ
      IF(KS.NE.IS) GOTO 1
      NR=NR+1
      KO(NR)=K1+10*K2+100*K3+1E3*K4+1E4*K5+1E5*K6+1E6+K7+1E7*K8
   1  CONTINUE
      END
```

4. REIHENUMKEHRUNG MIT REDUCE

Neben dem auf Rekursionsformeln von ***KÖNIG-WEISE*** (1951) aufbauenden Algorithmus zur Reihenumkehrung gibt es die Möglichkeit, die Gleichungen zur Lösung des Umkehrproblems von Grund auf rechnergestützt entwickeln zu lassen.

Diese Vorgehensweise ermöglicht das Programmsystem Reduce (***HEARN***, 1968), welches speziell für die rechnergestützte algebraische Formelmanipulation entwickelt wurde. Die Fähigkeiten von Reduce lassen sich in etwa wie folgt zusammenfassen:

Reduce ist als dialogfähiger Interpreter am Rechenzentrum der HSBw auf dem Rechner B 7800 installiert. Als Daten können z.Bsp. geklammerte Ausdrücke mit freien Variablen und Zahlen eingeführt werden, über die sich mathematische Umformungen durchführen lassen. Diese Umformungen sind mit den Begriffen Verknüpfung, Vereinfachung, Substitution und Anwendung von Operatoren beschreibbar.

Mit Reduce ist es möglich, die Programmierung des Umkehrproblems auf die untere übersichtliche Stufe der algebraischen Lösungsherleitung zu stellen. Damit brauchen keine Rekursionsformeln mehr gesucht werden, welche die Grundlage des unter 1. beschriebenen, rein numerisch arbeitenden Algorithmus darstellen.

5. LÖSUNGSWEG

Im Hinblick auf die rechnergestützte Reihenumkehrung mit Reduce ist es zunächst notwendig, eine Lösungsstrategie bereitzustellen, die sich möglichst bequem mit den in Reduce gegebenen Sprachelementen in ein Programm umsetzen läßt.

Einem Algorithmus, der dabei mit möglichst wenig Such-, Sortier- und Substitutionsvorgängen auskommt, wird dabei der Vorzug zu geben sein. Gleichzeitig sollte der Bedarf an Rechenzeit und Speicherkapazität nicht außer Acht gelassen werden.

Unter diesen Gesichtspunkten hat es sich als vorteilhaft erwiesen, die bis

zur Ordnungszahl n gesuchte Lösung nicht in einem Guß zu erzeugen, sondern vielmehr die einzelnen Ordnungen, was durchaus möglich ist, getrennt zu behandeln.

Praktisch werden für die Ordnungszahlen $n \leq 4$ nacheinander geschlossene Lösungen erzeugt, während für die höheren Ordnungen, um Speicherplatz zu sparen, ein System linearer Gleichungen angegeben wird, das zur Lösungsfindung nach einem bekannten Verfahren rein numerisch nach den gesuchten Unbekannten aufzulösen ist.

Zur numerischen Auswertung der gefundenen Gleichungen wird durch Reduce ein fortrancompatibles File erzeugt.

6. ANGABE DER GRUNDFORMELN

Die Umkehrung der Reihen erfolgt mit Hilfe der Methode des Koeffizientenvergleichs.

Ausgehend von den Reihen (2-1)

$$\begin{aligned} u = \; & a_{10}\, x + a_{01}\, y \\ & + a_{20}\, x^2 + a_{11}\, xy + a_{02}\, y^2 \\ & + a_{30}\, x^3 + a_{21}\, x^2 y + a_{12}\, xy^2 + a_{03}\, y^3 \\ & \;\; \cdot \\ & \;\; \cdot \\ & \;\; \cdot \end{aligned} \qquad (2\text{-}1a)$$

$$\begin{aligned} v = \; & b_{10}\, x + b_{01}\, y \\ & + b_{20}\, x^2 + b_{11}\, xy + b_{02}\, y^2 \\ & + b_{30}\, x^3 + b_{21}\, x^2 y + b_{12}\, xy^2 + b_{03}\, y^3 \\ & \;\; \cdot \\ & \;\; \cdot \\ & \;\; \cdot \end{aligned} \qquad (2\text{-}1b)$$

werden die umgekehrten Reihen (2-2) mit den unbekannten Koeffizienten c_{ij} und d_{ij} gesucht.

$$
\begin{aligned}
x = \;& c_{10}\,u + c_{01}\,v \\
&+ c_{20}\,u^2 + c_{11}\,uv + c_{02}\,v^2 \\
&+ c_{30}\,u^3 + c_{21}\,u^2v + c_{12}\,uv^2 + c_{03}\,v^3 \\
&\;\;\vdots
\end{aligned}
\qquad (2\text{-}2a)
$$

$$
\begin{aligned}
y = \;& d_{10}\,u + d_{01}\,v \\
&+ d_{20}\,u^2 + b_{11}\,uv + d_{02}\,v^2 \\
&+ d_{30}\,u^3 + b_{21}\,u^2v + d_{12}\,uv^2 + d_{03}\,v^3 \\
&\;\;\vdots
\end{aligned}
\qquad (2\text{-}2b)
$$

Am Beispiel der Ordnungen $n \leq 2$ soll das allgemein anwendbare Lösungsprinzip im folgenden demonstriert werden.

Das Rückeinsetzen von (2-1) in (2-2a) ergibt, sortiert nach Potenzen von x und y:

$$
\begin{aligned}
x = \;& x\;(a_{10}\,c_{10}+b_{10}\,c_{10}) \\
&+ y\;(a_{01}\,c_{10}+b_{01}\,c_{10}) \\
&+ x^2\,(a_{20}\,c_{10}+b_{20}\,c_{01}+a_{10}{}^2 \quad c_{20}+\,a_{10}\,b_{10} \quad c_{11}+b_{10}{}^2 \quad c_{02}) \\
&+ xy\,(a_{11}\,c_{10}+b_{11}\,c_{01}+2a_{10}\,a_{01}\;c_{20}+(a_{10}\,b_{01}+a_{01}\,b_{10})\;c_{11}+2b_{10}\,b_{01}\;c_{02}) \\
&+ y^2\,(a_{02}\,c_{10}+b_{02}\,c_{01}+a_{01}{}^2 \quad c_{20}+\,a_{01}\,b_{01} \quad c_{11}+b_{01}{}^2 \quad c_{02})
\end{aligned}
\qquad (6\text{-}1)
$$

Das Rückeinsetzen von (2-1) in (2-2b) ergibt analog:

$$
\begin{aligned}
y = \;& x\;(a_{10}\,d_{10}+b_{10}\,d_{10}) \\
&+ y\;(a_{01}\,d_{10}+b_{01}\,d_{10}) \\
&+ x^2\,(a_{20}\,d_{10}+b_{20}\,d_{01}+a_{10}{}^2 \quad d_{20}+\,a_{10}\,b_{10} \quad d_{11}+b_{10}{}^2 \quad d_{02}) \\
&+ xy\,(a_{11}\,d_{10}+b_{11}\,d_{01}+2a_{10}\,a_{01}\;d_{20}+(a_{10}\,b_{01}+a_{01}\,b_{10})\;d_{11}+2b_{10}\,b_{01}\;d_{02}) \\
&+ y^2\,(a_{02}\,d_{10}+b_{02}\,d_{01}+a_{01}{}^2 \quad d_{20}+\,a_{01}\,b_{01} \quad d_{11}+b_{01}{}^2 \quad d_{02})
\end{aligned}
\qquad (6\text{-}2)
$$

Der Koeffizientenvergleich in (6-1) und (6-2) führt zunächst auf zwei lineare Gleichungssysteme für die 10 Unbekannten c_{ij} und d_{ij} für $n \leq 2$.

Sie lauten in Matrizenschreibweise, ausgehend von (6-1)

$$\begin{bmatrix} a_{10} & b_{10} & 0 & 0 & 0 \\ a_{01} & b_{01} & 0 & 0 & 0 \\ a_{20} & b_{20} & a_{10}^2 & a_{10}b_{10} & b_{10}^2 \\ a_{11} & b_{11} & 2a_{10}a_{01} & a_{10}b_{01}+a_{01}b_{10} & 2b_{10}b_{01} \\ a_{02} & b_{02} & a_{01}^2 & a_{01}b_{01} & b_{01}^2 \end{bmatrix} \cdot \begin{bmatrix} c_{10} \\ c_{01} \\ c_{20} \\ c_{11} \\ c_{02} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad (6\text{-}3)$$

oder abgekürzt : $KO \cdot C = LX$ (6-3a)

Ausgehend von (6-2) erhält man

$$KO \cdot D = LY \qquad (6\text{-}4)$$

mit

$$D = \begin{bmatrix} d_{10} \\ d_{01} \\ d_{20} \\ d_{11} \\ d_{02} \end{bmatrix} \qquad LY = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Aus (6-3a) und (6-4) ließen sich die Unbekannten geschlossen herleiten:

$$C = KO^{-1} \cdot LX$$

$$D = KO^{-1} \cdot LY \ .$$

Der Rechenaufwand verringert sich jedoch wesentlich, wenn man in den Gleichungssystemen (6-3a) und (6-4) jeweils nur die zur gleichen Ordnungszahl n gehörenden Unbekannten gemeinsam betrachtet. Die Matrizen KO bzw. KO und L sind dabei den Systemen (6-5) und (6-6) bzw. (6-5a) und (6-6a) gemein.

$n = 1:$ $$C_{n=1} = KO_{n=1}^{-1} \cdot LX_{n=1} \tag{6-5}$$

$$D_{n=1} = KO_{n=1}^{-1} \cdot LY_{n=1} \tag{6-6}$$

$n = 2:$ $$C_{n=2} = KO_{n=2}^{-1} \cdot (-L_{n=2}) \cdot C_{n=1} \tag{6-5a}$$

$$D_{n=2} = KO_{n=2}^{-1} \cdot (-L_{n=2}) \cdot D_{n=1} \tag{6-6a}$$

mit

$$KO_{n=1} = \begin{bmatrix} a_{10} & b_{10} \\ a_{01} & b_{01} \end{bmatrix} \qquad LX_{n=1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad C_{n=1} = \begin{bmatrix} c_{10} \\ c_{01} \end{bmatrix}$$

$$LY_{n=1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad D_{n=1} = \begin{bmatrix} d_{10} \\ d_{01} \end{bmatrix}$$

$$KO_{n=2} = \begin{bmatrix} a_{10}^2 & a_{10}\,b_{10} & b_{10}^2 \\ 2\,a_{10}\,a_{01} & a_{10}\,b_{01} + a_{01}\,b_{10} & 2\,b_{10}\,b_{01} \\ a_{01}^2 & a_{01}\,b_{01} & b_{01}^2 \end{bmatrix}$$

$$L_{n=2} = \begin{bmatrix} a_{20} & b_{20} \\ a_{11} & b_{11} \\ a_{02} & b_{02} \end{bmatrix} \qquad C_{n=1} = \begin{bmatrix} c_{10} \\ c_{01} \end{bmatrix}$$

$$D_{n=1} = \begin{bmatrix} d_{10} \\ d_{01} \end{bmatrix}$$

Die Gleichungen zur Bestimmung der unbekannten Koeffizienten c_{ij} und d_{ij} unterscheiden sich nur in den Vektoren $LX_{n=1}$ und $LY_{n=1}$. Damit läßt sich eine einfache Beziehung zur Gewinnung der Unbekannten d_{ij} aus schon berechneten Koeffizienten c_{ij} angeben. Man ersetzt dazu in (6-5a) lediglich c_{ij} durch d_{ij} und berücksichtigt bei n = 1, daß

$$d_{10} = \frac{c_{10}}{b_{01}} \cdot (-b_{10})$$

und (6-7)

$$d_{01} = \frac{c_{01}}{-a_{01}} \cdot (-a_{10})$$

wird.

Für die Ordnungen $n \leq 4$ läßt sich die angegebene Lösungsstrategie ohne weiteres in Reduce umsetzen. Für $n \geq 5$ wird aus Gründen der Speicherplatzersparnis nurmehr $KO_{n \geq 5}$ und $L_{n \geq 5}$ aufgestellt und in einem fortrancompatiblen File abgelegt. Die Gleichungssysteme (6-3a) bzw. (6-4) für $n \geq 5$ werden dann rein numerisch aufgelöst.

7. REDUCE PROGRAMM ANWEISUNGEN

```
BEGIN
 ARRAY A( 6, 6), B( 6, 6), C( 6, 6):
               C(0,1):=C01;  C(0,2):=C02;  C(0,3):=C03 ; C(0,4):=C04;
                             C(0,5):=C05;  C(0,6):=C06;  C(1,5):=C15;
 C(1,0):=C10;  C(1,1):=C11;  C(1,2):=C12;  C(1,3):=C13;  C(1,4):=C14;
 C(2,0):=C20;  C(2,1):=C21;  C(2,2):=C22;  C(2,3):=C23;  C(2,4):=C24;
 C(3,0):=C30;  C(3,1):=C31;  C(3,2):=C32;  C(3,3):=C33;
 C(4,0):=C40;  C(4,1):=C41;  C(4,2):=C42;
 C(5,0):=C50;  C(5,1):=C51;  C(6.0):=C60;
               A(0,1):=A01;  A(0,2):=A02;  A(0,3):=A03 ; A(0,4):=A04;
                             A(0,5):=A05;  A(0,6):=A06;  A(1,5):=A15;
 A(1,0):=A10;  A(1,1):=A11;  A(1,2):=A12;  A(1,3):=A13;  A(1,4):=A14;
 A(2,0):=A20;  A(2,1):=A21;  A(2,2):=A22;  A(2,3):=A23; A(2,4):=A24;
 A(3,0):=A30;  A(3,1):=A31;  A(3,2):=A32; A(3,3):=A33;
 A(4,0):=A40;  A(4,1):=A41; A(4,2):=A42;
 A(5,0):=A50;  A(5,1):=A51; A(6.0):=A60;
               B(0,1):=B01;  B(0,2):=B02;  B(0,3):=B03 ; B(0,4):=B04;
                             B(0,5):=B05;  B(0,6):=B06; B(1,5):=B15;
 B(1,0):=B10;  B(1,1):=B11;  B(1,2):=B12;  B(1,3):=B13;  B(1,4):=B14;
 B(2,0):=B20;  B(2,1):=B21;  B(2,2):=B22;  B(2,3):=B23; B(2,4):=B24;
 B(3,0):=B30;  B(3,1):=B31;  B(3,2):=B32; B(3,3):=B33;
 B(4,0):=B40;  B(4,1):=B41;  B(4,2):=B42;
 B(5,0):=B50;  B(5,1):=B51;  B(6.0):=B60;
 COMMENT                     ORIGINALREIHEN********************;
 IORDM:=3;
 BEGIN INTEGER IORD;
 FOR IORD:=IORD STEP -1 UNTIL      1 DO BEGIN
 OUT T; WRITE" ABLEITUNG DER ",IORD," TEN ORDNUNG";OUT INVSO;
 WEIGHT U=1,V=1;
 WTLEVEL IORD;
 X:=0;    OFF RESUBS;
          FOR I:=0 STEP 1 UNTIL IORD DO BEGIN
          FOR J:=0 STEP 1 UNTIL IORD -I DO BEGIN
          X:=X+  A(I,J)* U**I * V**J;
          END J; END I;
 Y:=0;
          FOR I:=0 STEP 1 UNTIL IORD DO BEGIN
          FOR J:=0 STEP 1 UNTIL IORD -I DO BEGIN
          Y:=Y+  B(I,J)* U**I * V**J;
          END J; END I;
```

```
         ON RESUBS;
COMMENT              SUBSTITUIERTE REIHE ********************;
UI:=0;
          FOR I:=0 STEP 1 UNTIL IORD DO BEGIN
          FOR J:=0 STEP 1 UNTIL IORD -I DO BEGIN
          UI1:=C(I,J)* X**I * Y**J;
    WTLEVEL IORD-1;
          UI2:=C(I,J)* X**I * Y**J;
    WTLEVEL IORD;
            AKT:=UI1-UI2;
          UZ:=AKT+UI; CLEAR AKT,UI,UI1,UI2;
          UI:=UZ;  CLEAR UZ;
          END J; END I;
CLEAR X,Y;
        BEGIN ARRAY SU(55),SUV(55),ERG(1);
        BEGIN INTEGER ISPA,IZ;
COMMENT              ORDNEN NACH POTENZEN VON U UND V *******;
              COEFF(UI,U,SU) ; CLEAR UI;
        SUI:=0;
              FOR I:=1 STEP 1 UNTIL IORD+1 DO BEGIN
              SZ:=SUI+SU(I-1) ; CLEAR SUI;
              SUI:=SZ; CLEAR SZ;               END I; CLEAR SU;
              COEFF(SUI,V,SUV); CLEAR SUI;
              IZ:=IORD+1;
             BEGIN MATRIX KO(IZ,IZ),LS(IZ,1);LOES(IZ,1);
COMMENT               AUFBAU DES GLEICHUNGSSYSTEMS ***********;
               FOR I:=1 STEP 1 UNTIL IZ DO BEGIN
        SULS:=0 ;   ISPA:=0;
              FOR J:= IORD STEP -1 UNTIL 0 DO BEGIN
              K:=IORD-J;
              COEFF(SUV(I-1), C(J,K), ERG);
              ISPA:=ISPA+1;   KO(I,ISPA):=ERG(1);
              SZ:=SULS+ERG(1)*C(J,K) ;CLEAR SULS,ERG(0),ERG(1);
              SULS:=SZ;CLEAR SZ;
                                                 END J;
              LS(I,1):=SUV(I-1) - SULS;  CLEAR SULS;
                                                END I;
CLEAR SUV,ERG;
           IF IORD=1 THEN LS(2,1):=-1;
```

```
COMMENT              AUFLOESUNG DES GLEICHUNGSSYST. *********;
OFF GCD;
           LOES:=KO**(-1) *(-1)*LS;
ON GCD;
COMMENT              NATUERLICHE AUSGABE DER KOEFFIZIENTEN **;
           FOR I:= IORD STEP -1 UNTIL 0 DO BEGIN
               J:=IORD-I; CLEAR C(I,J);
               C(I,J):=LOES(J+1,1);
               WRITE "  C",I,J,"=  ",C(I,J); END I;
CLEAR KO,LS,LOES;         END  MAT;  END INT;
                          END AR;
END IORD;
COMMENT              FORTRANCOMPATIBLE AUSGABE DER KOEFF. ***;
ON LIST;OUT ROUTF; ON FORT; !*CARDNO:=1; !*BLKSIZE:=36;
    FOR K:=1 STEP 1 UNTIL IORDM DO BEGIN
           FOR I:= K    STEP -1 UNTIL 0 DO BEGIN
               J:= K  -I;
   WRITE C(I,J); WRITE "      C",I,J,"=ANS";
END I ;END K;OFF LIST;OFF FORT;              OUT INVSO;
END INT; END;
```

8. ERGEBNISAUSDRUCK (n ≤ 3)

```
                 3                    2                    2                    3
C30=  (B30*B01 *C01 - B21*B10*B01 *C01 - B20*B10*B01 *A01*C11 + B20*B01 *A10*
                        2                    2                          2
       C11 + B12*B10 *B01*C01 + B11*B10 *B01*A01*C11 - B11*B10*B01 *A10*C11
                 3                3                    3                 3
        - B10 *B03*C01 - B10 *B02*A01*C11 - B10 *A03*C10 - 2*B10 *A02*A01*C20
                 2                        2                      2
        + B10 *B02*B01*A10*C11 + B10 *B01*A12*C10 + 2*B10 *B01*A11*A01*C20 +
                 2                          2                      2
       2*B10 *B01*A10*A02*C20 - B10*B01 *A21*C10 - 2*B10*B01 *A20*A01*C20 - 2
                  2                    3                  3                    3    3
       *B10*B01 *A11*A10*C20 + B01 *A30*C10 + 2*B01 *A20*A10*C20)/(B10 *A01
                    2             2             2    2           3    3
          - 3*B10 *B01*A10*A01  + 3*B10*B01 *A10 *A01 - B01 *A10 )
                        2                                                2
C21=  ( - 3*B30*B01 *A01*C01 + 2*B21*B10*B01*A01*C01 + B21*B01 *A10*C01 - 2*
                      2                                2                3
       B20*B10*B01 *A01*C02 + 2*B20*B10*B01*A01 *C11 + 2*B20*B01 *A10*C02 - 2
                  2                        2
       *B20*B01 *A10*A01*C11 - B12*B10 *A01*C01 - 2*B12*B10*B01*A10*C01 + 2*
                 2                        2    2                      2
       B11*B10 *B01*A01*C02 - B11*B10 *A01 *C11 - 2*B11*B10*B01 *A10*C02 + 2*
                 2    2             3                    3                      2
       B11*B01 *A10 *C11 - 2*B10 *B02*A01*C02 - B10 *A02*A01*C11 + 3*B10 *B02
                              2                          2                      2
       *A10*C01 + 2*B10 *B02*B01*A10*C02 + 2*B10 *B02*A10*A01*C11 + B10 *B01*
                               2                        2                    2
       A11*A01*C11 + B10 *B01*A10*A02*C11 - B10 *A12*A01*C10 - 2*B10 *A11*
          2             2                        2
       A01 *C20 + 3*B10 *A10*A03*C10 + 4*B10 *A10*A02*A01*C20 - 2*B10*B02*B01
          2                2                          2
       *A10 *C11 - B10*B01 *A20*A01*C11 - B10*B01 *A11*A10*C11 + 2*B10*B01*
                                            2
       A21*A01*C10 + 4*B10*B01*A20*A01 *C20 - 2*B10*B01*A12*A10*C10 - 4*B10*
                 2                3                      2                    2
       B01*A10 *A02*C20 + B01 *A20*A10*C11 - 3*B01 *A30*A01*C10 + B01 *A21*
                              2                          2         2           3    3
       A10*C10 - 4*B01 *A20*A10*A01*C20 + 2*B01 *A11*A10 *C20)/(B10 *A01  - 2
          2             2             2    2           3    3
       *B10 *B01*A10*A01  + 3*B10*B01 *A10 *A01 - B01 *A10 )
                        2                    2
C12=  (3*B30*B01*A01 *C01 - B21*B10*A01 *C01 - 2*B21*B01*A10*A01*C01 + 4*B20*
                      2                    3                  2
       B10*B01*A01 *C02 - B20*B10*A01 *C11 - 4*B20*B01 *A10*A01*C02 + B20*B01
                  2                                                2                2
       *A10*A01 *C11 + 2*B12*B10*A10*A01*C01 + B12*B01*A10 *C01 - 2*B11*B10 *
          2                        2                  2    2                    2
       A01 *C02 + B11*B10*A10*A01 *C11 + 2*B11*B01 *A10 *C02 - B11*B01*A10 *
```

```
                                  2                                2          2          2
        A01*C11 + 4*B10 *B01*A10*A01*C02 - B10 *A11*A01 *C11 + 2*B10 *A10*A02*

                                  2                                2                   2
        A01*C11 - 3*B10*B03*A10 *C01 - 4*B10*B02*B01*A10 *C02 - B10*B02*A10 *

                                  2                                2
        A01*C11 + 2*B10*B01*A20*A01 *C11 - 2*B10*B01*A10 *A02*C11 - B10*A21*

           2                      3
        A01 *C10 - 2*B10*A20*A01 *C20 + 2*B10*A12*A10*A01*C10 + 2*B10*A11*A10*

           2                 2                      2                        3
        A01 *C20 - 3*B10*A10 *A03*C10 - 2*B10*A10 *A02*A01*C20 + B02*B01*A10 *

                  2                          2          2                     2
        C11 - 2*B01 *A20*A10*A01*C11 + B01 *A11*A10 *C11 + 3*B01*A30*A01 *C10

         - 2*B01*A21*A10*A01*C10 + 2*B01*A20*A10*A01 *C20 + B01*A12*A10 *C10

                        2                     3                3    3        2
         - 2*B01*A11*A10 *A01*C20 + 2*B01*A10 *A02*C20)/(B10 *A01  - 3*B10 *

                    2          2    2          3    3
           B01*A10*A01  + 3*B10*B01 *A10 *A01 - B01 *A10 )

                  3                    2                      3
C03=   ( - B30*A01 *C01 + B21*A10*A01 *C01 - 2*B20*B10*A01 *C02 + 2*B20*B01*

                 2                 2                          2
        A10*A01 *C02 - B12*A10 *A01*C01 + 2*B11*B10*A10*A01 *C02 - 2*B11*B01*

           2                      2                      3
        A10 *A01*C02 - 2*B10*B02*A10 *A01*C02 - B10*A20*A01 *C11 + B10*A11*A10

           2                 2                    3                      3
        *A01 *C11 - B10*A10 *A02*A01*C11 + B03*A10 *C01 + 2*B02*B01*A10 *C02

                        2                      2                    3
         + B01*A20*A10*A01 *C11 - B01*A11*A10 *A01*C11 + B01*A10 *A02*C11 -

                  3                    2                 2                3
        A30*A01 *C10 + A21*A10*A01 *C10 - A12*A10 *A01*C10 + A10 *A03*C10)/(

             3    3          2             2             2    2           3    3
          B10 *A01  - 3*B10 *B01*A10*A01  + 3*B10*B01 *A10 *A01 - B01 *A10 )

                  2                             2                2
C20=   ( - B20*B01 *C01 + B11*B10*B01*C01 - B10 *B021*C01 - B10 *A02*C10 + B10*

                        2                2    2                          2
        B01*A11*C10 - B01 *A20*C10)/(B10 *A01  - 2*B10*B01*A10*A01 + B01 *

          2
        A10 )

C11=   (2*B20*B01*A01*C01 - B11*B10*A01*C01 - B11*B01*A10*C01 + 2*B10*B02*A10*

        C01 - B10*A11*A01*C10 + 2*B10*A10*A02*C10 + 2*B01*A20*A01*C10 - B01*

                          2    2                          2    2
        A11*A10*C10)/(B10 *A01  - 2*B10*B01*A10*A01 + B01 *A10 )

                  2                          2                2
C02=   ( - B20*A01 *C01 + B11*A10*A01*C01 - B02*A10 *C01 - A20*A01 *C10 + A11*

                        2                2    2                          2
        A10*A01*C10 - A10 *A02*C10)/(B10 *A01  - 2*B10*B01*A10*A01 + B01 *

          2
        A10 )

C10=   B10/(B10*A01 - B01*A10)

C01=   ( - A10)/(B10*A01 - B01*A10)
```

9. FORTRANCOMPATIBLES AUSGABEFILE

```
      ANS=B10/(B10*A01-B01*A10)
      C10 = ANS
      ANS=(-A10)/(B10*A01-B01*A10)
      C01 = ANS
      ANS2=-B10**2*A02*C10+B10*B01*A11*C10-B01**2*A20*C10
      ANS1=-B20*B01**2*C01+B11*B10*B01*C01-B10**2*B02*C01+ANS2
      ANS3=B10**2*A01**2-2.*B10*B01*A10*A01+B01**2*A10**2
      ANS=ANS1/ANS3
      C20 = ANS
      ANS3=2.*B01*A20*A01*C10-B01*A11*A10*C10
      ANS2=2.*B10*B02*A10*C01-B10*A11*A01*C10+2.*B10*A10*A02*C10+ANS3
      ANS1=2.*B20*B01*A01*C01-B11*B10*A01*C01-B11*B01*A10*C01+ANS2
      ANS4=B10**2*A01**2-2.*B10*B01*A10*A01+B01**2*A10**2
      ANS=ANS1/ANS4
      C11 = ANS
      ANS2=-A20*A01**2*C10+A11*A10*A01*C10-A10**2*A02*C10
      ANS1=-B20*A01**2*C01+B11*A10*A01*C01-B02*A10**2*C01+ANS2
      ANS3=B10**2*A01**2-2.*B10*B01*A10*A01+B01**2*A10**2
      ANS=ANS1/ANS3
      C02 = ANS
      ANS10=B01**3*A30*C10+2.*B01**3*A20*A10*C20
      ANS9=-2.*B10*B01**2*A20*A01*C20-2.*B10*B01**2*A11*A10*C20+ANS10
      ANS8=2.*B10**2*B01*A10*A02*C20-B10*B01**2*A21*C10*ANS9
      ANS7=B10**2*B01*A12*C10+2.*B10**2*B01*A11*A01*C20*ANS8
      ANS6=-2.*B10**3*A02*A01*C20+B10**2*B02*B01*A10*C11+ANS7
      ANS5=-B10**3*B02*A01*C11-B10**3*A03*C10+ANS6
      ANS4=-B11*B10*B01**2*A10*C11-B10**3*B03*C01+ANS5
      ANS3=B12*B10**2*B01*C01+B11*B10**2*B01*A01*C11+ANS4
      ANS2=-B20*B10*B01**2*A01*C11+B20*B01**3*A10*C11+ANS3
      ANS1=B30*B01**3*C01-B21*B10*B01**2*C01+ANS2
      ANS12=3.*B10*B01**2*A10**2*A01-B01**3*A10**3
      ANS11=B10**3*A01**3*3.*B10**2*B01*A10*A01**2+ANS12
      ANS=ANS1/ANS11
      C30 = ANS
      ANS18=-4.*B01**2*A20*A10*A01*C20+2.*B01**2*A11*A10**2*C20
      ANS17=-3.*B01**2*A30*A01*C10+B01**2*A21*A10*C10+ANS18
      ANS16=-4.*B10*B01*A10**2*A02*C20+B01**3*A20*A10*C11+ANS17
      ANS15=4.*B10*B01*A20*A01**2*C20-2.*B10*B01*A12*A10*C10+ANS16
      ANS14=-B10*B01**2*A11*A10*C11+2.*B10*B01*A21*A01*C10+ANS15
      ANS13=-2.*B10*B02*B01*A10**2*C11-B10*B01**2*A20*A01*C11+ANS14
      ANS12=3.*B10**2*A10*A03*C10+4.*B10**2*A10*A02*A01*C20+ANS13
      ANS11=-B10**2*A12*A01*C10-2.*B10**2*A11*A01**2*C20+ANS12
      ANS10=B10**2*B01*A11*A01*C11+B10**2*B01*A10*A02*C11+ANS11
      ANS9=2.*B10**2*B02*B01*A10*C02+2.*B10**2*B02*A10*A01*C11+ANS10
      ANS8=-B10**3*A02*A01*C11+3.*B10**2*B03*A10*C01+ANS9
      ANS7=b11*B01**2*A10**2*C11-2.*B10**3*B02*A01*C02+ANS8
      ANS6=-B11*B10**2*A01**2*C11-2.*B11*B10*B01**2*A10*C02+ANS7
      ANS5=-2.*B12*B10*B01*A10*C01+2.*B11*B10**2*B01*A01*C02+ANS6
      ANS4=-2.*B20*B01**2*A10*A01*C11-B12*B10**2*A01*C01+ANS5
      ANS3=2.*B20*B10*B01*A01**2*C11+2.*B20*B01**3*A10*C02+ANS4
      ANS2=B21*B01**2*A10*C01-2.*B20*B10*B01**2*A01*C02+ANS3
      ANS1=-3.*B30*B01**2*A01*C01+2.*B21*B10*B01*A01*C01+ANS2
      ANS20=3.*B10*B01**2*A10**2*A01-B01**3*A01**3
      ANS19=B10**3*A01**3-3.*B10**2*B01*A10*A01**2+ANS20
      ANS=ANS1/ANS19
      C21 = ANS
      ANS18=-2.*B01*A11*A10**2*A01*C20+2.*B01*A10**3*A02*C20
      ANS17=2.*B01*A20*A10*A01**2*C20+B01*A12*A10**2*C10+ANS18
      ANS16=3.*B01*A30*A01**2*C10-2.*B01*A21*A10*A01*C10+ANS17
      ANS15=-2.*B01**2*A20*A10*A01*C11+B01**2*A11*A10**2*C11+ANS16
      ANS14=-2.*B10*A10**2*A02*A01*C20+B02*B01*A10**3*C11+ANS15
      ANS13=2.*B10*A11*A10*A01**2*C20-3.*B10*A10**2*A03*C10+ANS14
      ANS12=-2.*B10*A20*A01**3*C20+2.*B10*A12*A10*A01*C10+ANS13
      ANS11=-2.*B10*B01*A10**2*A02*C11-B10*A21*A01**2*C10+ANS12
      ANS10=-B10*B02*A10**2*A01*C11+2.*B10*B01*A20*A01**2*C11+ANS11
      ANS9=-3.*B10*B03*A10**2*C01-4.*B10*B02*B01*A10**2*C02+ANS10
      ANS8=-B10**2*A11*A01**2*C11+2.*B10**2*A10*A02*A01*C11+ANS9
      ANS7=-B11*B01*A10**2*A01*C11+4.*B10**2*B02*A10*A01*C02+ANS8
      ANS6=B11*B10*A10*A01**2*C11+2.*B11*B01**2*A10**2*C02+ANS7
      ANS5=B12*B01*A10**2*C01-2.*B11*B10**2*A01**2*C02+ANS6
       ⋮
```

10. RECHENBEISPIEL

```
A01= -.7092698198E-05          C10 =  .1000387081E+01
A02= -.5015497109E-11          C01 =  .7098218033E-05
A03= -.3156113334E-17          C20 = -.1035207723E-10
A04= -.7335016873E-22          C11 =  .1426211536E-10
A05= -.2716600776E-28          C02 =  .5021453347E-11
A06=  .1688648318E-33          C30 = -.1184557927E-15
A10=  .9996130690              C21 =  .7789958971E-18
A11= -.1424597097E-10          C12 = -.5114478950E-16
A12=  .5106478688E-16          C03 =  .3160691065E-17
A13=  .1690336265E-21          C40 = -.3155432848E-21
A14= -.1437580520E-27          C31 = -.6415790289E-21
A15= -.5242857708E-23          C22 =  .5480308467E-22
A20=  .1033993031E-10          C13 = -.1693672629E-21
A21= -.7792943490E-18          C04 =  .7349206342E-22
A22= -.5470029296E-22          C50 =  .1280808102E-26
A23= -.1315251295E-27          C41 =  .1579797195E-27
A24=  .1081359085E-32          C32 =  .4013304855E-27
A30=  .1182734564E-15          C23 =  .1318484064E-27
A31=  .6403324352E-21          C14 =  .1440863833E-27
A32= -.4004257193E-27          C05 =  .2723092426E-28
A33= -.8984499696E-33          C60 =  .4327061305E-32
A40=  .3149475892E-21          C51 =  .8554337853E-32
A41= -.1575617739E-27          C42 =  .2875553813E-32
A42= -.2867481436E-32          C33 =  .9012181756E-33
A50= -.1277788023E-26          C24 = -.1084345344E-32
A51= -.8530581670E-32          C15 =  .5257250507E-33
A60= -.4315360349E-32          C06 = -.1693278181E-33

B01=  .9996091430              D10 = -.1130412103E-04
B02=  .1529333647E-11          D01 =  .1000391010E+01
B03=  .5960105100E-17          D20 =  .3671789347E-11
B04= -.2191574770E-22          D11 = -.2828106023E-10
B05= -.6407959514E-28          D02 = -.1531385710E-11
B06=  .4973183580E-34          D30 =  .8988223651E-16
B10=  .1129533053E-04          D21 =  .9385826560E-16
B11=  .2824794822E-10          D12 = -.2646192470E-16
B12=  .2642255421E-16          D03 = -.5969793158E-17
B13= -.3980863818E-22          D40 = -.4729511184E-22
B14= -.1653786080E-27          D31 =  .1842482307E-21
B15=  .4578328787E-33          D22 =  .1409294825E-21
B20= -.3667078371E-11          D13 =  .3989216248E-22
B21= -.9370866472E-16          D04 =  .2195790366E-22
B22= -.1406477969E-21          D50 = -.1037948924E-26
B23=  .5083238457E-27          D41 = -.1007191929E-26
B24=  .1404680777E-32          D32 =  .8041395093E-28
B30= -.8974275384E-16          D23 = -.5095319068E-27
B31= -.1838884067E-21          D14 =  .1657604740E-27
B32= -.8025165527E-28          D05 =  .6422888931E-28
B33= -.1579932931E-32          D60 = -.7893472861E-33
B40=  .4719913338E-22          D51 = -.2832528819E-32
B41=  .1004776345E-26          D42 = -.3000033471E-33
B42=  .2988233112E-33          D33 =  .1584283436E-32
B50=  .1035512623E-26          D24 = -.1408508306E-32
B51=  .2824479104E-32          D15 = -.4591159481E-33
B60=  .7872299793E-33          D06 = -.4987282041E-34
```

11. VERGLEICH DER BEIDEN VERFAHREN

Eine abschließende Wertung der Verfahren soll dem interessierten Anwender überlassen bleiben, der sich, je nach Programmiererfahrung, lieber ein kompliziertes Fortranprogramm anhand der Rekursionsformeln zusammenstellt oder aber einen übersichtlichen Lösungsweg in Reduce umsetzt, mit dem Nachteil, daß bei großem Speicherplatzbedarf erst im zweiten Schritt ein Fortranprogramm entsteht, das allerdings wegen seiner automatischen Erstellung sofort fehlerfrei ist.

Ein objektiver Vergleich ist über die Rechenzeiten möglich.

Rechenzeiten: n =	3	4	5	6	
PT(N) :	0,6	1,5	7,8	54,5	sek
PT(R) :	58,8	7:59,0	14:8,7	27:0	min:sek
PT(P) :	0,1	0,2	0,3	0,6	sek

Ein Vergleich der Rechenzeiten zeigt, daß die numerische Lösung (N) gegenüber der Erzeugung des algebraischen Programms mit Reduce (R) wesentlich schneller ist, während sie gegenüber der Rechnung mit diesem Programm (P) wesentlich langsamer ist. Die algebraische Lösung ist daher für einen vorgegebenen Grad der Reihe ab einer bestimmten Anzahl X von Benutzungen rechenzeit-günstiger. Mit

$$X \cdot PT(N) = PT(R) + X \cdot PT(P)$$

wird

$$X = PT(R) / \left(PT(N) - PT(P)\right) \ .$$

Die Bevorzugung der einen oder anderen Lösung könnte daher, wenn das Programmsystem Reduce zur Verfügung steht, in Abhängigkeit von dem zur Verfügung stehenden Speicherplatz und von der Benutzungshäufigkeit bestimmt werden.

LITERATUR

HEARN, A.C.: ***Reduce, A User oriented Interactive System for Algebraic Simplification***. In Interactive Systems for Experimental Applied Mathematics, Academic Press New York 1968

KÖNIG-WEISE: ***Mathematische Grundlagen der Höheren Geodäsie und Kartographie, Band I.*** Springer, Berlin/Göttingen/Heidelberg 1951

ZUR BESTIMMUNG DER SIGNIFIKANTEN PARAMETER IN APPROXIMATIONSFUNKTIONEN

von Peter SCHWINTZER

SUMMARY

To solve the problem of how to find the significant parameters in empirical functions, a modified Cholesky-decomposition of the quadratic form of the parameter vector is used. The decomposition results in a set of uncorrelated values, which can be investigated one by one in a multiple statistical test. This procedure finally leads to the desired combination of the relevant parameters.

A numerical example demonstrates the steps of computations and the practicability of the proposed method.

ZUSAMMENFASSUNG

Bei der Anwendung von Approximationsfunktionen stellt sich häufig das Problem, Aussagen über die Signifikanz der gewählten Parameter zu treffen. Da die Parameter durch das Ausgleichungsverfahren in der Regel algebraisch korreliert sind, können sie nicht unabhängig voneinander betrachtet werden. In diesem Beitrag wird vorgeschlagen, die quadratische Form der Parameter deshalb über eine modifizierte Cholesky-Zerlegung in unkorrelierte und gleichgenaue Anteile zu zerlegen. In einem multiplen statistischen Test können diese Anteile anschließend einzeln auf Signifikanz untersucht werden. Das Ergebnis ist die gesuchte Kombination der Parameter, die modellrelevant sind.

An einem praktischen Beispiel wird der Ablauf der Berechnungen und die Wirksamkeit des Verfahrens gezeigt.

1. EINLEITUNG

Eine Approximationsfunktion oder eine ausgleichende Funktion beschreibt den funktionalen Zusammenhang zwischen einer diskreten Zahl von Meßwerten und gewissen Einflußgrößen. Über den geeigneten Funktionstyp (z.B. Polynomreihe, Fourrierreihe, Exponentialfunktion) muß a priori aufgrund bekannter Gesetzmäßigkeiten oder anhand des graphischen Verlaufs der Meßwerte entschieden werden. Die unbekannten Parameter der Funktion werden in einer Ausgleichung geschätzt.

Approximationsfunktionen finden in vielen Bereichen der Geodäsie Verwendung, z.B. bei der Kalibrierung geodätischer Instrumente (***SCHWARZ*** 1983), bei der Analyse von Deformationsmessungen (***FLOTRON*** 1981), bei der Grundstücksbewertung (***BAUER*** 1978) oder bei der Messung physikalischer Vorgänge (***JAKOB*** 1979).

Je nach Aufgabenstellung wird die ausgleichende Funktion für die Interpolation bzw. Extrapolation von Funktionswerten oder als generalisierende Funktion für die Interpretation einer gemessenen Erscheinung benötigt.

Da die a priori Festlegung auf Art und Zahl der Parameter in der ausgleichenden Funktion immer mit gewissen Unsicherheiten verbunden ist, ist es sinnvoll, nach der Ausgleichung zu untersuchen, welche der Parameter wirklich modellrelevant, d.h. signifikant sind und welche aus der Ausgleichung wieder eliminiert werden können. Die Reduzierung des Parametervektors um die nichtsignifikanten Parameter erhöht den Freiheitsgrad in der Ausgleichung und führt zu zuverlässigeren Schätzwerten für die verbliebenen Parameter und deren Genauigkeiten. Außerdem werden die Interpretationsmöglichkeiten verbessert, und es wird, was für Interpolationszwecke wichtig ist, zwischen den Stützpunkten ein glatterer Kurvenverlauf erzielt.

2. UNTERSCHIEDLICHE AUSGLEICHUNGSFORMEN BEI APPROXIMATIONSFUNKTIONEN

Der funktionale Zusammenhang läßt sich allgemein darstellen durch

$$y_i = f(\boldsymbol{p}^T, \boldsymbol{x}_i^T) \quad . \tag{2-1}$$

Hierin bedeuten:

y_i - Funktionswert

$\boldsymbol{p}$ - u x 1 Vektor der gesuchten Parameter

$\boldsymbol{x}_i$ - bekannter m x 1 Vektor der Einflußgrößen $\left[\boldsymbol{x}_i^T = (x_{1_i} \ldots x_{m_i})\right]$

Enthält der Vektor $\boldsymbol{x}_i$ mehr als eine Komponente, so spricht man von einem mehrdimensionalen Problem.

Ist die Funktion linear in den Parametern, so stellen die Funktionswerte Linearkombinationen der Komponenten des Parametervektors dar

$$y_i = \boldsymbol{g}^T(\boldsymbol{x}_i^T) \cdot \boldsymbol{p} \quad . \tag{2-2}$$

Bei nichtlinearen Ausgangsfunktionen muß vor der Ausgleichung nach Einführung von Näherungswerten für die Parameter ($\boldsymbol{p} = \boldsymbol{p}_0 + d\boldsymbol{p}$) linearisiert werden:

$$y_i = f(\boldsymbol{p}_0^T, \boldsymbol{x}_i^T) + \left(\frac{\partial f(\boldsymbol{p}^T, \boldsymbol{x}_i^T)}{\partial \boldsymbol{p}}\right)_0^T \cdot d\boldsymbol{p} \tag{2-3}$$

Bezeichnet man den u x 1 Vektor der Differentialquotienten mit $\overline{\boldsymbol{g}}(\boldsymbol{p}_0^T, \boldsymbol{x}_i^T)$ und substituiert $\overline{y}_i = y_i - f(\boldsymbol{p}_0^T, \boldsymbol{x}_i^T)$, so erhält man eine zu (2-2) analoge Darstellung

$$\overline{y}_i = \overline{\boldsymbol{g}}^T(\boldsymbol{p}_0^T, \boldsymbol{x}_i^T) \cdot d\boldsymbol{p} \quad . \tag{2-4}$$

Zu diesem funktionalen Modell gehören je nach Art des Problems unterschiedliche stochastische Modellvorstellungen, so daß sich die folgenden Ausgleichungsfälle unterscheiden lassen:

a) Die Funktionswerte y_i sind gemessene fehlerbehaftete Größen; die Einflußgrößen sind nichtstochastischer Art oder können aufgrund ihrer hohen Genauigkeit praktisch als fehlerfrei betrachtet werden.

   Daraus ergibt sich mit $\boldsymbol{e}$ als Fehlervektor für die Ausgleichung das

Modell

$$\bar{\boldsymbol{y}} - \boldsymbol{e} = \boldsymbol{\mathcal{A}} \cdot d\boldsymbol{p} \quad , \quad \mathcal{P}_{yy} = \mathcal{Q}_{yy}^{-1} = \mathcal{Q}_{ee}^{-1} \tag{2-5}$$

Die Anwendung der Methode der kleinsten Quadrate auf das Verbesserungsgleichungssystem

$$\bar{\boldsymbol{y}} + \boldsymbol{v} = \boldsymbol{\mathcal{A}} \cdot d\hat{\boldsymbol{p}} \quad , \quad \mathcal{P}_{yy} = \mathcal{Q}_{yy}^{-1} \tag{2-6}$$

liefert die Schätzungen $d\hat{\boldsymbol{p}}$ und $\hat{\boldsymbol{e}} = -\boldsymbol{v}$ für den unbekannten Parametervektor $d\boldsymbol{p}$ und den Fehlervektor $\boldsymbol{e}$.

$\bar{\boldsymbol{y}}$ ist der n x 1 Beobachtungsvektor (abzüglich der mit den Näherungswerten berechneten Funktionswerte), $\boldsymbol{v}$ der n x 1 Verbesserungsvektor und $\boldsymbol{\mathcal{A}}$ die n x u Koeffizientenmatrix. Die Koeffizientenmatrix $\boldsymbol{\mathcal{A}}$ enthält zeilenweise die Vektoren $\bar{\boldsymbol{g}}^{T}(\boldsymbol{p}_0^{T}, \boldsymbol{x}_i^{T})$ der Differentialquotienten. Die Beobachtungsgenauigkeit wird bis auf den zu schätzenden Varianzfaktor durch die Kofaktorenmatrix $\mathcal{Q}_{yy}$ ausgedrückt.

b) Ist eine Trennung in unabhängige und abhängige Variable von der Problemstellung her nicht gegeben, so sind sowohl die Komponenten des Vektors $\boldsymbol{y}$ als auch die Komponenten des Vektors $\boldsymbol{x}^{T} = (\boldsymbol{x}_1^{T}, \boldsymbol{x}_2^{T}, \ldots, \boldsymbol{x}_n^{T})$ als fehlerhafte Beobachtungen zu behandeln. Ein Anwendungsbeispiel findet man bei ***CASPARY/SCHWINTZER*** (1981 b).

Dem Vektor $\boldsymbol{y}$ wird die n x n Varianz-Kovarianz-Matrix $s_0^2 \mathcal{Q}_{yy}$ und dem n x m Vektor $\boldsymbol{x}$ die (n x m) x (n x m) Varianz-Kovarianz-Matrix $s_0^2 \mathcal{Q}_{xx}$ zugeordnet. In (2-3) ist dann zusätzlich noch nach den Komponenten des Vektors $\boldsymbol{x}$ partiell zu differenzieren. Die dabei entstehende Bedingungsgleichung mit Unbekannten lautet:

$$y_i - f(\boldsymbol{p}_0^{T}, \boldsymbol{x}_i^{T}) - e_{y_i} = \left(\frac{\partial f(\boldsymbol{p}^{T}, \boldsymbol{x}_i^{T})}{\partial \boldsymbol{p}}\right)_0 \cdot d\boldsymbol{p} + \left(\frac{\partial f(\boldsymbol{p}^{T}, \boldsymbol{x}_i^{T})}{\partial \boldsymbol{x}_i}\right)_{x_i} \cdot (-\boldsymbol{e}_{x_i}) \tag{2-7}$$

$$y_i \qquad - e_{y_i} = \bar{\boldsymbol{g}}^{T}(\boldsymbol{p}_0^{T}, \boldsymbol{x}_i^{T}) \qquad \cdot d\boldsymbol{p} + \bar{\boldsymbol{h}}^{T}(\boldsymbol{p}^{T}, \boldsymbol{x}_i^{T}) \qquad \cdot (-\boldsymbol{e}_{x_i})$$

Nach Übergang auf die Matrizendarstellung existiert für das System

der Bedingungsgleichungen die Darstellung

$$\bar{\boldsymbol{y}} - \boldsymbol{e}_y = \boldsymbol{\mathcal{A}} \cdot d\boldsymbol{p} + \boldsymbol{\mathcal{B}} \cdot \left(-\boldsymbol{e}_y\right) \quad , \tag{2-8}$$

in der $\boldsymbol{\mathcal{B}}$ eine n x (n x m) Matrix ist, die zeilenweise die Vektoren $\bar{\boldsymbol{h}}^T\left(\boldsymbol{p}^T, \boldsymbol{x}_i^T\right)$ der Differentialquotienten enthält. Für die Komponenten des Vektors $\boldsymbol{x}$, die in der i-ten Bedingungsgleichung nicht auftreten, ist eine Null einzutragen. Die Parameterschätzung kann über eine Ausgleichung nach bedingten Beobachtungen mit Unbekannten erfolgen. Sie läßt sich jedoch auch auf eine Ausgleichung nach vermittelnden Beobachtungen zurückführen. Substituiert man nämlich

$$\boldsymbol{e} = \boldsymbol{e}_y - \boldsymbol{\mathcal{B}}\boldsymbol{e}_x \tag{2-9}$$

und wendet auf diese Gleichung das Fehlerfortpflanzungsgesetz an

$$\boldsymbol{Q}_{ee} = \boldsymbol{Q}_{yy} + \boldsymbol{\mathcal{B}}\,\boldsymbol{Q}_{xx}\boldsymbol{\mathcal{B}}^T \quad , \tag{2-10}$$

so läßt sich das Modell nach Übergang auf die Schätzwerte mit $\boldsymbol{v} = -\hat{\boldsymbol{e}}$ in einer (2-6) entsprechenden Form formulieren:

$$\bar{\boldsymbol{y}} + \boldsymbol{v} = \boldsymbol{\mathcal{A}} \cdot d\hat{\boldsymbol{p}} \qquad , \qquad \boldsymbol{\mathcal{P}}_{yy} = \boldsymbol{Q}_{ee}^{-1} = \left(\boldsymbol{Q}_{yy} + \boldsymbol{\mathcal{B}}\,\boldsymbol{Q}_{xx}\boldsymbol{\mathcal{B}}^T\right)^{-1} \tag{2-11}$$

Da die Matrix $\boldsymbol{\mathcal{B}}$ und damit auch die Gewichtsmatrix $\boldsymbol{\mathcal{P}}_{yy}$ von den unbekannten Parametern p_i abhängig sind, erhält man die endgültigen Gewichte iterativ mit schrittweise verbesserten Näherungswerten für p_i. Bei der ausgleichenden Geraden müssen die Verbesserungsgleichungen erst mit den nach (2-10) zu bestimmenden Gewichten homogenisiert werden. Die sich dann ergebende nichtlineare funktionale Beziehung ist zu linearisieren und kann in eine der vermittelnden Ausgleichung angepaßte Form gebracht werden.

c) Die Funktionswerte y_i sind als praktisch fehlerfrei zu betrachten; die Variablen x_i sind fehlerfreie Beobachtungen.

Die Behandlung dieses Falls ergibt sich durch Spezialisierung der unter (2-8) bis (2-11) genannten Gleichungen, wenn $\boldsymbol{e}_y = 0$ und

$Q_{yy} = \mathcal{O}$ gesetzt werden.

d) Das Modell (2-6) wird um einen n x 1 Vektor z von Zufallsparametern erweitert, der nicht funktional erfaßbare, zufällig wirkende Modellrestfehler auffangen soll:

$$\bar{y} - e_y = \mathcal{A} \cdot d\mathcal{p} + z \tag{2-12}$$

Als a priori Information werden der Erwartungswert von z und die Varianz-Kovarianz-Matrix als bekannt vorausgesetzt

$$E(z) = \mathcal{O} \quad , \quad Var(z) = \sigma_0^2 Q_{zz} \ . \tag{2-13}$$

Zwischen den Meßfehlern und den Zufallsparametern bestehen keine Korrelationen

$$Kov(e_y, z) = \mathcal{O} \ . \tag{2-14}$$

Ein solches Modell wird wegen der Kombination von deterministischen Parametern und Zufallsparametern als gemischtes Modell und in Verbindung mit Interpolation und Extrapolation auch als Kollokationsmodell bezeichnet.

Auch dieses Modell läßt sich wieder auf die Form (2-6) zurückführen, indem man

$$e = e_y + z \tag{2-15}$$

substituiert und nach Anwendung des Fehlerfortpflanzungsgesetzes

$$Q_{ee} = Q_{yy} + Q_{zz} \tag{2-16}$$

auf das Modell

$$\bar{y} - e = \mathcal{A} \cdot d\mathcal{p} \ , \ \mathcal{P}_{yy} = Q_{ee}^{-1} \tag{2-17}$$

übergeht. Für die Ausgleichung ergibt sich das Verbesserungsglei-

chungssystem:

$$\bar{\boldsymbol{y}} + \boldsymbol{v} = \boldsymbol{A} \cdot d\hat{\boldsymbol{p}} \quad , \quad \boldsymbol{P}_{yy} = \left(\boldsymbol{Q}_{yy} + \boldsymbol{Q}_{zz}\right)^{-1} \tag{2-18}$$

Die Parameter können in allen Fällen über eine Ausgleichung nach vermittelnden Beobachtungen geschätzt werden. Unterschiede bestehen formal lediglich in der a priori aufzustellenden Gewichtsmatrix $\boldsymbol{P}_{yy}$ des Beobachtungsvektors $\bar{\boldsymbol{y}}$. Die folgenden Ableitungen schließen daher alle vier Fallunterscheidungen ein.

3. QUADRATISCHE FORMEN

Zwischen den quadratischen Formen der Vektoren $\boldsymbol{v}$, $\bar{\boldsymbol{y}}$ und $d\hat{\boldsymbol{p}}$ besteht in einer Ausgleichung nach vermittelnden Beobachtungen die grundlegende Beziehung

$$\boldsymbol{v}^T \boldsymbol{P}_{yy}\, \boldsymbol{v} = \bar{\boldsymbol{y}}^T \boldsymbol{P}_{yy}\, \bar{\boldsymbol{y}} - d\hat{\boldsymbol{p}}^T \boldsymbol{Q}_{\hat{p}\hat{p}}^{-1}\, d\hat{\boldsymbol{p}} \tag{3-1}$$

mit $\boldsymbol{Q}_{\hat{p}\hat{p}}^{-1} = \boldsymbol{N} = \boldsymbol{A}^T \boldsymbol{P}_{yy} \boldsymbol{A}$ als Normalgleichungsmatrix. Bei einem in den Parametern linearen Funktionsansatz (z.B. Polynomansatz) vereinfacht sich die Beziehung mit $\boldsymbol{p}_0 = \boldsymbol{0}$ zu

$$\boldsymbol{v}^T \boldsymbol{P}_{yy}\, \boldsymbol{v} = \boldsymbol{y}^T \boldsymbol{P}_{yy}\, \boldsymbol{y} - \hat{\boldsymbol{p}}^T \boldsymbol{Q}_{\hat{p}\hat{p}}^{-1}\, \hat{\boldsymbol{p}} \quad . \tag{3-2}$$

Bei nichtlinearen Funktionen ist es zuweilen sinnvoll, für eine Gruppe der Parameter Näherungswerte einzuführen, $\hat{\boldsymbol{p}}_1 = (\boldsymbol{p}_1)_0 + d\hat{\boldsymbol{p}}_1$, während für eine zweite Gruppe, deren Signifikanz nicht evident ist, die Näherungswerte Null ausreichend sind ($\hat{\boldsymbol{p}}_2 = \boldsymbol{0} + \hat{\boldsymbol{p}}_2$). Nur diese Parametergruppe soll dann in die Signifikanzuntersuchungen einbezogen werden. Partitioniert man den Parametervektor $d\hat{\boldsymbol{p}}$ und die Formmatrix $\boldsymbol{N} = \boldsymbol{Q}_{\hat{p}\hat{p}}^{-1}$ entsprechend den beiden Parametergruppen $d\hat{\boldsymbol{p}}_1$ und $\hat{\boldsymbol{p}}_2$ mit u_1 und u_2 Komponenten, so lautet die quadratische Form der Parameter

$$q_{\hat{p}} = d\hat{\boldsymbol{p}}^T \boldsymbol{Q}_{\hat{p}\hat{p}}^{-1}\, d\hat{\boldsymbol{p}} = \left(d\hat{\boldsymbol{p}}_1^T \hat{\boldsymbol{p}}_2^T\right) \begin{pmatrix} \boldsymbol{N}_{11} & \boldsymbol{N}_{12} \\ \boldsymbol{N}_{21} & \boldsymbol{N}_{22} \end{pmatrix} \cdot \begin{pmatrix} d\hat{\boldsymbol{p}}_1 \\ \hat{\boldsymbol{p}}_2 \end{pmatrix} \quad . \tag{3-3}$$

Diese quadratische Form läßt sich in zwei unabhängige Summanden zerlegen

$$q_{\hat{p}} = \left(d\hat{p}_1 + N_{11}^{-1}N_{12}\hat{p}_2\right)^T N_{11}\left(d\hat{p}_1 + N_{11}^{-1}N_{12}\hat{p}_2\right) + \hat{p}_2^T\left(N_{22} - N_{21}N_{11}^{-1}N_{12}\right)\hat{p}_2 \qquad (3\text{-}4)$$

wie man sich durch Ausmultiplizieren leicht überzeugt. Wegen $N_{21}N_{11}^{-1}N_{12} = Q_{22}^{-1}$, wobei Q_{22} die auf die Parametergruppe $\hat{p}_2$ bezogene Submatrix der Kofaktorenmatrix $Q_{\hat{p}\hat{p}}$ bedeutet, und mit den Substitutionen $d\bar{\hat{p}}_1 = d\hat{p}_1 + N_{11}^{-1}N_{12}\hat{p}_2$ und $N_{11} = \bar{Q}_{11}^{-1}$ kann (3-1) in die folgende Form gebracht werden:

$$v^T P_{yy} v = \bar{y}^T P_{yy} \bar{y} - d\hat{p}_1^T \bar{Q}_{11}^{-1} d\hat{p}_1 - \hat{p}_2^T Q_{22}^{-1} \hat{p}_2 \qquad (3\text{-}5)$$

An dieser Beziehung ist abzulesen, daß die quadratische Form der Residuen $v^T P_{yy} y$ genau um den Betrag $\hat{p}_2^T Q_{22}^{-1} \hat{p}_2$ ansteigt, wenn die Parameter $\hat{p}_2$ in einer Neuausgleichung a priori gleich Null gesetzt werden. Der Parametervektor der ersten Gruppe ändert sich von $d\hat{p}_1$ in $d\bar{\hat{p}}_1$ und die zugehörige Kofaktorenmatrix von Q_{11} in $\bar{Q}_{11}$. Eine solche Zerlegung in zwei Parametergruppen läßt sich natürlich auch bei einem von Hause aus linearen Zusammenhang durchführen.

$$v^T P_{yy} v = y^T P_{yy} y - \bar{\hat{p}}_1^T \bar{Q}_{11}^{-1} \bar{\hat{p}}_1 - \hat{p}_2^T Q_{22}^{-1} \hat{p}_2 \ , \qquad (3\text{-}6)$$

wenn sich die Signifikanzuntersuchungen nur auf die Parametergruppe $\hat{p}_2$ beziehen soll. Anderenfalls ist im folgenden $Q_{22} = Q_{\hat{p}\hat{p}}$, $\hat{p}_2 = \hat{p}$ und $u_2 = u$ zu sehen.

4. GLOBALTEST DER ZU UNTERSUCHENDEN PARAMETER

Eine Schätzung für die Varianz der Gewichtseinheit mit $f_1 = n - u$ Freiheitsgraden erhält man aus der Quadratsumme der Verbesserungen,

$${}_1s_0^2 = \frac{v^T P_{yy} v}{f_1} \ , \ f_1 = n - u \ . \qquad (4\text{-}1)$$

Eine zweite, davon unabhängige Schätzung ist unter der Nullhypothese $H_0\{E(\hat{p}_2) = 0\}$, die Parameter sind nicht signifikant, durch

$${}_2s_0^2 = \frac{\hat{p}_2^T Q_{22}^{-1} \hat{p}_2}{u_2} \qquad (4\text{-}2)$$

gegeben. Der Quotient

$$F = \frac{{}_2s_0^2}{{}_1s_0^2} \tag{4-3}$$

gehört dann, wenn die Nullhypothese zutrifft, einer zentralen F-Verteilung mit den Freiheitsgraden $f_2 = u_2$ und $f_1 = n - u$ an:

$$F \sim F_{f_2,f_1} \tag{4-4}$$

Die empirische Testgröße F wird mit dem bei der vorzugebenden Sicherheitswahrscheinlichkeit $S = 1-\alpha$ zu bildenden Fraktilwert der F-Verteilung verglichen

$$F \leftrightarrow F_{\alpha,f_2,f_1} \ . \tag{4-5}$$

Gilt $F \leq F_\alpha$, so wird die Nullhypothese angenommen und die zweite Parametergruppe kann, da sie nicht modellrelevant ist, aus der Ausgleichung gestrichen werden.

5. EINZELTESTS DER PARAMETER NACH EINER CHOLESKY-ZERLEGUNG MIT PIVOTSUCHE

Muß die Nullhypothese wegen $F > F_{\alpha,f_2,f_1}$ abgelehnt werden, sind die für den Ausfall des Tests verantwortlichen signifikanten Parameter ausfindig zu machen. Da die Parameter korreliert sind, können sie nicht unmittelbar, etwa durch Anwendung des t-Tests, einzeln getestet werden. Die quadratische Form $\hat{\boldsymbol{p}}_2^T \boldsymbol{Q}_{22}^{-1} \hat{\boldsymbol{p}}_2$ kann jedoch in eine Quadratsumme von u_2 unkorrelierte und gleichgenaue Summanden zerlegt werden. Dazu denke man sich die Formmatrix $\boldsymbol{Q}_{22}^{-1}$ über eine Cholesky-Zerlegung als Produkt zweier Dreiecksmatrizen ausgedrückt:

$$\boldsymbol{Q}_{22}^{-1} = \boldsymbol{C}^T \boldsymbol{C} \tag{5-1}$$

Die rechte obere Dreiecksmatrix $\boldsymbol{C}$ entsteht durch Reduktion der Matrix $\boldsymbol{Q}_{22}^{-1}$ nach dem Gaußschen Algorithmus und anschließender Division jeder Zeile durch die Wurzel ihres Hauptdiagonalelements. Transformiert man den ur-

sprünglichen Parametervektor $\hat{\boldsymbol{p}}_2$ nach der Vorschrift

$$\boldsymbol{r} = \boldsymbol{C}\hat{\boldsymbol{p}}_2 \ , \tag{5-2}$$

so zerfällt die quadratische Form der Parameter $\hat{\boldsymbol{p}}_2$ in die Quadratsumme von u_2 Summanden r_i^2 :

$$\hat{\boldsymbol{p}}_2^T \boldsymbol{Q}_{22}^{-1} \hat{\boldsymbol{p}}_2 = \hat{\boldsymbol{p}}_2^T \boldsymbol{C}^T \boldsymbol{C} \hat{\boldsymbol{p}}_2 = \boldsymbol{r}^T \boldsymbol{r} = r_1^2 + \ldots r_{u_2}^2 \tag{5-3}$$

Jeder Komponente $(\hat{p}_2)_i$ des Vektors $\hat{\boldsymbol{p}}_2$ wird auf diese Weise ein Anteil r_i^2 am Gesamtwert der quadratischen Form zugeordnet.

Die Komponenten des Vektors $\boldsymbol{r}$ haben die folgenden charakteristischen Eigenschaften:

- Die Komponenten r_i sind unkorreliert und gleichgenau mit $s_{r_i}^2 = {}_1s_0^2$, wie die Anwendung des Fehlerfortpflanzungsgesetzes auf (5-2) beweist:

$$\begin{aligned} \boldsymbol{Q}_{rr} &= \boldsymbol{C}\,\boldsymbol{Q}_{22}\,\boldsymbol{C}^T = \boldsymbol{C}(\boldsymbol{C}^T\boldsymbol{C})^{-1}\boldsymbol{C} = \boldsymbol{C}\,\boldsymbol{C}^{-1}(\boldsymbol{C}^T)^{-1}\boldsymbol{C}^T \\ \boldsymbol{Q}_{rr} &= \boldsymbol{\mathcal{E}} \end{aligned} \tag{5-4}$$

- Die den Parametern $(p_2)_1, \ldots, (p_2)_j$ zugeordneten Komponenten $r_1, \ldots, r_j$ ändern im Gegensatz zu den Parametern in einer Neuausgleichung ihren Wert nicht, wenn die in der Reihenfolge weiter hinten stehenden Parameter $(p_2)_{j+1}, \ldots, (p_2)_{u_2}$ gestrichen werden.

- Die quadratische Form der Verbesserungen steigt bei Streichung der Parameter $(p_2)_{j+1}, \ldots, (p_2)_{u_2}$ genau um den Betrag $\sum_{i=j+1}^{u_2} r_i^2$ an.
  Daraus folgt, die quadratische Form der Verbesserungen nimmt mit wachsendem Index i von Parameter zu Parameter jeweils um den Betrag r_i^2 ab, wenn bei jedem neu hinzukommenden Parameter neu ausgeglichen wird.

- Aus den bisher genannten Eigenschaften ergibt sich unmittelbar, daß die Komponenten des Vektors $\boldsymbol{r}$ nicht invariant gegenüber Veränderungen

der Reihenfolge der Parameter sind.

Bevor die unkorrelierten r_i auf Signifikanz getestet werden können, müssen sie in eine sinnvolle Reihenfolge gebracht werden. Diese wird so festgelegt, daß der Parameter $(\hat{p}_2)_k$, für den sich, wenn er an die letzte Stelle des Parametervektors getauscht wird, der kleinste Anteil $(r^2_{u_2})_k$ ergibt, die letzte Stelle beibehält. Für diesen Parameter ist die Wahrscheinlichkeit, nicht signifikant zu sein, am größten. Die vorletzte Stelle nimmt aus den noch verbleibenden u_2-1 Parametern derjenige ein, dessen Anteil $r^2_{u_2-1}$ minimal wird. Nach dieser Vorschrift wird weiter verfahren, bis schließlich auch die bis zuletzt offenbleibenbe führende Position im Parametervektor mit dem Anteil r_i^2 vergeben ist.

Das führt zu der folgenden Strategie bei der Zerlegung der quadratischen Form $\hat{\boldsymbol{p}}_2^T \boldsymbol{Q}_{22}^{-1} \hat{\boldsymbol{p}}_2$ (Cholesky-Zerlegung mit Pivotsuche):

Bei der Reduktion der Matrix $\boldsymbol{Q}_{22}^{-1}$ nach Cholesky entsteht in der Dreiecksmatrix $\boldsymbol{C}$ als letztes Hauptdiagonalelement $c_{u_2u_2}$ die Wurzel aus dem Reziprokwert des Kofaktors des an letzter Stelle stehenden Parameters,

$$c_{u_2u_2} = q_{u_2u_2}^{-1/2} \quad . \tag{5-5}$$

Wegen $r_{u_2} = c_{u_2u_2} \cdot (\hat{p}_2)_{u_2}$ sind die Anteile $(r_{u_2})_i$, die sich ergeben, wenn jeder Parameter nacheinander an die letzte Stelle gesetzt wird, zu berechnen aus

$$(r^2_{u_2})_i = \frac{(\hat{p}_2)_i^2}{q_{ii}} \quad , \; i = 1, \ldots, u_2 \; . \tag{5-6}$$

Die Kofaktoren q_{ii} sind die Hauptdiagonalelemente der Matrix $\boldsymbol{Q}_{22}$. Der Parameter $(\hat{p}_2)_k$, für den $(r^2_{u_2})_k = \mathrm{Min}\left[(r^2_{u_2})_i\right]$ gilt, behält die letzte Position.

Bevor die vorletzte Position besetzt wird, wird gedanklich eine Neuausgleichung ohne den Parameter $(\hat{p}_2)_k$ durchgeführt. Analog zu der in Gleichung (3-5) durchgeführten Zerlegung ergibt sich die neue

$(u_2\text{-}1)\times(u_2\text{-}1)$ Kofaktorenmatrix $\bar{\mathcal{Q}}_{22}$ für die in dem Vektor $\hat{\boldsymbol{p}}_2$ verbleibenden u_2-1 Parameter zu

$$\bar{\mathcal{Q}}_{22} = \mathcal{Q}_{rr} - \boldsymbol{q}_{rk} \cdot q_{kk}^{-1} \cdot \boldsymbol{q}_{rk}^{T} \quad . \tag{5-7}$$

Die Matrix $\mathcal{Q}_{rr}$ und der Vektor $\boldsymbol{q}_{rk}$ entstehen durch Partitionierung und eventuelles Umsortieren der Matrix $\mathcal{Q}_{22}$

$$\mathcal{Q}_{22} = \begin{pmatrix} \mathcal{Q}_{rr} & \boldsymbol{q}_{rk} \\ \boldsymbol{q}_{rk}^{T} & q_{kk} \end{pmatrix} . \tag{5-8}$$

Führt man eine entsprechende Partitionierung auch für den Vektor $\hat{\boldsymbol{p}}_2$ und die Matrix $\mathcal{Q}_{22}^{-1}$ durch

$$\hat{\boldsymbol{p}}_2 = \begin{pmatrix} \hat{\boldsymbol{p}}_r \\ \hat{p}_k \end{pmatrix} \quad , \quad \bar{\mathcal{Q}}_{22} = \begin{pmatrix} \mathcal{N}_{rr} & \boldsymbol{n}_{rk} \\ \boldsymbol{n}_{rk}^{T} & n_{kk} \end{pmatrix} \quad , \tag{5-9}$$

so erhält man die neue Schätzung des $(u_2\text{-}1)\times 1$ Parametervektors $\hat{\hat{\boldsymbol{p}}}_2$ aus den bisherigen Ergebnissen nach der Formel

$$\hat{\hat{\boldsymbol{p}}}_2 = \hat{\boldsymbol{p}}_r + \bar{\mathcal{Q}}_{22} \cdot \boldsymbol{n}_{rk} \cdot \hat{p}_k \quad . \tag{5-10}$$

Mit den Hauptdiagonalelementen der nach (5-7) zu bildenden Kofaktorenmatrix $\bar{\mathcal{Q}}_{22}$ und dem Parametervektor $\hat{\hat{\boldsymbol{p}}}_2$ werden dann die Anteile

$$\left(r_{u_2\text{-}1}\right)_i = \frac{\left(\hat{\hat{p}}_2\right)_i^2}{\bar{q}_{ii}} \quad , \quad i = 1, \ldots, k\text{-}1, k\text{+}1, \ldots, u_2$$

berechnet, die sich ergeben, wenn die verbleibenden Parameter nacheinander an die vorletzte Stelle gesetzt werden. Der Parameter mit dem kleinsten Anteil behält diese Position.

Ausgehend von $\bar{\mathcal{Q}}_{22}$, $\hat{\hat{\boldsymbol{p}}}_2$ und $\mathcal{N}_{rr} = \bar{\mathcal{Q}}_{22}^{-1}$ wird dann wie zuvor beschrieben die Reduktion des Parametervektors weitergeführt und die drittletzte Position besetzt.

Dieses Verfahren wird fortgesetzt bis alle Parameter abgearbeitet sind und die gesuchten Anteile r_i^2 $(i = 1, \ldots, u_2)$ als ihrer Bedeutung nach sortierte Folge vorliegen. Wie sich aus den genannten Eigenschaften der Variablen r_i ergibt, ist die Folge nicht unbedingt auch der Größe nach geordnet.

Da die Komponenten des Vektors $\boldsymbol{r}$ unkorreliert sind, können sie einzeln auf Signifikanz getestet werden. Die Nullhypothese lautet für jede Komponente $H_0\{E(r_i) = 0\}$. Da es sich um einen multiplen Test handelt und im Prinzip jede Komponente r_i getestet wird, ist jeder Einzeltest mit der Sicherheitswahrscheinlichkeit

$$S = 1 - \alpha = (1 - \bar{\alpha})^{1/u_2} \tag{5-11}$$

durchzuführen (***KRÜGER*** 1980), wenn das Gesamtrisiko für einen Fehler erster Art von insgesamt $\bar{\alpha}$ nicht überschritten werden soll.

Mit dem letzten Anteil $r_{u_2}^2$ beginnend werden die Testgrößen

$$t_{u_2} = \frac{|r_{u_2}|}{{}_1s_0} \quad \text{oder} \quad F_{u_2} = \frac{r_{u_2}^2}{{}_1s_0^2} = t_{u_2}^2 \tag{5-12}$$

gebildet, die unter der Nullhypothese einer t-Verteilung mit $f_1 = n - u$ Freiheitsgraden bzw. einer F-Verteilung mit dem Freiheitsgrad eins im Zähler und f_1 im Nenner angehören. Die Signifikanzschwellen $t_{\alpha/2,f_1}$ bzw. $F_{\alpha,1,f_1}$ werden der t- oder F-Verteilung entnommen.
Wegen $t_{\alpha/2,f_1}^2 = F_{\alpha,1,f_1}$ sind beide Testmöglichkeiten gleichwertig und führen zu identischen Ergebnissen.

Wird die Nullhypothese für r_{u_2} angenommen, kann der zugeordnete Parameter aus der Ausgleichung eliminiert werden. Dadurch erhöht sich der Freiheitsgrad in der Ausgleichung um eins und die Quadratsumme der Residuen nimmt um den Betrag $r_{u_2}^2$ zu, so daß man als neue Schätzung für die Varianz der Gewichtseinheit

$$\left(s_0^2\right)_{u_2-1} = \frac{\boldsymbol{v}^T\boldsymbol{P}\boldsymbol{v} + r_{u_2}^2}{f_1 + 1} \tag{5-13}$$

erhält. Die Testgrößen für den vorletzten Anteil $r^2_{u_2-1}$ lauten dann unter Berücksichtigung der neuen Varianzschätzung

$$t_{u_2-1} = \frac{|r_{u_2-1}|}{(s_0)_{u_2-1}} \quad \text{oder} \quad F_{u_2-1} = \frac{r^2_{u_2-1}}{(s^2_0)_{u_2-1}} \ , \tag{5-14}$$

die mit den Fraktilwerten $t_{\alpha/2,f_1+1}$ bzw. $F_{\alpha,1,f_1+1}$ verglichen werden. Wird auch in diesem Schritt die Signifikanzschwelle nicht überschritten, erhöht sich die Quadratsumme der Residuen um den zusätzlichen Betrag $r^2_{u_2-1}$ bei einem Freiheitsgrad von f_1+2.

Als allgemeines Bildungsgesetz für die Varianzschätzung der Gewichtseinheit erhält man auf diese Weise für den zweiten und alle folgenden Testschritte

$$(s^2_0)_i = \frac{\boldsymbol{v}^T\boldsymbol{P}\boldsymbol{v} + r^2_{u_2} + r^2_{u_2-1} + \ldots + r^2_{i+1}}{f_1 + u_2 - i} \ . \tag{5-15}$$

Der Wert $(s^2_0)_i$ ist die a posteriori Varianzschätzung einer Ausgleichung mit u_1+i Parametern. Der Index i wird beginnend mit $i=u_2-1$ in jedem Testschritt um eins erniedrigt. Die Testgrößen lauten in allgemeiner Form

$$t_i = \frac{|r_i|}{(s_o)_i} \quad \text{oder} \quad F_i = \frac{r^2_i}{(s^2_0)_i} \ , \tag{5-16}$$

die mit den jeweils neu zu bildenden Fraktilwerten $t_{\alpha/2,f_1+u_2-i}$ bzw. $F_{\alpha,1,f_1+u_2-i}$ zu vergleichen sind, bis die Signifikanzschwelle an der Stelle i = j überschritten wird. Da die Folge sortiert ist, muß für alle weiter vorn stehenden Parameter ebenfalls angenommen werden, daß sie sich signifikant von Null unterscheiden, auch wenn ihre Testgrößen im einzelnen unterhalb der Signifikanzschwelle bleiben sollten.

Diese Teststrategie findet man im Zusammenhang mit Orthogonalpolynomen bei ***ANDERSON*** (1962).

Mit $r_1, \ldots, r_j$ ist die Kombination der j Parameter gefunden, die modellrelevant sind:

$r_1^2, \ldots, r_j^2$	$r_{j+1}^2, \ldots, r_{u_2}^2$
$F_j > F_{\alpha,1,f_1+u_2-j}$	$F_{j+1} \le F_{\alpha,1,f_1+u_2-j-1}, \ldots, F_{u_2} \le F_{\alpha,1,f_1}$
signifikant	nicht signifikant

(5-17)

Die in ihrem Wert sich ändernden noch verbleibenden $u_1 + j$ Parameter können in einer abschließenden Ausgleichung geschätzt werden.

Falls in die Zerlegung alle Parameter einbezogen werden ($u_2=u$), entsteht der durch die Signifikanzuntersuchungen verkürzte neue Parametervektor $\hat{\boldsymbol{p}}$ und dessen Kofaktorenmatrix ohne Neuausgleichung im Zuge der Berechnung der Anteile r_i^2.

Zwischen den quadratischen Formen des neuen Verbesserungsvektors $\bar{\boldsymbol{v}}$ und des ursprünglichen Verbesserungsvektors $\boldsymbol{v}$ besteht die Relation

$$\bar{\boldsymbol{v}}^T \boldsymbol{P}_{yy} \bar{\boldsymbol{v}} = \boldsymbol{v}^T \boldsymbol{P}_{yy} \boldsymbol{v} + \sum_{i=j+1}^{u_2} r_i^2 \,, \tag{5-18}$$

und der Freiheitsgrad in der Ausgleichung erhöht sich auf

$$\bar{f}_1 = f_1 + u_2 - j \,. \tag{5-19}$$

Das beschriebene Verfahren, Cholesky-Zerlegung mit Pivotsuche, kann als eine Verallgemeinerung der Orthogonalpolynome bei ausgleichenden Polynomen (***WOLF*** 1975) angesehen werden. Es läßt sich jedoch auch auf mehrdimensionale Probleme und auf beliebige Funktionstypen anwenden. Außerdem werden die Parameter ihrer Bedeutung nach umsortiert, während man bei Orthogonalpolynomen voraussetzt, daß mit wachsenden Potenzen die Wahrscheinlichkeit der Signifikanz abnimmt.

6. ANMERKUNGEN

Aus der Statistik-Literatur sind weitere Strategien bekannt, um aus einer gegebenen Zahl von Einflußgrößen die optimale Kombination zu finden.

Am rechenaufwendigsten ist die Lösung, alle möglichen Kombinationen (bei u Parametern 2^u Kombinationen) durchzurechnen und anhand eines vorgegebenen Kriteriums, z.B. Minimierung von s_0^2, das Optimum auszuwählen (all possible subsets regression). Daneben gibt es die schrittweise Regression, wo in jedem Schritt jeweils ein weiterer Parameter in die Ausgleichung einbezogen wird und dann geprüft wird, welche der schon im Modell befindlichen Parameter wieder eliminiert werden können. Für die Aufnahme und die Streichung eines Parameters werden bestimmte an der F-Verteilung orientierte Schwellenwerte festgesetzt (stepwise regression). Eine vereinfachte Version der schrittweisen Regression erhält man, wenn auf die Elimination von einmal aufgenommenen Parametern verzichtet wird (forward selection). Diese Vorgehensweise ist in ***CASPARY/SCHWINTZER*** (1981a) beschrieben. Es hat sich gezeigt, daß bei dieser Version die endgültige Lösung noch Parameter enthalten kann, die, ohne daß sich die a posteriori Varianzschätzung signifikant ändert, eliminiert werden können. Die umgekehrte Vorgehensweise, die von einer Ausgleichung mit allen in Frage kommenden Einflußgrößen ausgeht, ist Gegenstand dieses Beitrages (backward elimination).

BECK (1978) hat gezeigt, daß es keine objektiv beste Lösung gibt und daß die Unterschiede zwischen den Methoden vom praktischen Standpunkt betrachtet unbedeutend sind. Lediglich in extremen Fällen können größere Differenzen auftreten. In manchen Fällen ist auch kein eindeutiges Ergebnis gegeben, zum Beispiel wenn einzelne Parameter wegen ihrer hohen Korrelation gegenseitig ersetzbar sind. Es können dann neben dem gefundenen Ergebnis andere von der Ausgleichung her gesehen gleichwertige Kombinationen existieren, die der Problemstellung möglicherweise gerechter werden.

Allen Methoden ist gemeinsam, daß sie pragmatisch gehandhabt werden müssen und das Ergebnis kritisch auf den Einzelfall bezogen zu interpretieren ist.

7. BEISPIEL

Mit einem inertialen Vermessungssystem (PADS IMU 007) sind für Genauigkeitsuntersuchungen auf einer Traverse in Maryland (USA) die Koordinaten von elf lage- und höhenmäßig bekannten Punkten gemessen worden. Die Abweichungen der gemessenen Koordinaten von den Sollkoordinaten lassen systematische Einflüsse vermuten. Über die Untersuchungsergebnisse berichtet

TINDALL (1982).

Es werden hier nur die einem Originalprotokoll entnommenen Höhenbeobachtungen einer herausgegriffenen Traversenmessung betrachtet. Am Anfangspunkt P_A wurde das Instrument auf die Sollkoordinaten abgeglichen, so daß insgesamt zehn Höhendifferenzen für die weitere Untersuchung zur Verfügung stehen. In der folgenden Tabelle sind die Meßergebnisse und einige denkbare Einflußgrößen zusammengestellt. Es bedeuten:

$dH_i = (H_{Ist} - H_{Soll})_i$

s_i – Weglänge vom Anfangspunkt bis zum Punkt P_i

ΔH_i – positive Höhenänderung von Punkt P_{i-1} zum Punkt P_i

$\widetilde{\Delta H_i}$ – negative Höhenänderung von Punkt P_{i-1} zum Punkt P_i

t_i – Zeitverbrauch vom Anfangspunkt bis zur Messung am Punkt P_i

Pkt.	dH_i [m]	s_i [km]	ΔH_i [km]	$\widetilde{\Delta H_i}$ [km]	t_i [sek·10^{-4}]
1	+ 0,1	2.620	+ 0,0157		0,1820
2	+ 0,7	7,595		- 0,0088	0,2456
3	+ 1,3	10,260	+ 0,0189		0,2837
4	+ 1,4	13,542	+ 0,0257		0,3580
5	+ 1,7	15,521	+ 0,0183		0,4196
6	+ 1,9	20,801	+ 0,0534		0,5492
7	+ 2,6	24,548	+ 0,0080		0,7009
8	+ 2,1	28,821		- 0,0434	0,8468
9	+ 1,7	33,459		- 0,0322	0,9856
10	+ 2,4	40,580		- 0,0248	1,1244

Tabelle 1: Beobachtungen und Einflußgrößen

Es wird die folgende lineare ausgleichende Funktion gewählt:

$$dH_i + v_i = \hat{a}\cdot s_i + \hat{b}\cdot \sum_{j=1}^{i} \Delta H_j + \hat{c}\cdot \sum_{j=1}^{i} \widetilde{\Delta H_j} + \hat{d}\cdot t_i + \hat{e}\cdot t_i^2 + \hat{f}\cdot t_i^3 \qquad (7\text{-}1)$$

Alle Messungen werden als gleichgenau und unkorreliert behandelt (P = E). Die ersten drei Glieder auf der rechten Seite sind aus ***TINDALL*** (1982) ent-

nommen. Zusätzlich wurde die Gleichung um die zeitabhängigen Glieder erweitert. In einer Ausgleichung nach der Methode der kleinsten Quadrate wird eine Schätzung für den Parametervektor

$$\boldsymbol{p}^T = (a, b, c, d, e, f) \qquad (7\text{-}2)$$

gesucht.

Auf die Begründung des Ansatzes und auf die Bedeutung der Parameter braucht hier nicht eingegangen zu werden. Die Einflußgrößen und Beobachtungen sind mit den in der Tabelle 1 angegebenen Dimensionen in die Ausgleichung eingeführt worden.

Im folgenden werden die Ergebnisse der Ausgleichung, der Cholesky-Zerlegung mit Pivotsuche und der Signifikanztests tabellarisch zusammengestellt:

- Ausgleichung gemäß Gleichung (7-1)

$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{d}$	$\hat{e}$	$\hat{f}$
+ 0,237	- 15,1	+ 35,1	- 2,2	+ 2,6	- 1,5
± 0,042	± 6,1	± 8,4	± 1,4	± 2,9	± 1,9

Tabelle 2: Vorläufiges Ausgleichungsergebnis

$$\boldsymbol{v}^T\boldsymbol{v} = \boldsymbol{dH}^T\boldsymbol{dH} - \hat{\boldsymbol{p}}^T\boldsymbol{Q}_{\hat{p}\hat{p}}^{-1}\hat{\boldsymbol{p}}$$

$$0{,}084 = 30{,}470 - 30{,}386$$

$$s_0^2 = 0{,}084/f = 0{,}021 \quad ; \quad f = 10 - 6 = 4$$

$$s_0 = \pm 0{,}14\ \text{m}$$

- Cholesky-Zerlegung mit Pivotsuche

Zerlegung von $\hat{\boldsymbol{p}}\,\boldsymbol{Q}_{\hat{p}\hat{p}}^{-1}\hat{\boldsymbol{p}} = 30{,}386$

Berechnungs-folge:	6.	5.	4.	3.	2.	1.	
Parameter	r_1^2	r_2^2	r_3^2	r_4^2	r_5^2	r_6^2	Parameter
a	27,980	13,774	1,404	0,759	0,762	0,654	a
b			[0,188]	0,176	0,158	0,130	b
c		[2,154]	1,212	1,384	0,360	0,370	c
d				[0,043]	0,045	0,055	d
e					[0,009]	0,018	e
f						[0,012]	f

Tabelle 3: Berechnung der Anteile r_i^2

Mit dem Ergebnis gilt die folgende Zerlegung:

$$\hat{\boldsymbol{p}}\, Q_{\hat{p}\hat{p}}^{-1}\, \hat{\boldsymbol{p}} = (r_1)_a^2 + (r_2)_c^2 + (r_3)_b^2 + (r_4)_d^2 + (r_5)_e^2 + (r_6)_f^2$$

$$30{,}386 = 27{,}980 + 2{,}154 + 0{,}188 + 0{,}043 + 0{,}009 + 0{,}012$$

- Signifikanztests

Testgröße: $F_i = \dfrac{r_i^2}{(s_0^2)_i}$

Gesamtsicherheitswahrscheinlichkeit: $S = 1-\overline{\alpha} = 94\,\%$

Sicherheitswahrscheinlichkeit für Einzeltests: $S = (1-\alpha)^{1/6} = 1-\alpha = 99\,\%$

Parameter	a	c	b	d	e	f
i	1	2	3	4	5	6
r_i^2	27,980	2,154	0,188	0,043	0,009	0,012
$(s_0^2)_i$	0,277	0,042	0,021	0,018	0,019	0,021
F_i	101,0	51,3	9,0	2,4	0,5	0,6
		>	<	<	<	<
$F_{1\%,1,10-i}$		11,3	12,2	13,7	16,3	21,2

Tabelle 4: Ergebnis der Signifikanzuntersuchungen

signifikante Parametergruppe: a, c

nicht signifikante Parametergruppe: b, d, e, f

endgültiges Ausgleichungsmodell mit dem neu zu schätzenden Parametervektor $\hat{\boldsymbol{p}}^T = (\hat{\bar{a}}, \hat{\bar{c}})$:

$$dH_i + v_i = \hat{\bar{a}} \cdot s_i + \hat{\bar{c}} \cdot \sum_{j=1}^{i} \widetilde{\Delta H}_j \qquad (7\text{-}3)$$

- Ausgleichung gemäß Gleichung (7-3)

$\hat{\bar{a}}$	$\hat{\bar{c}}$
0,112	+ 21,4
± 0,006	± 3,0

Tabelle 5: endgültiges Ausgleichungsergebnis

$$\bar{\boldsymbol{v}}^T \bar{\boldsymbol{v}} = \boldsymbol{v}^T \boldsymbol{v} + r_3^2 + r_4^2 + r_5^2 + r_6^2$$

$$0{,}336 = 0{,}084 + 0{,}188 + 0{,}043 + 0{,}009 + 0{,}012$$

oder

$$\bar{\boldsymbol{v}}^T \bar{\boldsymbol{v}} = d\mathcal{H}^T d\mathcal{H} - r_1^2 - r_2^2$$

$$0{,}336 = 30{,}470 - 27{,}980 - 2{,}154$$

$$\bar{s}_0^2 = 0{,}336/\bar{f} = 0{,}042 \quad ; \quad \bar{f} = 10 - 2 = 8$$

$$\bar{s}_0 = \pm 0{,}20\,\text{m}$$

Man beachte die erheblichen Ergebnisänderungen für die Schätzungen der Parameter a und c in den beiden Modellen (7-1) und (7-3). Das ist eine Folge der zum Teil großen Korrelationen zwischen den ursprünglich sechs Parametern (z.B. Korrelationskoeffizient $r_{\hat{a},\hat{d}} = -0{,}86$). Die mittleren Fehler der verbleibenden zwei Parameter sind gegenüber der ersten vorläufigen Schätzung

kleiner geworden.

LITERATUR

ANDERSON, T.W.: ***The choice of the degree of a polynominal regression as a multiple decision problem***. Ann. Math. Stat. 33, 255-265, 1962

BERK, K.N.: ***Comparing Subset Regression Procedures***. Technometrics 20, 1-6, 1978

BAUER, W.: ***Verkehrsermittlung bebauter Grundstücke mit Hilfe automatisch erstellter Vergleichswerttabellen***. Allgemeine Vermessungsnachrichten 85, 138-149, 1978

CASPARY, W., und P. SCHWINTZER: ***Bestimmung von Einzelpunktbewegungen und von Relativbewegungen zweier Netzteile in geodätischen Deformationsnetzen***. Zeitschrift für Vermessungswesen 106, 277-288, 1981 a

CASPARY, W., und P. SCHWINTZER: ***An extension of Chronometric Gyroscope Observation Methods***. The Canadian surveyor 35, 364-372, 1981 b

FLOTRON, A.: ***Interpretation von Deformationsmessungen mittels ausgleichender Funktionen***. In Conzett/Mathias/Schmidt (Hrsg.), Ingenieurvermessung 80, Beiträge zum VIII. Internationalen Kurs für Ingenieurvermessung, Bd. 2, Dümmler, Bonn, E 6/1 - E 6/14, 1981

JAKOB, B.: ***Anwendung der Theorie der ausgleichenden Funktion zur Bestimmung der plausibelsten Temperaturfunktion***. Vermessungstechnik 27, 152-156, 1979

KRÜGER, J.: ***Elimination grober Beobachtungsfehler in geodätischen Netzen***. In Pelzer (Hrsg.), Geodätische Netze in Landes- und Ingenieurvermessung, Wittwer, Stuttgart, 499-509, 1980

SCHWARZ, W.: ***Die Additionskorrektion bei elektrooptischen Distanzmessern***. Allgemeine Vermessungsnachrichten 90, 54-63, 1983

TINDALL, T.O.: ***Evaluation of the Position and Azimuth Determining System's Potential for Higher Accuracy Survey***. Proceedings of the ACSM-Spring Meeting 1982

WOLF, H.: ***Ausgleichungsrechnung – Formeln zur praktischen Anwendung***. Dümmler, Bonn 1975

DAS PROGRAMMSYSTEM NOPTI II ZUM ENTWURF DES OPTIMALEN BEOBACHTUNGSPLANES GEODÄTISCHER LAGENETZE NACH INVARIANTEN KRITERIEN

von R. BAUMER u. H. HEISTER

SUMMARY

The optimization program NOPTI II was developed for the second order design of geodetic networks. The program allows to minimize different invariant objective functions, the choice of which controls the accuracy of the network. In several examples it is demonstrated how this optimization technique effects, via the chosen objective function, the accuracy and the distribution of the observations.

ZUSAMMENFASSUNG

Zum optimalen Entwurf des Beobachtungsplanes geodätischer Lagenetze wurde das Programmsystem NOPTI II erstellt. Es ermöglicht unter verschiedenen invarianten Zielfunktionen skalare Genauigkeitsmaße des Netzes zu optimieren. An unterschiedlichen Beispielen wird sowohl die Arbeitsweise der Programme als auch die Auswirkung der unterschiedlichen Zielfunktionen auf das Optimierungsergebnis erläutert.

1. EINFÜHRUNG

Für den Entwurf geodätischer Netze ist neben der Festlegung der Konfiguration auch die Bestimmung des Beobachtungsplanes mit Analysen über die zu erwartende Lagegenauigkeit der Netzpunkte von großer Bedeutung. Um hier zu computerfreundlichen Lösungsalgorithmen zu kommen, ist es notwendig, möglichst objektive und invariante Zielfunktionen und Restriktionen zu formulieren, damit das optimale Modell den tatsächlichen Verhältnissen möglichst nahe kommt. Hierfür wurden in den letzten Jahren eine Anzahl von Verfahren vorgestellt, die in ***GRAFAREND*** e.a. (1979) und ***SCHMIDT*** (1982) nahezu vollständig erfaßt wurden.

Ist die Konfiguration eines Netzes doch weitgehend abhängig von der Topographie, der Problemstellung des Vermessungsobjektes o.ä., so läßt der Beobachtungsplan durch Bewertung der Beobachtungen, Auswahl des Instrumentariums oder Festlegung des Aufwandes noch Möglichkeiten zur individuellen Verbesserung der Netzgenauigkeit offen. Die Komplexität dieser Entscheidungskette und die oft fachfremden Lösungstheorien haben die praktische Anwendung doch wesentlich eingeschränkt.

Ziel dieser Beitrags ist, dem Praktiker ein anwendungsfreundliches Programmsystem vorzustellen, das sich besonders durch einfache Bedienung - interaktive Eingabe und Führung durch das Programm - auszeichnet und eine intensive Beschäftigung mit Genauigkeitskriterien und Optimierungsalgorithmen erübrigt. Diese Übersichtlichkeit bedingt natürlich eine gewisse Festlegung à priori von Optimalitätskriterien; durch gewisse Steuerparameter kann jedoch jederzeit eine Ausweitung in der Wahl bestimmter Zielfunktionen und Beobachtungsverfahren vorgenommen werden. Die hier verwendeten Zielfunktionen zur Charakterisierung der Netzgenauigkeit sind skalare Genauigkeitsmaße; sie beziehen sich auch auf solche Funktionen der Unbekannten, die datumsinvariant und somit unverzerrt schätzbar sind. Sie decken die meisten Anwendungsfälle ab und sind auch sinnvoll, damit z.B. eine Änderung in der Annahme von Bezugspunkt und Orientierung des Netzes keinen Einfluß auf die optimale Beobachtungsanordnung hat.

Nach Vorgabe der Näherungskoordinaten der Netzpunkte, ihrer Sichtverbindungen, der à priori Genauigkeiten aller Messungsgrößen - festgelegt durch das

einsetzbare Instrumentarium - , der Beobachtungsmethode sowie des vorläufig festgelegten Meßaufwandes wird mit Hilfe der dynamischen Optimierung nach Wahl eines skalaren Genauigkeitskriteriums der optimale Beobachtungsplan erstellt. Hierbei können auch gewisse Zuverlässigkeitskriterien berücksichtigt werden; dies kann z.B. in der Weise geschehen, daß Beobachtungen, die für die Verbesserung der Netzgenauigkeit keinen wesentlichen Beitrag liefern, jedoch für die Kontrolle und Aufdeckung grober Fehler wichtig sind, im Optimierungsalgorithmus nicht eliminiert werden.

Da das Optimierungsverfahren bereits in ***HEISTER*** (1978) eingehend beschrieben wurde, wird hier nur der prinzipielle Lösungsweg, soweit er zum Verständnis der Programmbeschreibung notwendig ist, skizziert.

2. LÖSUNGSWEG UND ZIELFUNKTIONEN

Zur endgültigen Festlegung eines Lagenetzes von u Netzunbekannten x_i, die durch den Vektor $\boldsymbol{u}$ dargestellt werden, sollen n Beobachtungsgrößen l_i bestimmt werden, die in dem Vektor $\boldsymbol{\ell}$ zusammengefaßt werden. Im allgemeinen liegt zur Berechnung des Unbekanntenvektors $\boldsymbol{u}$ eine Überbestimmung vor, so daß sich hierfür nach den bekannten Regeln der Methode der kleinsten Quadrate die Lösung ergibt

$$\boldsymbol{u} = \left(\boldsymbol{\mathcal{A}}^{\mathsf{T}}\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}\right)^{-1}\boldsymbol{\mathcal{A}}^{\mathsf{T}}\boldsymbol{\mathcal{P}}\boldsymbol{\ell}$$
$$\text{oder} \quad \boldsymbol{u} = \left(\boldsymbol{\mathcal{A}}^{\mathsf{T}}\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{A}}\right)^{+}\boldsymbol{\mathcal{A}}^{\mathsf{T}}\boldsymbol{\mathcal{P}}\boldsymbol{\ell} \quad . \tag{1}$$

Darin bedeutet $\boldsymbol{\mathcal{A}}$ die Design-Matrix und

$$\boldsymbol{\mathcal{P}} = \mathrm{diag}\left(p_1, p_2, \ldots, p_n\right) \tag{2}$$

die Gewichtsmatrix der Beobachtungen l_i.

$\boldsymbol{\mathcal{A}}$ ist durch die Vorgabe der Näherungskoordinaten und möglichen Beobachtungen festgelegt, $\boldsymbol{\mathcal{P}}$ durch die à priori Genauigkeiten der Messungen. Werden zur Lagebestimmung Richtungen oder Winkel sowie Strecken zugelassen, dann sind die Gewichte je nach Beobachtungsart (Richtungen, Winkel, Strecken) festgelegt durch

$$P_R = \frac{const}{m_R{}^2} \quad , \quad P_W = \frac{const}{m_W{}^2} \quad , \quad P_S = \frac{const}{m_S{}^2} \quad . \tag{3}$$

Dabei soll für die einzelnen mittleren Fehler à priori folgender Ansatz gelten:

$$\left.\begin{aligned} m_R^2 &= \frac{m_{or}^2}{w_r} + \sigma_r^2 \quad , \\ m_W^2 &= 2\, m_R^2 \quad , \\ m_S^2 &= \frac{m_{os}}{w_s} + \sigma_s^2 \quad , \quad \sigma_s^2 = b^2 \cdot e \quad . \end{aligned}\right\} \tag{4}$$

Hierin bedeutet m_{or} der zufällige Fehleranteil einer Zielung (Richtungsmessung und σ_r ein systematischer Anteil, der gleichbleibend und unabhängig von der Wiederholungszahl w_r ist. Für die Varianz der Streckenmessung ist ein entsprechender Ansatz gewählt worden, wobei σ_s ein von der Entfernung e abhängiger systematischer Anteil ist. Somit ergibt sich für die Gewichtsfestlegung folgende allgemeine Beziehung

$$P_i = \frac{const \cdot w_i}{m_o^2 + w_i \sigma^2} \quad . \tag{5}$$

Kann die Genauigkeit der Richtungs- oder Winkelmessung durch Erhöhung der Wiederholungszahlen - natürlich nur in vertretbaren Grenzen - gesteigert werden, so ist dieses Fehlermodell besonders beim Einsatz von elektrooptischen Distanzmeßgeräten im allgemeinen einzuschränken. Es wird hier vorgeschlagen, die für das verwendete EDM-Gerät bewährte Meßmethode bei allen vorgesehenen Streckenmessungen gleichbleibend anzuwenden, sodaß im Normalfall für w_s nur eins oder null wählbar ist.

Der Gesamtaufwand S aller Beobachtungen ergibt sich dann in Abhängigkeit der Wiederholungszahlen zu

$$S = A_R + A_S$$

wobei

$$\left.\begin{aligned} A_R &= a_r \sum_{i=1}^{r} w_i \ , \\ A_S &= a_s \sum_{i=1}^{s} w_i \ , \\ r + s &= n \ . \end{aligned}\right\} \qquad (6)$$

Hierin bedeuten a_r und a_s Aufwandsfaktoren für die Richtungs- und Streckenmessung. Im Normalfall kann

$$a_r = a_s = 1$$

gesetzt werden, so daß

$$S = \sum_{i=1}^{n} w_i \ . \qquad (7)$$

Oder der Aufwand der Streckenmessung wird in Einheiten der Richtungsmessung festgelegt, so daß

$$S = \sum_{i=1}^{r} w_i + k \sum_{i=1}^{s} w_i \qquad (8)$$

$$\text{mit } k = \frac{a_s}{a_r} \ .$$

Unter Einhaltung dieses Gesamtaufwandes S sollen nun gewisse Genauigkeitsmaße optimiert werden.

Ausgehend von der Matrix der Kofaktoren

$$Q_{xx} = \left(\mathcal{A}^T \mathcal{P} \mathcal{A}\right)^{-1}$$

oder (9)

$$Q_{xx} = \left(\mathcal{A}^T \mathcal{P} \mathcal{A}\right)^{+}$$

läßt sich mit Hilfe der Funktionsmatrix $\mathcal{F}$ die Matrix der Funktionsgewichte

berechnen:

$$\mathcal{Q}_{FF} = \mathcal{F}\,\mathcal{Q}_{xx}\,\mathcal{F}^{T} \quad . \tag{10}$$

Als skalares Genauigkeitsmaß für ein geodätisches Netz wird in allgemeinster Form die $\mathrm{sp}\,\mathcal{Q}_{FF}$ eingesetzt.

Damit läßt sich das Optimierungsproblem folgendermaßen formulieren

$$\text{Min. } Z = \mathrm{sp}\,\mathcal{Q}_{FF} \tag{12a}$$

unter der Bedingung, daß

$$A_R + A_S \leq S \quad . \tag{12b}$$

Die Wiederholungszahlen w_i für die Beobachtungen, die die Gewichte p_i festlegen und sowohl die Zielfunktion (12a) als auch die Restriktion (12b) verändern, werden jetzt im Optimierungsalgorithmus so bestimmt, daß das angestrebte Minimum des skalaren Genauigkeitsmaßes erreicht wird.

Für die Praxis sind nun solche Genauigkeitsmaße von Interesse, die das Netz in seiner Gesamtheit möglichst scharf und umfassend beschreiben. Dazu eignen sich besonders die skalaren Größen, die datumsunabhängig sind. Es lassen sich somit aus (10) für die Zielfunktion (12a) folgende Spezialfälle ableiten:

1. $\mathcal{F} = \mathcal{E} \qquad \mathcal{Q}_{FF} = \mathcal{Q}_{xx}$

$$\mathrm{sp}\,\mathcal{Q}_{FF} = \mathrm{sp}\,\mathcal{Q}_{xx} \tag{13}$$

2. $\mathcal{F} = \mathbf{A}_R$

$\mathbf{A}_R$ ist die Submatrix in der Design-Matrix $\mathbf{A}$ für Richtungen

$$\mathcal{Q}_{FF} = \mathbf{A}_R\,\mathcal{Q}_{xx}\,\mathcal{A}_R^{T} = \mathcal{Q}_{RR}$$

$$\mathrm{sp}\,\mathcal{Q}_{FF} = \mathrm{sp}\,\mathcal{Q}_{RR} \tag{14}$$

3. $\mathcal{F} = A_W$

A_W ist die Submatrix für Winkel

$$Q_{FF} = A_W Q_{xx} \mathcal{A}_W^T = Q_{WW}$$

$$sp\, Q_{FF} = sp\, Q_{WW} \qquad (15)$$

4. $\mathcal{F} = A_S$

A_S ist die Submatrix für Strecken

$$Q_{FF} = A_S Q_{xx} \mathcal{A}_S^T = Q_{SS}$$

$$sp\, Q_{FF} = sp\, Q_{SS} \qquad (16)$$

5. $\mathcal{F} = \mathcal{A}$

$$Q_{FF} = \mathcal{A} Q_{xx} \mathcal{A}^T = Q_{RS}$$

$$sp\, Q_{FF} = sp\, Q_{RS} \qquad (17)$$

6. $\mathcal{F} = \mathcal{S}$ ($\mathcal{S}$ ist die Matrix der Koeffizienten von Sonderfunktionen)

$$Q_{FF} = \mathcal{S} Q_{xx} \mathcal{S}^T = Q_S$$

$$sp\, Q_{FF} = sp\, Q_S \qquad (18)$$

Berücksichtigt man in (12) unter Vorgabe des Aufwandes S diese skalaren Genauigkeitsmaße, kann durch geeignete Manipulation der Wiederholungszahlen - dies geschieht mit Hilfe der dynamischen Optimierung - das Minimum bestimmt werden.

3. DAS PROGRAMMSYSTEM NOPTI II

Auf der Grundlage des dargestellten mathematischen Konzepts wurde das dialogfähige Programmsystem NOPTI II so konzipiert, daß es auch für den in der mathematischen Optimierung weniger vorgebildeten Nutzer einfach und praxisorientiert zu handhaben ist. Der gesamte Arbeitsablauf kann der Abbildung 1 entnommen werden.

Die Eingabedatei enthält im wesentlichen die Erkundungsergebnisse zur Festlegung der Netzkonfiguration sowie Genauigkeitsmaße für die Beobachtungen, die sich aus dem einsetzbaren Instrumentarium ergeben. Die maximalen

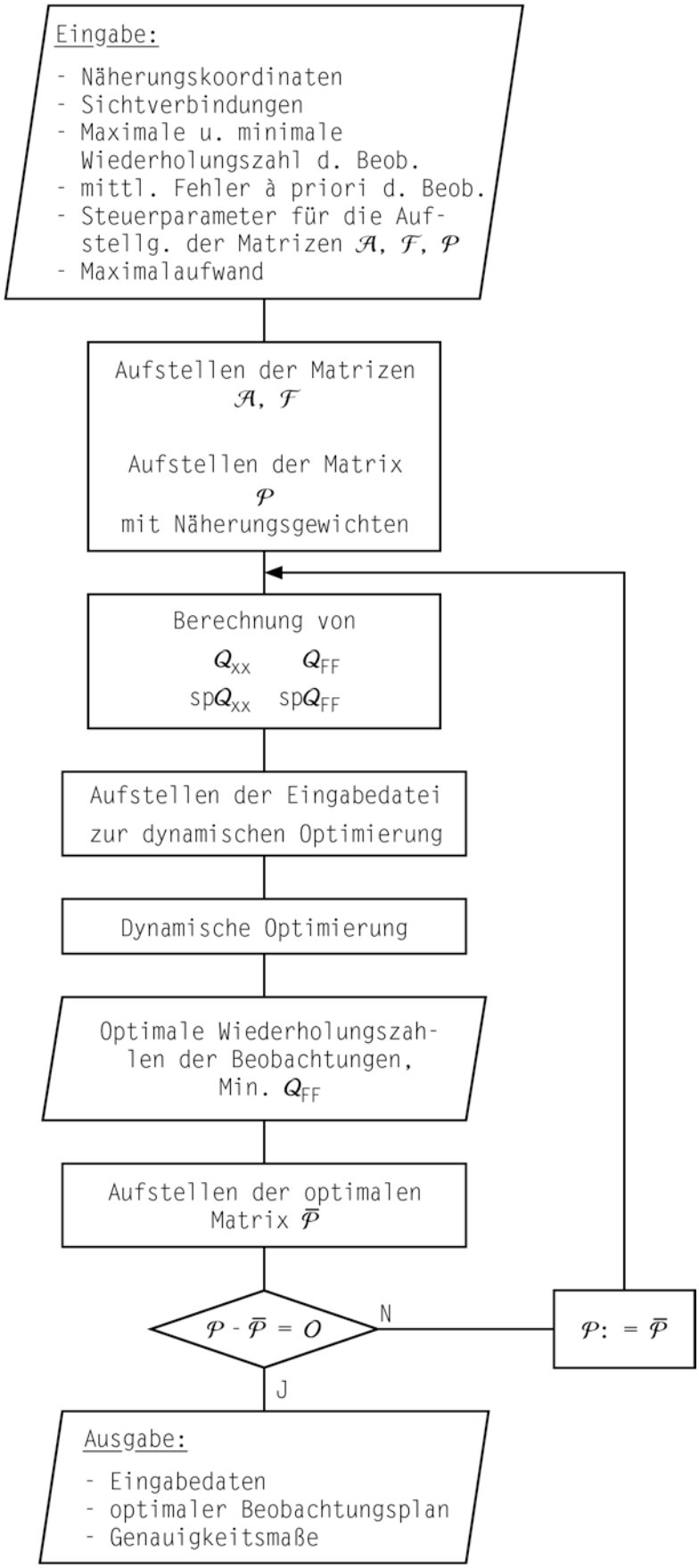


Abb. 1: Ablaufplan des Programms NOPTI II

und minimalen Wiederholungszahlen sowie der Gesamtaufwand sind nach Erfahrungswerten und unter Berücksichtigung der Zuverlässigkeit festzulegen. Über Steuerparameter kann dann bestimmt werden, welches Optimierungsziel angestrebt wird, d.h. welches der skalaren Genauigkeitsmaße (13) bis (17) gewählt wird. Zusätzlich ist dann noch die Möglichkeit gegeben, andere Funktionen gemäß (18) der Unbekannten zu berücksichtigen; die Koeffizienten von $\mathcal{S}$ können dann nach einem vorgegebenen Schema eingegeben werden.

Die Optimierung selbst läuft iterativ ab, da zunächst die Gewichtsmatrix $\mathcal{P}$ durch die nicht optimierten Wiederholungszahlen w_i nur genähert berechnet werden kann; normalerweise führt man zuerst die Wiederholungszahlen ein, die einer Gleichverteilung entsprechen. Danach werden dann im wesentlichen drei Steuerprogramme aktiviert. Das erste erstellt die für den jeweiligen Iterationsschritt aktuellen Dateien für die Optimierung. Das zweite Programm führt die Minimierung der Zielfunktion nach der Methode der dynamischen Optimierung durch und im dritten werden die Ergebnisse des letzten und vorletzten Iterationsschrittes auf Übereinstimmung überprüft. Stimmen die Werte der Wiederholungszahlen überein, ist das Iterationsziel erreicht, so daß die Ausgabedatei erstellt und alle notwendigen Daten auf einem Pageprinter in übersichtlicher Form ausgedruckt werden können. Das Verfahren konvergiert sehr schnell, so daß in den meisten Fällen nach drei bis vier Iterationsschritten der Endzustand erreicht wird.

Entspricht das Ergebnis nicht den Genauigkeitsvorstellungen, können die Eingabedaten, besonders aber der Maximalaufwand, geändert werden, um das Programmsystem dann erneut zu starten.

4. BEISPIELE

Anhand verschiedener Netzkonfigurationen soll der Programmausdruck und die Wirkungsweise verschiedener Zielfunktionen auf die Verteilung der Beobachtungen und die Genauigkeitsmaße näher erläutert werden.

Zunächst wurde das Huaytapallana-Netz (***WELSCH, HEISTER, BAUMER*** 1983) unter verschiedenen Genauigkeitsannahmen der beobachtbaren Meßgrößen optimiert. Dieses Netz (s. Anhang, Abb. 7) dient der Überwachung einer aktiven tektonischen Verwerfung in den peruanischen Anden. Da das Meßgebiet sowohl an Beobachter als auch an Instrumentarium extreme Anforderungen stellt, erscheint es besonders sinnvoll, eine Aufwandsoptimierung durchzuführen. Weil einerseits höchste Genauigkeitsforderungen zur Bestimmung signifikanter Verschiebungsparameter gestellt werden, andererseits aber der Beobachtungszeitraum begrenzt ist, soll das Optimierungsergebnis

- vor Beginn der Meßkampagne eine Entscheidungshilfe zur Instrumentenauswahl und Organisationsplanung sein

- im Felde dem verantwortlichen Beobachter als Grundlage für die vorrangige Beobachtung von Meßgrößen dienen.

Der Meßaufwand, der hierfür anzusetzen ist, konnte aus vorausgegangenen Messungen zuverlässig abgeleitet werden. Er wurde gemäß (7) ermittelt. Die optimierten Genauigkeitsmaße sind der Tabelle 1 zu entnehmen. Dabei wurden drei Optimierungsfälle (I, II, III) unterschieden, bedingt durch die unterschiedliche Kombination von Strecken- und Winkelmeßgeräten. $\overline{A}$ und $\overline{B}$ sind das Mittel der großen und kleinen Halbachse der Fehlerellipse, $\overline{PF}$ das Mittel des mittleren Punktfehlers. Die letzte Spalte gibt den Optimierungsgewinn der Zielfunktion wieder, bezogen auf eine Gleichverteilung der Beobachtungen. Als Zielfunktion wurde hier das durch (13) festgelegte skalare Genauigkeitsmaß gewählt. Der Optimierungsgewinn von 30 %, der nahezu gleichgroß in allen Fällen erreicht wird, bedeutet, daß die $\mathrm{sp}\,\boldsymbol{Q}_{xx}$ um ca. ein Drittel verkleinert wird. Bei der Genauigkeitskombination I wurde an den Einsatz eines elektronischen Tachymetersystems gedacht. Im Anhang sind dem Ausdruck zur Eingabe die möglichen Sichtverbindungen zu entnehmen. Weiterhin sind die Wiederholungszahlen angegeben, die als Startlösung zur Berechnung der $\boldsymbol{P}$-Matrix dienen. In den beiden letzten Spalten ist die mög-

liche Variationsbreite für die Wiederholungszahlen von Richtungen und Strecken ausgedruckt. Die Ausgabe zur dynamischen Optimierung enthält neben dem Koordinatenverzeichnis und den Genauigkeitsmaßen die Liste des optimierten Beobachtungsplanes. Bei der vorgegebenen Beobachtungsgenauigkeit wird trotz der ungünstigen Netzausdehnung eine relativ homogene Genauigkeit erzielt, das mittlere Achsverhältnis der Fehlerellipsen liegt bei 1.3.

Mittlere Fehler à priori	$\overline{A}$ [mm]	$\overline{B}$ [mm]	$\overline{PF}$ [mm]	$\overline{A}/\overline{B}$	Optim. Gewinn %
I m_{or} = 0.7 mgon m_{os} = 3.0 mm	1.2 (1.9)	1.0 (1.2)	1.6 (2.2)	1.3 (1.7)	31
II m_{or} = 0.3 mgon m_{os} = 7.0 mm	1.4 (2.0)	0.9 (1.3)	1.6 (2.3)	1.6 (2.2)	29
III m_{or} = 0.3 mgon m_{os} = 1.5 mm	0.6 (0.9)	0.5 (0.6)	0.7 (1.0)	1.3 (1.6)	30

Die in ( ) angegebenen Werte sind Maximalwerte

Tab. 1: Optimierte Genauigkeitsmaße bei unterschiedlichen Kombinationen von Strecken- und Winkelmeßinstrumenten

Entscheidet man sich für eine getrennte Gerätekombination (II) mit genauerem Winkelmeß- und ungenauerem Streckenmeßteil, so bewirkt dies einen leichten Genauigkeitsabfall und ein ungünstigeres Verhältnis der Halbachsen der Fehlerellipsen. Steigert man wiederum die Streckenmeßgenauigkeit, wie im Fall III angegeben, dann erreicht man bezüglich der Homogenität und Isotropie ein zu Fall I vergleichbares Ergebnis. Die Genauigkeit wird jedoch durchschnittlich um den gleichen Faktor gesteigert, wie die Streckenmeßgenauigkeit verbessert wurde. Dies bedeutet, daß ein Netz vorgegebener Konfiguration und Ausdehnung im wesentlichen nur durch die Streckenmessung in seiner Genauigkeit verbessert werden kann.

Ein zweites Beispiel - ein schematisches Netz (s. Abb. 2) - soll die Auswir-

BEOBACHTUNGS-ART		1. RICHTUNGEN					2. STRECKEN					3. RICHTUNGEN UND STRECKEN				
ZIELFUNKTION		E	R	S	W	S1	E	R	S	W	S1	E	R	S	W	S1
PUNKT NR.	A [mm]	2.2	2.3	2.4	2.4	2.5	2.1	2.1	2.1	2.0	2.3	2.2	2.5	2.2	2.6	3.0
①	B [mm]	2.1	2.1	2.1	2.1	2.2	1.8	1.8	1.8	1.8	1.7	2.0	1.9	2.1	1.9	1.8
	T [gon]	1	149	26	147	125	45	45	37	79	79	121	120	125	124	112
	A [mm]	2.8	3.0	2.7	3.0	4.3	2.4	2.4	2.5	2.4	2.3	2.1	2.2	2.1	2.0	3.8
②	B [mm]	1.6	1.6	1.7	1.5	1.5	1.8	1.8	1.8	1.8	2.0	1.8	1.7	1.8	1.7	1.7
	T [gon]	100	100	100	100	99	200	200	200	200	200	100	100	100	100	100
	A [mm]	2.2	2.3	2.4	2.4	2.5	2.1	2.1	2.1	2.0	2.3	2.2	2.5	2.2	2.6	3.0
③	B [mm]	2.1	2.1	2.1	2.0	2.1	1.8	1.8	1.8	1.8	1.7	2.0	1.9	2.1	1.9	1.8
	T [gon]	199	51	174	55	77	155	155	163	121	120	79	80	75	76	88
	A [mm]	2.2	2.3	2.4	2.4	2.5	2.1	2.1	2.1	2.0	2.3	2.2	2.5	2.2	2.6	2.9
④	B [mm]	2.1	2.2	2.1	2.0	2.2	1.8	1.8	1.8	1.8	1.7	2.0	1.9	2.1	1.9	1.9
	T [gon]	1	140	26	145	125	45	45	37	79	79	121	122	125	125	112
	A [mm]	2.8	3.0	2.7	3.0	4.3	2.4	2.4	2.5	2.4	2.3	2.1	2.2	2.1	2.0	3.6
⑤	B [mm]	1.6	1.6	1.7	1.5	1.5	1.8	1.8	1.8	1.8	2.0	1.8	1.7	1.8	1.7	1.6
	T [gon]	100	100	100	100	99	0	0	0	0	0	100	100	100	100	100
	A [mm]	2.2	2.3	2.4	2.4	2.5	2.1	2.1	2.1	2.0	2.3	2.2	2.5	2.2	2.6	2.9
⑥	B [mm]	2.1	2.2	2.1	2.1	2.1	1.8	1.8	1.8	1.8	1.7	2.0	1.9	2.1	1.9	1.9
	T [gon]	199	60	174	53	77	155	155	163	121	121	79	79	75	75	88
	$\overline{A}/\overline{B}$	1.27	1.33	1.31	1.42	1.75	1.20	1.20	1.23	1.18	1.26	1.14	1.30	1.10	1.30	1.82
	$\overline{PF}$ [mm]	3.1	3.2	3.2	3.2	3.7	2.8	2.8	2.8	2.8	2.9	2.9	3.0	3.0	3.0	3.6

Zielfunktion: E = $sp\,Q_{xx}$, R = $sp\,Q_{RR}$, S = $sp\,Q_{SS}$, W = $sp\,Q_{WW}$, S1 = $sp\,Q_{S1}$

Tab. 2: Zusammenstellung der Genauigkeitsmaße für das schematische Netz

kungen verschiedener Zielfunktionen auf die mittleren Punktfehler demonstrieren. Die Ergebnisse der Optimierungsvarianten sind in Tabelle 2 zusammengestellt. Insgesamt wurden 15 Fälle durchgerechnet, wobei fünf unterschiedliche Zielfunktionen unter Berücksichtigung von drei verschiedenen Beobachtungsverfahren minimiert wurden. Neben den bereits erwähnten Zielfunktionen (s. (13) - (16)) wurde noch die Sonderfunktion S1 optimiert. S1 bedeutet, daß die Matrix $\boldsymbol{F}$ die Koeffizienten der Strecken 1-6, 2-5 und 3-4 enthält und somit deren Gewichtsreziproken minimiert werden.

Für alle Netzpunkte sind dann die unter dem jeweiligen Optimalitätskriterium bestimmten Halbachsen A und B sowie ihre Richtungswinkel T zusammengestellt; ferner sind noch die bereits erwähnten Größen $\overline{A}/\overline{B}$ und $\overline{PF}$ angegeben. Der Gesamtaufwand wurde für alle Fälle auf S = 100 festgesetzt, die Genauigkeiten für die Richtungs- und Streckenmessungen mit $m_{or} = \pm 0.5\,\text{mgon}$ und $m_{os} = \pm 1.0\,\text{cm}$ angenommen. Die Wiederholungszahlen für die Richtungen wurden gleichbleibend für alle Fälle auf 0 bis 6 festgelegt; die Wiederholungszahlen für die Streckenmessung betrugen 0 und 1. Eine Ausnahme wurde jedoch beim reinen Streckennetz gemacht. Um hier die Schwachstellen eines Netzes erkennen zu können, ist ebenfalls eine Variationsbreite der Wiederholungszahlen von 0 bis 6 zugelassen.

Betrachtet man zunächst die beiden letzten Zeilen der Tabelle 1, so zeigt der mittlere Punktfehler $\overline{PF}$ innerhalb einer Beobachtungsart für die ersten vier Zielfunktionen wenig Veränderung; nur das Maß für die Isotropie $\overline{A}/\overline{B}$ wird unterschiedlich beeinflußt. Eine einheitliche Interpretation der Ergebnisse für die unterschiedlichen Beobachtungsarten ist nicht möglich, jedoch ist deutlich abzulesen, daß einerseits die Beobachtungsart 'Strecken', andererseits die Zielfunktion $\text{sp}\,\boldsymbol{Q}_{xx}$ einen starken Einfluß auf die Homogenität und Isotropie des Netzes haben. Bemerkenswert ist ebenfalls der Beitrag der einzelnen Punkte auf dieses summarische Ergebnis. Sind z.B. für Richtungs- oder Streckennetze die Punkte 2 und 5 diejenigen, die Extremwerte für Fehlerellipsen und Achsverhältnis annehmen, so ändert sich diese Aussage nahezu gänzlich bei der kombinierten Ausgleichung von Richtungen und Strecken. Bei der Optimierung der Sonderfunktion ist deutlich zu erkennen, daß das Optimum der ausgewählten Streckenfunktionsgewichte nur unter einer deutlichen Verschlechterung der Netzisotropie erreicht wird. Die Abb. 2 und 5 geben für jede Beobachtungsart sowie für die Sonderfunktion den optimierten Beobachtungsplan wieder. Hierbei wurden den einzelnen Sichtverbindungen die opti-

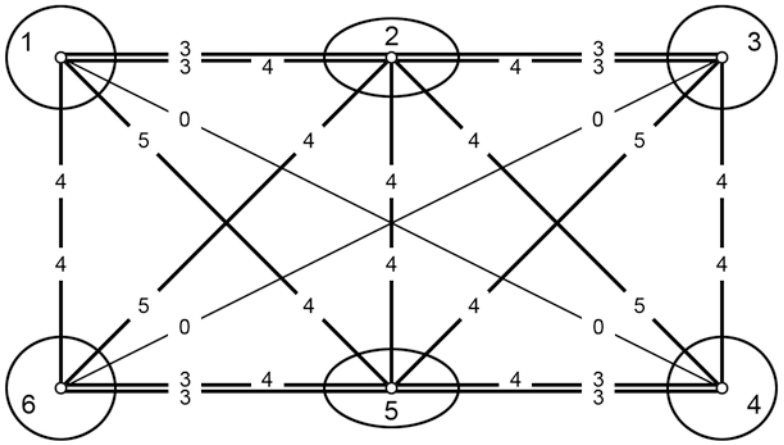


Abb. 2:

Optimaler Beobachtungsplan für ein Richtungsnetz
Min. sp Q_{xx}

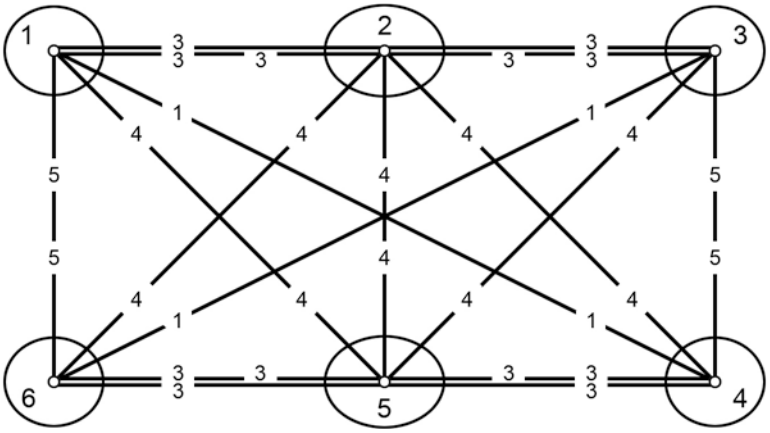


Abb. 3:

Optimaler Beobachtungsplan für ein Streckennetz
Min. sp Q_{WW}

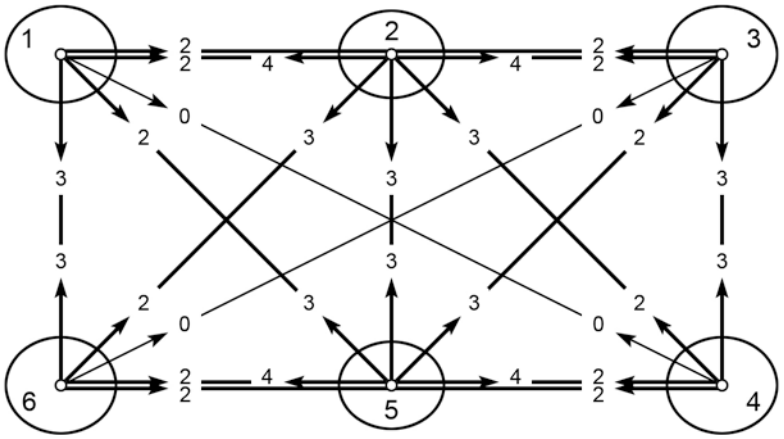


Abb. 4:

Optimaler Beobachtungsplan für ein kombiniertes Netz
Min. sp Q_{xx}

→ Streckenmessung

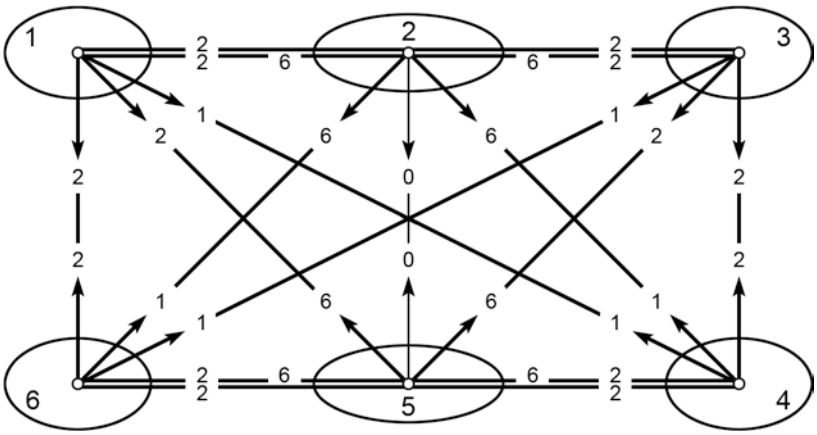


Abb. 5:

Optimaler Beobachtungsplan für ein kombiniertes Netz
Min. sp Q_{S1}

malen Wiederholungszahlen zugeordnet; schwarze Pfeile geben eine notwendige Streckenmessung an.

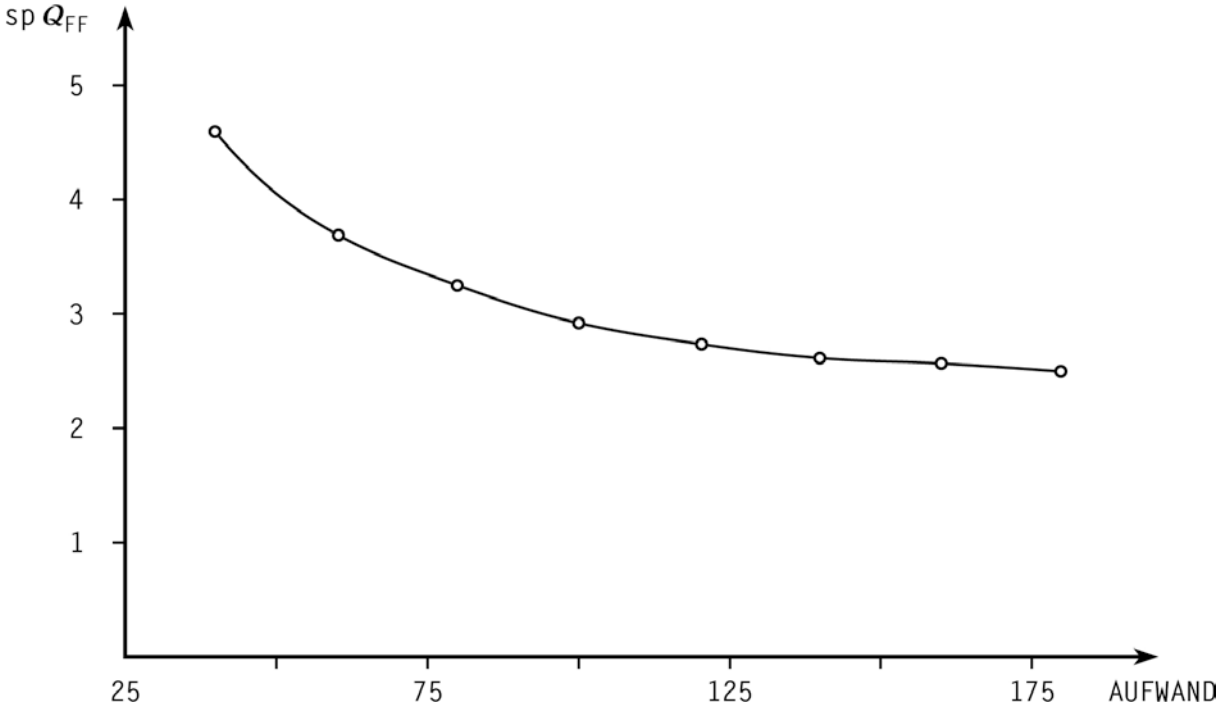


Abb. 6: Aufwands- und Genauigkeitssteigerung

Sind die vorangegangenen Beispiele der reinen Aufwandsoptimierung zuzurechnen, so kann mit dem Programmsystem NOPTI II auch eine Genauigkeitsoptimierung durchgeführt werden. In einem iterativen Rechenablauf wird dann die Aufwandszahl in vorgegebenen Schrittweiten so lange erhöht, bis der vorgegebene Wert der skalaren Genauigkeit für die Zielfunktion erreicht wird. Um einen Eindruck über die Wechselwirkung von Aufwands- und Genauigkeitssteigerung zu bekommen, gibt die Abb. 6 für das bereits vorgestellte schematische Netz den charakteristischen Kurvenverlauf wieder. Dieser zeigt auch deutlich, daß die Netzgenauigkeit ab einer gewissen Stufe nur noch unwesentlich gesteigert werden kann, wobei dann hierfür eine unverhältnismäßig hohe Beobachtungsgenauigkeit - denn Aufwand bedeutet auch Beobachtungsgenauigkeit - angesetzt werden muß.

5. SCHLUSSBEMERKUNG

Der vorliegende Beitrag sollte anhand verschiedener Beispiele das praxisorientierte Programmsystem NOPTI II vorstellen. Hierin wird unter Berücksichti-

gung unterschiedlicher invarianter Zielfunktionen der Beobachtungsplan von Lagenetzen optimiert. Dabei ist sowohl eine Aufwands- als auch Genauigkeitsoptimierung möglich. Die Beispiele zeigen, daß man unter realistischen Vorgaben auch Optimalität bezüglich Homogenität und Isotropie erreichen kann. Ein besonderer Anwendungsbereich dürfte die Planung von Sondernetzen sein, die - bedingt durch ihre Verwendung - bestimmte Optimalitätskriterien erfüllen müssen. Letztlich sei nochmals darauf hingewiesen, daß der mathematische Optimierungsalgorithmus als ein objektives Hilfsmittel für Planung, Genauigkeitsvoranschlag, Organisation und Durchführung der Feldmessungen verstanden werden muß.

LITERATUR

GRAFAREND, E. e.a., 1979, ***Optimierung geodätischer Meßoperationen***, Herbert Wichmann Verlag, Karlsruhe

HEISTER, H., 1978, ***Die diskrete dynamische Optimierung und ihre Anwendung beim geodätischen Netzentwurf***, AVN, 85, S. 64-81

SCHMITT, G., 1982, ***Optimization of Geodetic Networks***, Reviews of Geophysics and Space Physics, Vol. 20, Nr. 4, S. 877-884

WELSCH, W.M., HEISTER, H., BAUMER, R., 1983, ***Geodetic Monitoring of an Active Fault in the Peruvian Andes. The Huaytapallana Faults 1975-1982***, Zbl. Geol. Paläont. Teil I, S. 363-374

ANHANG

Ergebnisausdruck des optimierten Huaytapallana-Netzes

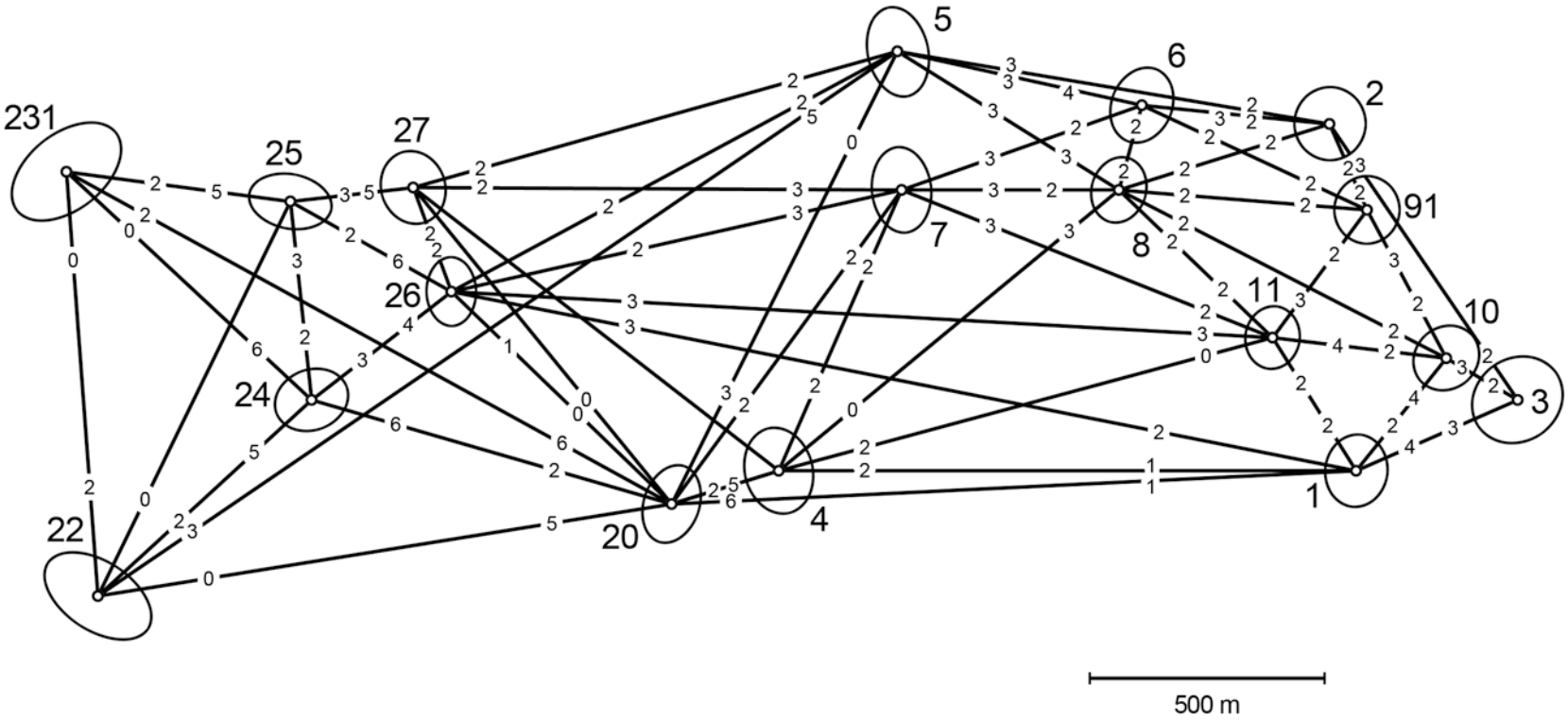


Abb. 7: Huaytapallana-Netz und optimierter Beobachtungsplan

KOORDINATENVERZEICHNIS

PKTNR	Y	X
1	2232.19900	1000.00900
2	2175.87500	1737.14600
3	2577.37800	1148.35800
4	999.99400	1000.00900
5	1252.82800	1893.09300
6	1774.56700	1778.64100
7	1262.64200	1596.97200
8	1725.75500	1596.57900
91	2253.88200	1554.87600
10	2424.44600	1237.11400
11	2053.95800	1282.11400
20	770.68100	925.87100
22	-453.38000	730.62700
231	-520.32400	1635.78400
24	2.58700	1148.97900
25	-42.91900	1571.81700
26	300.94900	1379.44600
27	219.44300	1602.23200

7.0000 = SORTG (0.1MGON)
0.0030 = SOSTR (M)
0.0000 = SIGMAR (0.1MGON)
0.0000 = SIGMAS (M)
1.0000 = CONSTR
1.0000 = CONSTS

$$P = \frac{CONST * W}{W*SIGMA^{**}2 + SO^{**}2}$$

W = WIEDERHOLUNGSZAHL

LFD.NR.	STANDPUNKT	ZIELPUNKT	WIEDERHOLUNGSZAHLEN RICHTUNGEN	WIEDERHOLUNGSZAHLEN STRECKEN	ZULAESSIGE WIEDERHOLUNGSZAHLEN RICHTUNGEN	ZULAESSIGE WIEDERHOLUNGSZAHLEN STRECKEN
1	1	20	3	1	0 6	0 2
2	1	4	3	1	0 6	0 2
3	1	26	3	1	0 6	0 2
4	1	11	3	1	0 6	0 2
5	1	10	3	1	0 6	0 2
6	1	3	3	1	0 6	0 2
7	2	3	3	1	0 6	0 2
8	2	91	3	1	0 6	0 2
9	2	8	3	1	0 6	0 2
10	2	6	3	1	0 6	0 2
11	2	5	3	1	0 6	0 2
12	3	1	3	KEINE BEOBACHTUNG	0 6	0 0

*	13	*	3	*	10	*	3	*	1	*	0	6	*	0	2	*
*	14	*	3	*	2	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	15	*	4	*	20	*	3	*	1	*	0	6	*	0	2	*
*	16	*	4	*	7	*	3	*	1	*	0	6	*	0	2	*
*	17	*	4	*	8	*	3	*	1	*	0	6	*	0	2	*
*	18	*	4	*	11	*	3	*	1	*	0	6	*	0	2	*
*	19	*	4	*	1	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	20	*	5	*	2	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	21	*	5	*	6	*	3	*	1	*	0	6	*	0	2	*
*	22	*	5	*	8	*	3	*	1	*	0	6	*	0	2	*
*	23	*	5	*	20	*	3	*	1	*	0	6	*	0	2	*
*	24	*	5	*	22	*	3	*	1	*	0	6	*	0	2	*
*	25	*	5	*	26	*	3	*	1	*	0	6	*	0	2	*
*	26	*	5	*	27	*	3	*	1	*	0	6	*	0	2	*
*	27	*	6	*	2	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	28	*	6	*	91	*	3	*	1	*	0	6	*	0	2	*
*	29	*	6	*	8	*	3	*	1	*	0	6	*	0	2	*
*	30	*	6	*	7	*	3	*	1	*	0	6	*	0	2	*
*	31	*	6	*	5	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	32	*	7	*	6	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	33	*	7	*	8	*	3	*	1	*	0	6	*	0	2	*
*	34	*	7	*	11	*	3	*	1	*	0	6	*	0	2	*
*	35	*	7	*	4	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	36	*	7	*	20	*	3	*	1	*	0	6	*	0	2	*
*	37	*	7	*	26	*	3	*	1	*	0	6	*	0	2	*
*	38	*	7	*	27	*	3	*	1	*	0	6	*	0	2	*
*	39	*	8	*	2	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	40	*	8	*	91	*	3	*	1	*	0	6	*	0	2	*
*	41	*	8	*	10	*	3	*	1	*	0	6	*	0	2	*
*	42	*	8	*	11	*	3	*	1	*	0	6	*	0	2	*
*	43	*	8	*	4	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	44	*	8	*	7	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	45	*	8	*	5	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	46	*	8	*	6	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	47	*	91	*	2	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	48	*	91	*	10	*	3	*	1	*	0	6	*	0	2	*
*	49	*	91	*	11	*	3	*	1	*	0	6	*	0	2	*
*	50	*	91	*	8	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	51	*	91	*	6	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	52	*	10	*	3	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	53	*	10	*	1	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	54	*	10	*	11	*	3	*	1	*	0	6	*	0	2	*
*	55	*	10	*	8	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	56	*	10	*	91	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	57	*	11	*	1	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	58	*	11	*	4	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*
*	59	*	11	*	26	*	3	*	1	*	0	6	*	0	2	*
*	60	*	11	*	7	*	3	*	KEINE BEOBACHTUNG	*	0	6	*	0	0	*

61	11	8	3	KEINE BEOBACHTUNG	0	6	0	0
62	11	91	3	KEINE BEOBACHTUNG	0	6	0	0
63	11	10	3	KEINE BEOBACHTUNG	0	6	0	0
64	20	22	3	1	0	6	0	2
65	20	24	3	1	0	6	0	2
66	20	231	3	1	0	6	0	2
67	20	26	3	1	0	6	0	2
68	20	27	3	1	0	6	0	2
69	20	5	3	KEINE BEOBACHTUNG	0	6	0	0
70	20	7	3	KEINE BEOBACHTUNG	0	6	0	0
71	20	4	3	KEINE BEOBACHTUNG	0	6	0	0
72	20	1	3	KEINE BEOBACHTUNG	0	6	0	0
73	22	231	3	1	0	6	0	2
74	22	25	3	1	0	6	0	2
75	22	24	3	1	0	6	0	2
76	22	5	3	KEINE BEOBACHTUNG	0	6	0	0
77	22	20	3	KEINE BEOBACHTUNG	0	6	0	0
78	231	25	3	1	0	6	0	2
79	231	20	3	KEINE BEOBACHTUNG	0	6	0	0
80	231	24	3	1	0	6	0	2
81	231	22	3	KEINE BEOBACHTUNG	0	6	0	0
82	24	22	3	KEINE BEOBACHTUNG	0	6	0	0
83	24	231	3	KEINE BEOBACHTUNG	0	6	0	0
84	24	25	3	1	0	6	0	2
85	24	26	3	1	0	6	0	2
86	24	20	3	KEINE BEOBACHTUNG	0	6	0	0
87	25	27	3	1	0	6	0	2
88	25	26	3	1	0	6	0	2
89	25	24	3	KEINE BEOBACHTUNG	0	6	0	0
90	25	231	3	KEINE BEOBACHTUNG	0	6	0	0
91	26	5	3	KEINE BEOBACHTUNG	0	6	0	0
92	26	7	3	KEINE BEOBACHTUNG	0	6	0	0
93	26	11	3	KEINE BEOBACHTUNG	0	6	0	0
94	26	1	3	KEINE BEOBACHTUNG	0	6	0	0
95	26	20	3	KEINE BEOBACHTUNG	0	6	0	0
96	26	24	3	KEINE BEOBACHTUNG	0	6	0	0
97	26	25	3	KEINE BEOBACHTUNG	0	6	0	0
98	26	27	3	1	0	6	0	2
99	27	5	3	KEINE BEOBACHTUNG	0	6	0	0
100	27	7	3	KEINE BEOBACHTUNG	0	6	0	0
101	27	26	3	KEINE BEOBACHTUNG	0	6	0	0
102	27	25	3	KEINE BEOBACHTUNG	0	6	0	0

ZIELFUNKTION: MINIMUM DER SPUR Qxx

Ende der Liste

KOORDINATENVERZEICHNIS FEHLERELLIPSEN

PKTNR	Y (M)	X (M)	T (GON)	A (MM)	B (MM)	A/B	PF (MM)
1	2232.19900	1000.00900	7.8635	1.10	0.95	1.16	1.46
2	2175.87500	1737.14600	2.1893	1.11	1.08	1.02	1.55
3	2577.37800	1148.35800	61.1813	1.42	1.25	1.14	1.89
4	999.99400	1000.00900	183.6181	1.31	1.03	1.27	1.66
5	1252.82800	1893.09300	185.9905	1.37	0.92	1.48	1.65
6	1774.56700	1778.64100	25.2378	1.17	0.92	1.27	1.49
7	1262.64200	1596.97200	193.3584	1.30	0.90	1.43	1.58
8	1725.75500	1596.57900	18.7896	1.02	0.83	1.24	1.31
91	2253.88200	1554.87600	2.1443	1.03	1.00	1.04	1.44
10	2424.44600	1237.11400	54.8811	1.05	0.94	1.11	1.41
11	2053.95800	1282.11400	9.7218	0.96	0.84	1.14	1.28
20	770.68100	925.87100	20.6202	1.20	0.86	1.39	1.48
22	-453.38000	730.62700	137.0052	1.81	1.05	1.73	2.09
231	-520.32400	1635.78400	57.4012	1.89	1.16	1.64	2.22
24	2.58700	1148.97900	76.9265	1.14	0.93	1.23	1.47
25	-42.91900	1571.81700	112.1195	1.26	0.83	1.52	1.50
26	300.94900	1379.44600	198.7484	1.04	0.77	1.34	1.30
27	219.44300	1602.23200	179.4037	1.11	0.97	1.14	1.47
			MITTEL	1.24	0.96	1.29	1.57

OPTIMIERUNGSGEWINN = 31 PROZENT (GEGENUEBER EINGABE)
ANZAHL DER ITERATIONEN = 4

LFD.NR.	STANDPUNKT	ZIELPUNKT	MITTLERE FEHLER RICHTUNGEN (0.1 MGON)	MITTLERE FEHLER STRECKEN (M)	WIEDERHOLUNGSZAHLEN RICHTUNGEN	WIEDERHOLUNGSZAHLEN STRECKEN
1	1	20	7.0000	0.0021	1	2
2	1	4	7.0000	0.0021	1	2

*	3	*	1	*	26	*	4.9497	*	0.0021	*	2	*	2	*	
*	4	*	1	*	11	*	4.9497	*	0.0021	*	2	*	2	*	
*	5	*	1	*	10	*	4.9497	*	0.0021	*	2	*	2	*	
*	6	*	1	*	3	*	3.5000	*	0.0021	*	4	*	2	*	
*	7	*	2	*	3	*	4.0415	*	0.0021	*	3	*	2	*	
*	8	*	2	*	91	*	4.9497	*	0.0021	*	2	*	2	*	
*	9	*	2	*	8	*	4.9497	*	0.0021	*	2	*	2	*	
*	10	*	2	*	6	*	4.9497	*	0.0021	*	2	*	2	*	
*	11	*	2	*	5	*	4.9497	*	0.0021	*	2	*	2	*	
*	12	*	3	*	1	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	13	*	3	*	10	*	4.9497	*	0.0021	*	2	*	2	*	
*	14	*	3	*	2	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	15	*	4	*	20	*	3.1305	*	0.0021	*	5	*	2	*	
*	16	*	4	*	7	*	4.9497	*	0.0021	*	2	*	2	*	
*	17	*	4	*	8	*	KEINE BEOBACHTUNG	*	0.0021	*	0	*	2	*	
*	18	*	4	*	11	*	4.9497	*	0.0021	*	2	*	2	*	
*	19	*	4	*	1	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	20	*	5	*	2	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	21	*	5	*	6	*	4.0415	*	0.0021	*	3	*	2	*	
*	22	*	5	*	8	*	4.0415	*	0.0021	*	3	*	2	*	
*	23	*	5	*	20	*	KEINE BEOBACHTUNG	*	0.0021	*	0	*	2	*	
*	24	*	5	*	22	*	3.1305	*	0.0021	*	5	*	2	*	
*	25	*	5	*	26	*	4.9497	*	0.0021	*	2	*	2	*	
*	26	*	5	*	27	*	4.9497	*	0.0021	*	2	*	2	*	
*	27	*	6	*	2	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	28	*	6	*	91	*	4.9497	*	0.0021	*	2	*	2	*	
*	29	*	6	*	8	*	4.9497	*	0.0021	*	2	*	2	*	
*	30	*	6	*	7	*	4.9497	*	0.0021	*	2	*	2	*	
*	31	*	6	*	5	*	3.5000	*	KEINE BEOBACHTUNG	*	4	*	0	*	
*	32	*	7	*	6	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	33	*	7	*	8	*	4.0415	*	0.0021	*	3	*	2	*	
*	34	*	7	*	11	*	4.0415	*	0.0021	*	3	*	2	*	
*	35	*	7	*	4	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	36	*	7	*	20	*	4.9497	*	0.0021	*	2	*	2	*	
*	37	*	7	*	26	*	4.0415	*	0.0021	*	3	*	2	*	
*	38	*	7	*	27	*	4.0415	*	0.0021	*	3	*	2	*	
*	39	*	8	*	2	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	40	*	8	*	91	*	4.9497	*	0.0030	*	2	*	1	*	
*	41	*	8	*	10	*	4.9497	*	0.0021	*	2	*	2	*	
*	42	*	8	*	11	*	4.9497	*	0.0021	*	2	*	2	*	
*	43	*	8	*	4	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	44	*	8	*	7	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	45	*	8	*	5	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	46	*	8	*	6	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	47	*	91	*	2	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	48	*	91	*	10	*	4.0415	*	0.0021	*	3	*	2	*	
*	49	*	91	*	11	*	4.9497	*	0.0021	*	2	*	2	*	
*	50	*	91	*	8	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	

*	51	*	91	*	6	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	52	*	10	*	3	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	53	*	10	*	1	*	3.5000	*	KEINE BEOBACHTUNG	*	4	*	0	*	
*	54	*	10	*	11	*	4.9497	*	0.0021	*	2	*	2	*	
*	55	*	10	*	8	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	56	*	10	*	91	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	57	*	11	*	1	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	58	*	11	*	4	*	KEINE BEOBACHTUNG	*	KEINE BEOBACHTUNG	*	0	*	0	*	
*	59	*	11	*	26	*	4.0415	*	0.0021	*	3	*	2	*	
*	60	*	11	*	7	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	61	*	11	*	8	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	62	*	11	*	91	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	63	*	11	*	10	*	3.5000	*	KEINE BEOBACHTUNG	*	4	*	0	*	
*	64	*	20	*	22	*	3.1305	*	0.0021	*	5	*	2	*	
*	65	*	20	*	24	*	4.9497	*	0.0021	*	2	*	2	*	
*	66	*	20	*	231	*	2.8577	*	0.0021	*	6	*	2	*	
*	67	*	20	*	26	*	KEINE BEOBACHTUNG	*	0.0021	*	0	*	2	*	
*	68	*	20	*	27	*	KEINE BEOBACHTUNG	*	0.0021	*	0	*	2	*	
*	69	*	20	*	5	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	70	*	20	*	7	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	71	*	20	*	4	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	72	*	20	*	1	*	2.8577	*	KEINE BEOBACHTUNG	*	6	*	0	*	
*	73	*	22	*	231	*	4.9497	*	0.0021	*	2	*	2	*	
*	74	*	22	*	25	*	KEINE BEOBACHTUNG	*	0.0021	*	0	*	2	*	
*	75	*	22	*	24	*	4.9497	*	0.0021	*	2	*	2	*	
*	76	*	22	*	5	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	77	*	22	*	20	*	KEINE BEOBACHTUNG	*	KEINE BEOBACHTUNG	*	0	*	0	*	
*	78	*	231	*	25	*	4.9497	*	0.0021	*	2	*	2	*	
*	79	*	231	*	20	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	80	*	231	*	24	*	KEINE BEOBACHTUNG	*	0.0021	*	0	*	2	*	
*	81	*	231	*	22	*	KEINE BEOBACHTUNG	*	KEINE BEOBACHTUNG	*	0	*	0	*	
*	82	*	24	*	22	*	3.1305	*	KEINE BEOBACHTUNG	*	5	*	0	*	
*	83	*	24	*	231	*	2.8577	*	KEINE BEOBACHTUNG	*	6	*	0	*	
*	84	*	24	*	25	*	4.9497	*	0.0021	*	2	*	2	*	
*	85	*	24	*	26	*	4.0415	*	0.0021	*	3	*	2	*	
*	86	*	24	*	20	*	2.8577	*	KEINE BEOBACHTUNG	*	6	*	0	*	
*	87	*	25	*	27	*	4.0415	*	0.0021	*	3	*	2	*	
*	88	*	25	*	26	*	4.9497	*	0.0021	*	2	*	2	*	
*	89	*	25	*	24	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	90	*	25	*	231	*	3.1305	*	KEINE BEOBACHTUNG	*	5	*	0	*	
*	91	*	26	*	5	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	92	*	26	*	7	*	4.9497	*	KEINE BEOBACHTUNG	*	2	*	0	*	
*	93	*	26	*	11	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	94	*	26	*	1	*	4.0415	*	KEINE BEOBACHTUNG	*	3	*	0	*	
*	95	*	26	*	20	*	7.0000	*	KEINE BEOBACHTUNG	*	1	*	0	*	
*	96	*	26	*	24	*	3.5000	*	KEINE BEOBACHTUNG	*	4	*	0	*	
*	97	*	26	*	25	*	2.8577	*	KEINE BEOBACHTUNG	*	6	*	0	*	
*	98	*	26	*	27	*	4.9497	*	0.0021	*	2	*	2	*	

99	27	5	4.9497	KEINE BEOBACHTUNG	2		0	
100	27	7	4.9497	KEINE BEOBACHTUNG	2		0	
101	27	26	4.9497	KEINE BEOBACHTUNG	2		0	
102	27	25	3.1305	KEINE BEOBACHTUNG	5		0	

7.0000 = SORTG (0.1MGON)
0.0030 = SOSTR (M)
0.0000 = SIGMAR (0.1MGON)
0.0000 = SIGMAS (M)
1.0000 = CONSTR
1.0000 = CONSTS

$$P = \frac{CONST * W}{W*SIGMA^{**}2 + SO^{**}2}$$

W = WIEDERHOLUNGSZAHL

358 = vorgegebener Maximalaufwand

ZIELFUNKTION: MINIMUM DER SPUR Qxx

Ende der Liste

EIN KONZEPT ZUR INTERAKTIVEN PLANUNG, AUSGLEICHUNG UND ANALYSE GEODÄTISCHER NETZE

von Rolf KÖNIG

SUMMARY

The interest in computer aided design, adjustment and analysis of geodetic networks has lately considerably increased. This paper, starting with the usual procedure of a surveying project, describes a concept for a modular programme by flow charts. The design problems are to be solved according to the trial-and-error manner. The interactions to be taken into consideration between human being and computer equipment have been made evident.

ZUSAMMENFASSUNG

Die computerunterstützte Planung, Ausgleichung und Analyse geodätischer Netze findet in letzter Zeit immer größeres Interesse. In diesem Beitrag wird, ausgehend vom üblichen Verlauf eines Vermessungsprojekts, ein Konzept für ein modulares Programmsystem anhand von Flußdiagrammen dargestellt. Die Designprobleme sollen nach dem Trial-and-error-Prinzip gelöst werden. Die zu berücksichtigenden Interaktionen zwischen Mensch und Maschine werden verdeutlicht.

1. EINFÜHRUNG

Bearbeitet der Ingenieur ein geodätisches Netz, so erlebt er den Ablauf des Vermessungsprojekts i.a. in den in BILD 1 dargestellten Schritten. Zu jedem Zeitpunkt kann er gezwungen sein, seine bisherigen Planungen und Arbeiten zu überdenken und zu ändern. Die folgenden Beispiele sollen dies verdeutlichen:

- Das Netz ist in der häuslichen Vorbereitung entworfen worden. Neubebauungen machen jedoch die Beobachtung einer Station unmöglich.
- Der Beobachtungsplan kann in der verfügbaren Zeitspanne nicht durchgeführt werden.
- Die Berechnung weist grobe Datenfehler aus. Wiederholungsmessungen werden nötig.
- Die Analyse ergibt, daß ein Punkt nicht ausreichend genau bestimmt wurde. Zusätzliche oder genauere Messungen schaffen Abhilfe.
- Einige Ziele des Projekts sind nicht erreicht worden. Zusätzliche Planungen und Messungen sind unumgänglich.
- Äußere Einflüsse wie geänderte Anforderungen des Auftraggebers können jederzeit auf das Projekt einwirken.

Diese komplexe Aufgabe mit ihren vielfältigen Interaktionen wird noch heute meist mit Hilfe von Erfahrung und Gefühl gelöst. Dabei bieten Computerunterstützung und vorhandene Rechenalgorithmen die Möglichkeit, die Erfahrung des Ingenieurs zu ergänzen und Ablauf und Lösung des Projekts zu optimieren. An dieser Stelle wird ein Konzept vorgestellt, das geeignet ist, ein geodätisches Netz im Dialog mit dem Computer zu planen, auszugleichen und zu analysieren. Dabei wird durch einen "Updating"-Prozeß ständig eine Anpassung an veränderte Gegebenheiten erreicht.

In BILD 1 wird deutlich, daß sich der Projektablauf in vier Blöcke gliedern läßt:

Phase I: Die Projektspezifikationen
Phase II: Die Netzplanung
Phase III: Die Ausführungsphase
Phase IV: Die Dokumentation und Präsentation der Ergebnisse.

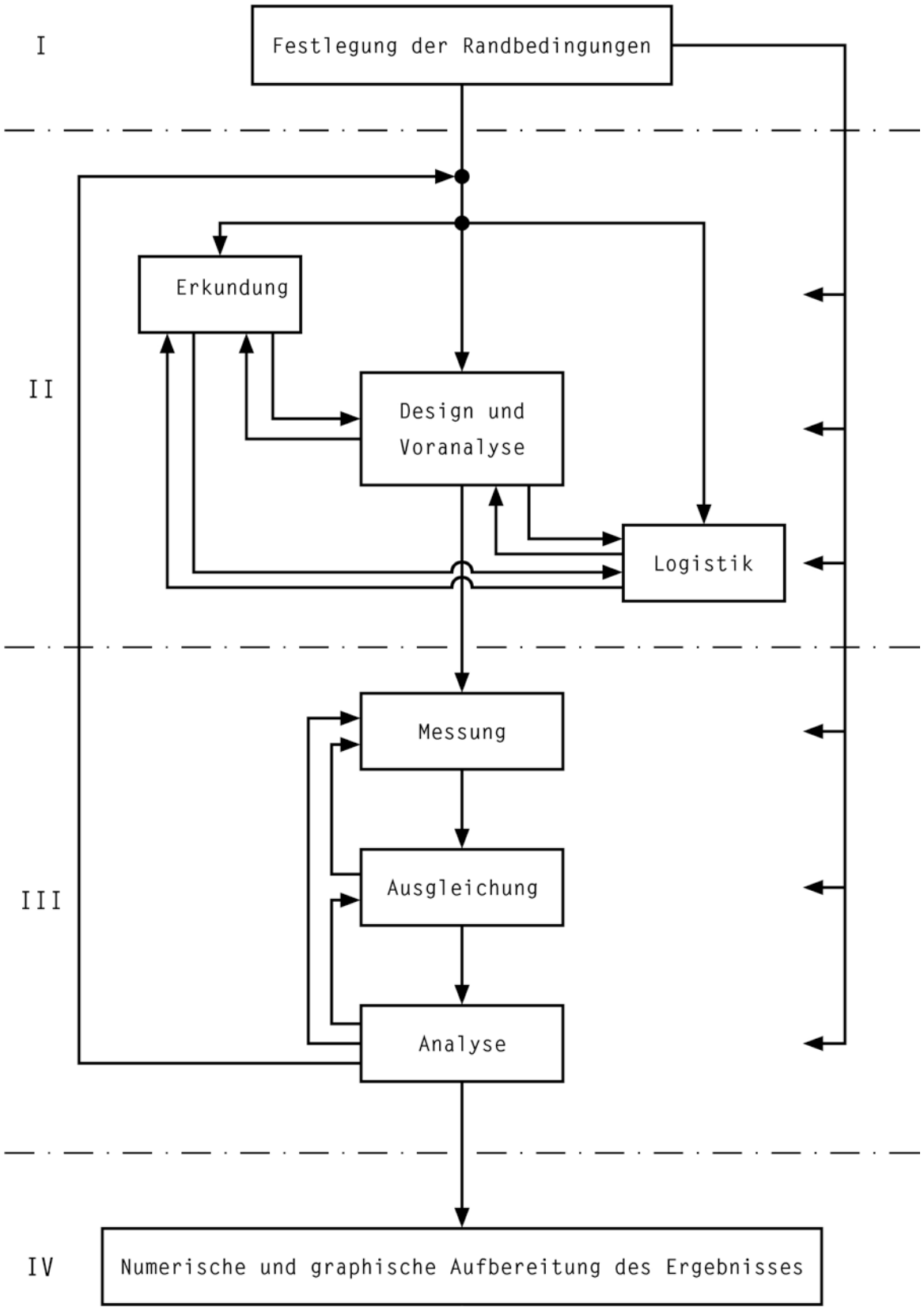

I
Festlegung der Randbedingungen
Erkundung
II
Design und
Voranalyse
Logistik
Messung
III
Ausgleichung
Analyse
IV
Numerische und graphische Aufbereitung des Ergebnisses

Bild 1

Zu Phase I mag die Bemerkung genügen, daß sich in Abhängigkeit vom Auftrag rein technische Anforderungen, wie z.B. Objektnähe, Punktdichte, Genauigkeit, Zuverlässigkeit, und organisationstechnische Anforderungen ergeben, wie z.B. Verfügbarkeit von Personal und Instrumenten, Zeitaufwand, Kosten.

Die Phasen II und III werden in den nächsten zwei Kapiteln näher betrachtet, während auf Phase IV nicht eingegangen wird.

2. DIE NETZPLANUNG

Zur Netzplanung sollen die Erkundung, die Lösung der Designprobleme O. bis 3. Ordnung (vgl. ***GRAFAREND*** u.a., 1979) und die Optimierung der Logistik gezählt werden. Der Entwurf eines Netzes, modern Design genannt, wird gewöhnlicherweise folgendermaßen klassifiziert:

Design O. Ordnung: Die Wahl eines Referenzsystems.

Design 1. Ordnung: Die Festlegung der Netzkonfiguration im Hinblick auf die möglichen Messungen.

Design 2. Ordnung: Die Verteilung des Meßaufwands bei vorgegebener Konfiguration, deshalb auch Gewichtsoptimierung genannt.

Design 3. Ordnung: Die Verdichtung eines bestehenden Netzes.

Um das Design mathematisch zu optimieren, müssen bestimmte Kriterien als Zielfunktionen formuliert werden. Diese Designkriterien lauten:

- Genauigkeit,
- Zuverlässigkeit,
- Wirtschaftlichkeit,
- Testbarkeit.

Letzteres soll bedeuten, daß das Netz so gestaltet sein muß, daß spezielle Hypothesen bei kleinstem Vertrauensbereich getestet werden können. Die Designkriterien haben einen teilweise konträren Einfluß auf die Optimierung (z.B. die Forderung von höchsten Genauigkeiten und minimalen Kosten). Die Zielfunktionen bestehen aus zwei grundlegenden Typen:

- Skalare Funktionen,
- Kriterionmatrizen.

Als skalare Funktionen sind Spur, Determinante und Funktionen der Eigenwerte der Varianz - Kovarianzmatrix der Unbekannten häufig in Gebrauch. Kriterionmatrizen können als idealisierte Varianz - Kovarianzmatrizen angesehen werden, die die geforderte Genauigkeitssituation des Netzes wiedergeben und durch den Designprozeß approximiert werden sollen.

Nach dem derzeitigen Stand der Wissenschaft gibt es abgesicherte mathematische Lösungen nur für spezielle Optimierungsprobleme (siehe hierzu ***CROSS***, 1983). So kann zum Beispiel bei fester Konfiguration und festgelegten Genauigkeitsanforderungen an die Lösungsparameter (Koordinatenunbekannte) eine Gewichtsoptimierung berechnet werden. ***NINKOV*** und ***SCHMITT*** (1983) stellen dazu eine Methode vor, die allerdings bei Netzen mit verschiedenen Beobachtungstypen auf Schwierigkeiten stößt. Das Verhalten der Methode von ***SCHAFFRIN*** (1981) scheint noch in der Praxis untersucht werden zu müssen. In der umfangreichen Literatur zum Design 2. Ordnung wurde dem Kriterium Zuverlässigkeit wenig Beachtung geschenkt. Die Zuverlässigkeit eines Netzes ist aber abhängig vom Beobachtungsplan und sollte deshalb in die Optimierung einbezogen werden. ***VAN MIERLO*** (1981) hat auf dieses Problem hingewiesen.

Ein mathematisches Modell zur geschlossenen Bearbeitung aller Aspekte ist zur Zeit wohl nicht realisierbar. Eine Lösungsmöglichkeit ist aber das sogenannte Monte-Carlo-Design. Durch einen Zufallsprozeß wird eine Menge von Designvarianten erzeugt. Die Ergebnisse werden dann im Hinblick auf die Zielfunktionen verglichen. Diese Methode hat den Vorteil, daß die optimale Variante unter gleichzeitiger Betrachtung aller Designkriterien ausgewählt wird. Allerdings stößt man hier sehr schnell an Grenzen, die durch die verfügbare Rechenzeit gesetzt sind. Das Ergebnis wird daher stets nur ein relatives Optimum sein können.

Der erfolgversprechendste Lösungsweg scheint deshalb die Simulation verbunden mit einer Auswahl nach dem Trial-and-error-Prinzip zu sein. Erfahrung und Sachverstand des Bearbeiters spielen dabei eine gewisse Rolle. Der bei dieser operationellen Vorgehensweise nicht zu vermeidende Mehraufwand wird bewußt in Kauf genommen. Das iterative Verfahren liefert zwar ebenfalls nicht die optimale Variante, jedoch kann die Bearbeitung abgebrochen werden, wenn ein zufriedenstellendes Ergebnis erzielt worden ist. ***MEPHAM*** und ***KRAKIWSKY*** (1982) haben diese Methode mit Erfolg angewandt.

Liegen Netzfigur und Beobachtungsplan fest, so kann die Logistik optimiert werden. Gemeint ist damit die Minimierung der Kosten der Feldarbeiten. Besonders beim Einsatz teurer Ausrüstungen, aber auch bei den Personalkosten können immense Einsparungen erzielt werden. ***STOLIKER*** und ***ANDERSON*** (1982) lieferten Beispiele für Optimierungsmethoden.

Die Interaktionen zwischen Bearbeiter und Computer während der Planungsphase soll BILD 2 verdeutlichen (Verfeinerung der Phase II in BILD 1). Die Entscheidungen und Planungen des Bearbeiters werden durch Software-Module unterstützt. Außerdem ermöglicht der Einsatz von entsprechender Hardware, wie graphisches Terminal und Digitizer, einen weniger problematischen Arbeitsablauf und visuelle Entscheidungshilfen (vgl. dazu Kapitel 4).

3. DIE AUSFÜHRUNGSPHASE

Die Planung wird nun in die Tat umgesetzt. Der Zusammenhang zwischen menschlichen Tätigkeiten und Entscheidungen und der EDV ist in BILD 3 dargestellt (Verfeinerung der Phase III in BILD 1).

Während der Meßkampagne erfolgt der Dialog mit dem Computer täglich. Dieses Vorgehen gewährleistet eine vollständige Dokumentation der durchgeführten Arbeiten, eine unmittelbare Analyse der Ergebnisse und ein frühzeitiges Aufdecken grober Meßfehler. Dabei erfolgt jeweils (nur) eine Aufdatierung (Updating) des bestehenden Normalgleichungsystems. D.h. das Normalgleichungssystem muß am Rechner vorliegen vom Zeitpunkt der Netzplanung bis zur endgültigen Analyse. Die Normalgleichungsmatrix wird nur einmal invertiert. Sequentielle Algorithmen, die das Zufügen bzw. Weglassen von Beobachtungen oder Positionen ermöglichen, werden eingesetzt (z.B. ***CLINE***, 1964). Erst wenn alle Messungen durchgeführt sind, können die abschließenden Auswertungen endgültige Aussagen zulassen.

4. HARD- UND SOFTWARE IM ZUSAMMENSPIEL

Die heute verfügbare Hardware gibt dem Bearbeiter ein Repertoire an die Hand, das seine Aufgabe erleichtert, unterstützt und komfortabel macht. Interessant sind die Möglichkeiten, die Digitizer und graphisches Terminal mit Hardcopy bieten. In BILD 4 ist als Beispiel die schrittweise Verfeinerung des Moduls "Anfangsdesign" dargestellt.

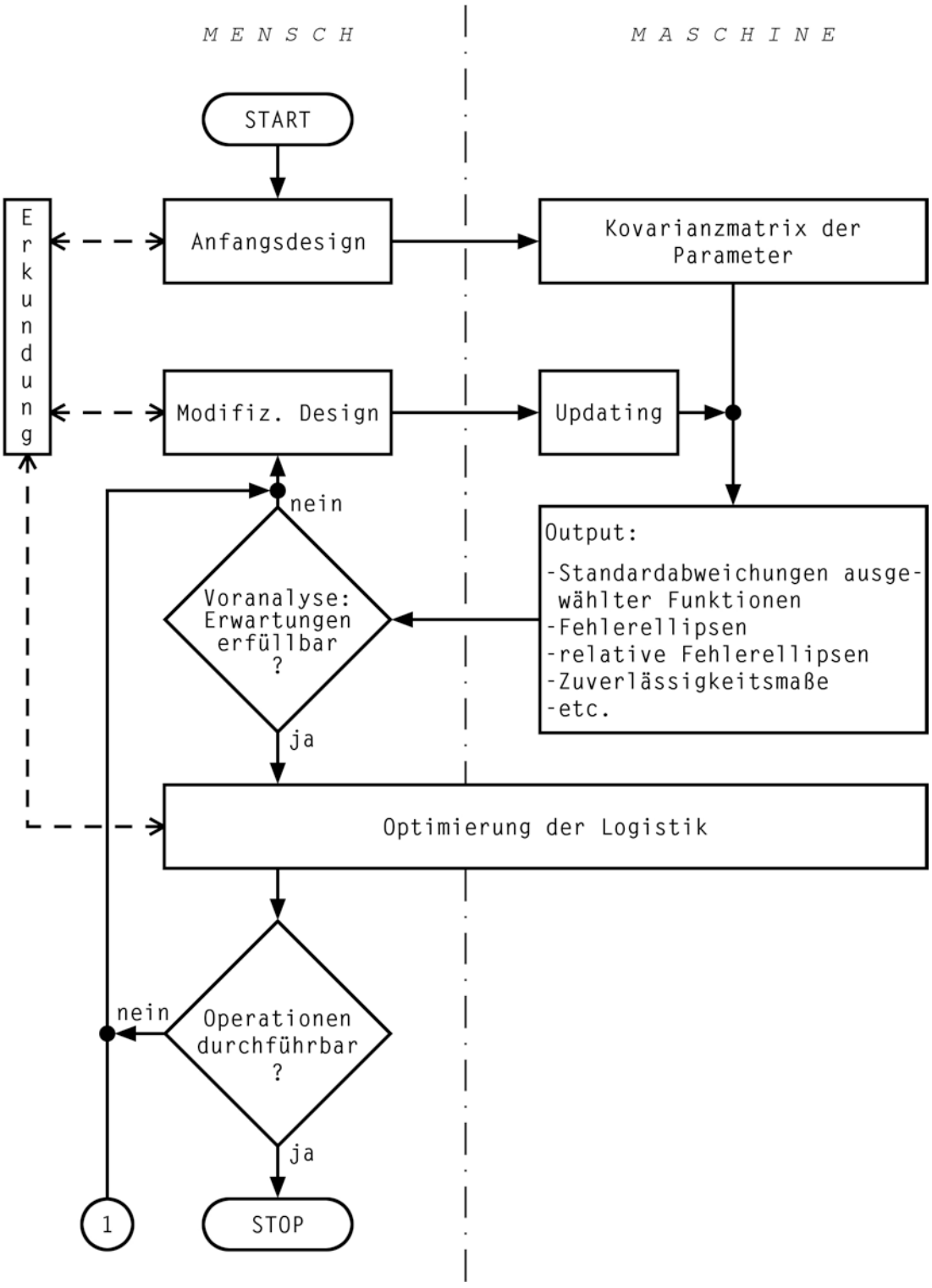


Bild 2 Phase II

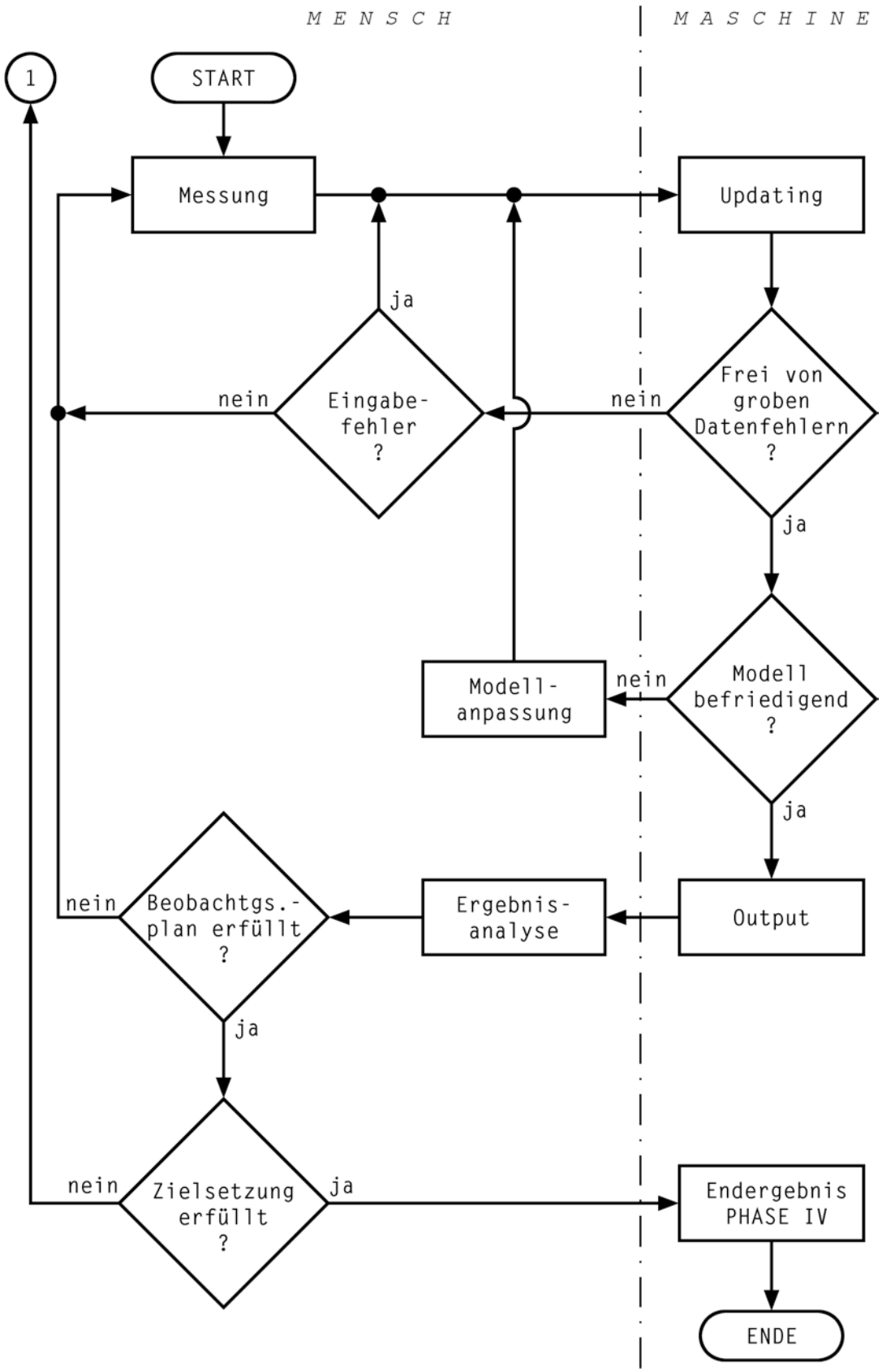


Bild 3 Phase III

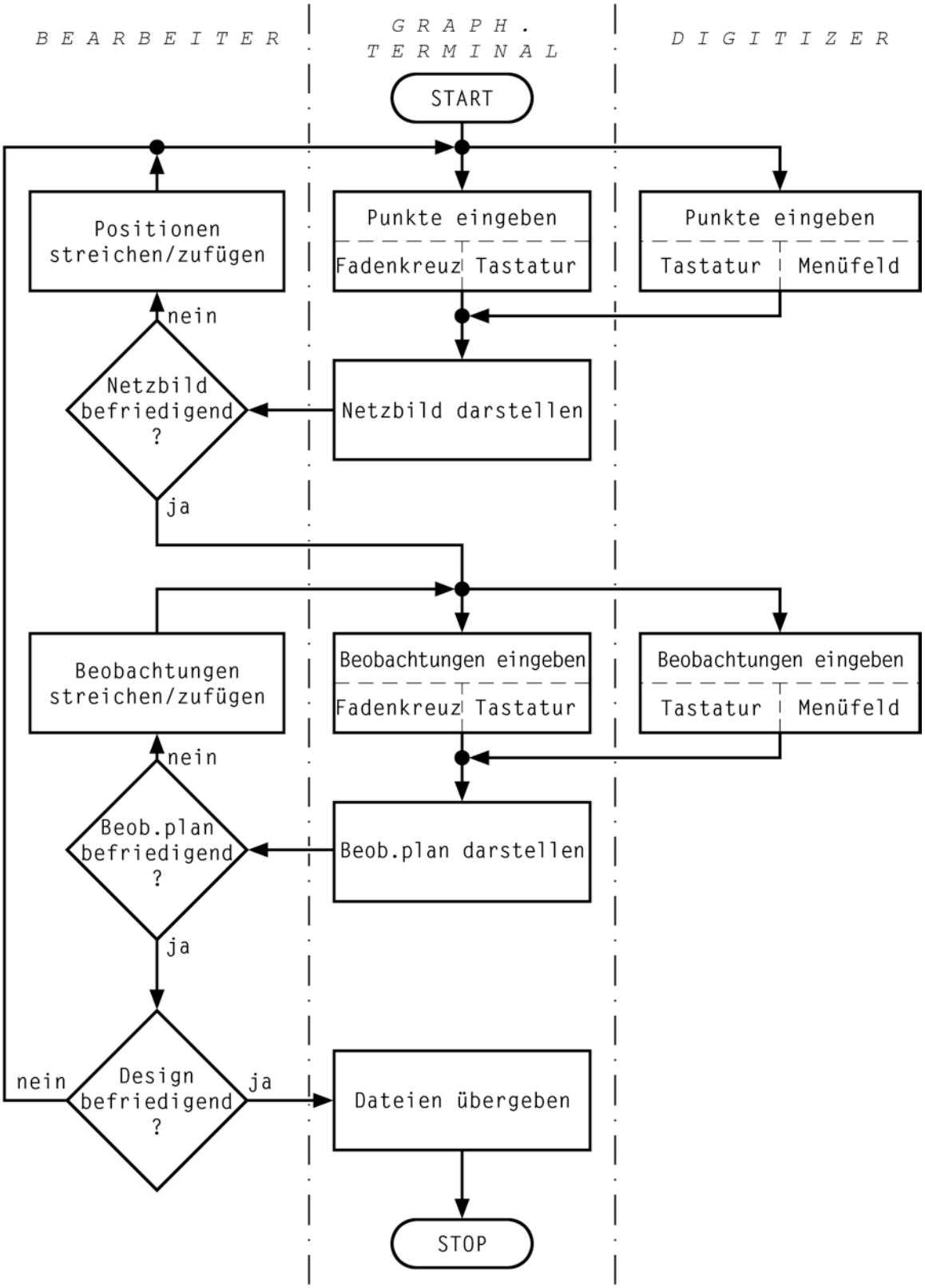


Bild 4 Schrittweise Verfeinerung des Moduls "Anfangsdesign"

Verfügbare Punktinformationen, z.B. Näherungskoordinaten aus früheren Ausgleichungen, können über die Tastatur eingegeben werden. Nach der Darstellung auf dem Bildschirm können mit dem Fadenkreuz weitere Punkte definiert und daraus Näherungskoordinaten bestimmt werden. Eine Alternative ist die Beschaffung von Näherungskoordinaten aus der Karte mit Hilfe des Digitizers. Ein Menüfeld macht die Eingabe zusätzlicher Informationen möglich. Die geplanten Beobachtungen können ebenfalls zahlenmäßig oder mit Hilfe des Fadenkreuzes oder Digitizers eingegeben werden.

6. SCHLUSSBEMERKUNGEN

Nach dem Prinzip der strukturierten Programmierung ist die schrittweise Verfeinerung der Module weiterzubetreiben, so daß ein benutzerfreundliches, anpassungsfähiges und wartungsfreundliches Programmsystem zu erwarten ist.

LITERATUR

CLINE, R.E.: ***Representations for the generalized inverse of a partitioned matrix***. Journal SIAM, Vol. 12, No. 3, 588-600, 1964

CROSS, P.A.: ***Computer aided design of geodetic networks***. Report of the IAG Special Study Group 1.59, Hamburg 1983

GRAFAREND, E., u.a.: ***Optimierung geodätischer Meßoperationen***. Wichmann, Karlsruhe 1979

MEPHAM, M.P.A., und KRAKIWSKY, E.J.: ***Interactive network design and analysis***. DGK, Reihe B, Heft 258/III, 90-96, München 1982

VAN MIERLO, J.: ***Second order design: precision and reliability aspects***. Allgemeine Vermessungsnachrichten 3, 95-101, 1981

NINKOV, T., und SCHMITT, G.: ***Eine Methode der Gewichtsoptimierung in geodätischen Netzen***. Allgemeine Vermessungsnachrichten 6, 216-222, 1983

SCHAFFRIN, B.: ***Some proposals concerning the diagonal second order design of geodetic networks***. Manuscripta Geodaetica 6, 303-326, 1981

STOLIKER, P.C., und ANDERSON, E.G.: ***Optimization of observing logistics in geodetic networks***. DGK, Reihe B, Heft 258/III, 22-29, München 1982

INVARIANZEIGENSCHAFTEN VON STRAINPARAMETERN

von Walter WELSCH

SUMMARY

The deformations of the earth's crust and other bodies can be investigated and described by the distortions of a geodetic monitoring network with the help of the analysis of infinitesimal strains. Doing this it is frequently desirable to transform the coordinate system of the network or to perform datum changes of the point displacements. Then the question of invariance and estimability of the strain parameters with respect to those transformations arises.

It is briefly described how strain parameters are derived from point displacements (x-method) and from distortions of distances and angles (l-method). On this basis the investigations are carried out. They result in the ascertainment that due to similarity transformations of the coordinate system coordinate related quantities, i.e. the strain tensor and pure shear, are likewise projected according to similarity, while object related quantities, i.e. dilatation, maximum shear strain and principal strains, remain invariant. Therefore, these coordinate system independent parameters are estimable in the sense of the theory of estimation. Furthermore, as a rule all parameters are invariant with respect to datum transformations of the point displacements.

ZUSAMMENFASSUNG

Bei der Untersuchung und Beschreibung von Deformationen, insbesondere auch im Bereich der rezenten Krustenbewegungen, mit Hilfe von Techniken der infinitesimalen Strainanalyse geodätischer Überwachungsnetze ist es häufig erwünscht, Transformationen des Koordinatensystems der Netze und Datumstransformationen der Koordinatenänderungen der Vergleichspunkte durchzuführen. Dabei erhebt sich die Frage der Invarianz und Schätzbarkeit der strainbeschreibenden Parameter gegenüber diesen Operationen.

Nach kurzer Darstellung der Ableitung der Strainparameter auf der Grundlage der Veränderungen der Koordinaten der Vergleichspunkte (x-Methode) und der Verzerrungen vergleichbarer Strecken und Winkel (l-Methode) werden die Untersuchungen durchgeführt. Als Ergebnis wird festgestellt, daß nach Transformationen des Koordinatensystems koordinatenbezogene Größen (Straintensor, reine Scherung) ähnlich abgebildet werden, während sich körperbezogene Grössen (dilatation, maximale Scherdehnung, Strainellipse) invariant verhalten. Diese sind somit im Sinne der Schätztheorie als vom Koordinatensystem unabhängige schätzbare Größen anzusehen. Gegenüber Datumstransformationen der Koordinatenänderungen verhalten sich im Regelfall alle Größen invariant.

1. STRAINPARAMETER NACH DER x-METHODE

Bei der Aufgabe, Deformationen eines Körpers zu beschreiben, geht man im allgemeinen von der Annahme linearer Beziehungen zwischen den Punkten, die den Körper vor und nach der Deformationen diskretisieren, aus (homogener Strain). Die Beziehungen vereinfachen sich noch bei kleinen Verformungen (infinitesimaler Strain).

Die grundlegenden Annahmen sind in der Transformationsvorschrift

$$\boldsymbol{x}' = \mathcal{F}\boldsymbol{x} + \boldsymbol{t} \tag{1-1}$$

formuliert (***WELSCH***, 1982). Dabei sind

$\boldsymbol{x}^{\top} = |\, x \;\; y \,|$, $\boldsymbol{x}'^{\top} = |\, x' \;\; y' \,|$ — Koordinatenvektoren vor bzw. nach der Deformation

$\boldsymbol{t}^{\top} = |\, t_x \; t_y \,|$ — Translationsvektor

$$\mathcal{F} = \begin{vmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{vmatrix}$$ — Rotations- und Verzerrungsmatrix (Deformationsmatrix) .

Für manche Zwecke ist es vorteilhaft, den Verschiebungsvektor $\boldsymbol{u}$ einzuführen

$$\boldsymbol{u} = \boldsymbol{x}' - \boldsymbol{x} = \left(\mathcal{F} - \mathcal{I}\right)\boldsymbol{x} + \boldsymbol{t} = \mathrm{d}\mathcal{F}\boldsymbol{x} + \boldsymbol{t} \tag{1-2}$$

($\mathcal{I}$ Einheitsmatrix).

Bei infinitesimalem Strain kann der Verschiebungsgradient $\mathrm{d}\mathcal{F}$ in eine Rotationsmatrix $\mathrm{d}\mathcal{R}$ mit kleinem Rotationswinkel ω und in eine symmetrische Verzerrungsmatrix $\mathcal{E}$ (Straintensor) aufgespalten werden:

$$\mathrm{d}\mathcal{F} = \mathrm{d}\mathcal{R} + \mathcal{E} \tag{1-3}$$

mit

$$\mathrm{d}\mathcal{R} = \begin{vmatrix} 0 & \omega \\ -\omega & 0 \end{vmatrix}, \quad \mathcal{E} = \begin{vmatrix} e_{xx} & e_{xy} \\ e_{xy} & e_{yy} \end{vmatrix} .$$

Die explizite Darstellung von (1-2) lautet damit

$$\begin{aligned} u_x &= e_{xx}\cdot x + e_{xy}\cdot y \qquad + \omega y + t_x \\ u_y &= \qquad e_{xy}\cdot x + e_{yy}\cdot y - \omega x \quad + t_y \end{aligned} \tag{1-4}$$

oder

$$\boldsymbol{u} = \mathcal{H}_1 \cdot \boldsymbol{p}_1 \tag{1-5}$$

mit

$$\mathcal{H}_1 = \begin{vmatrix} x & y & 0 & y & 1 & 0 \\ 0 & x & y & -x & 0 & 1 \end{vmatrix}$$

$$\boldsymbol{p}_1^T = | e_{xx} \;\; e_{xy} \;\; e_{yy} \;\; \omega \;\; t_x \;\; t_y | \; .$$

Die Transformation (1-5) stellt die Beziehung zwischen den Koordinatenänderungen $\boldsymbol{u}$ und den Transformationsparametern $\boldsymbol{p}_1$ dar (x-Methode). Dabei beschreiben die Parameter

$$\boldsymbol{e}_x^T = | e_{xx} \;\; e_{xy} \;\; e_{yy} | \tag{1-6}$$

die Verformung des Körpers, deren Ermittlung das eigentliche Anliegen einer Strainanalyse ist. Die Rotations- und Translationsparameter sind von untergeordneter Bedeutung, da sie eine den Körper nicht deformierende Starrkörperbewegung betreffen.

2. STRAINPARAMETER NACH DER l-METHODE

Außer aus den Änderungen der Koordinaten vergleichbarer Punkte können die wesentlichen Parameter $\boldsymbol{e}_x$ auch aus den Verzerrungen der die Punkte verbindenden Strecken und Winkel (l-Methode) gewonnen werden (***WELSCH***, 1983).

Dabei gilt für die Verzerrung dl allgemein

$$dl = \mathcal{L}\, \boldsymbol{u} \quad . \tag{2-1}$$

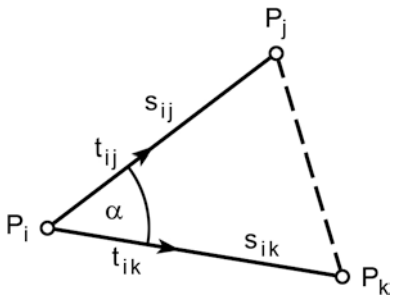


Fig. 1 Zur Strecken- bzw. Winkelverzerrung

Als Maß für Streckenverzerrungen dient die lineare Extension e, für die gilt

$$dl = e_{ij} = \frac{s'_{ij} - s_{ij}}{s_{ij}} \tag{2-2}$$

(s' verzerrte, s ursprüngliche Strecke),

$$\mathcal{L} = \mathcal{L}_e = \frac{1}{s_{ij}} \left| -\cos t_{ij} \quad -\sin t_{ij} \quad \cos t_{ij} \quad \sin t_{ij} \right| , \tag{2-3}$$

(t_{ij} Richtungswinkel)

$$\boldsymbol{u}^T = \left| u_{x_i} \quad u_{y_i} \quad u_{x_j} \quad u_{y_j} \right| . \tag{2-4}$$

Für Winkelverzerrungen

$$dl = g_{jik} = \alpha' - \alpha = \left(t'_{ik} - t'_{ij}\right) - \left(t_{ik} - t_{ij}\right) \tag{2-5}$$

(α' verzerrter, α ursprünglicher Winkel)

ist

$$\mathcal{L} = \mathcal{L}_g = \left| -\frac{\sin t_{ij}}{s_{ij}} + \frac{\sin t_{ik}}{s_{ik}} \;\vdots\; \frac{\cos t_{ij}}{s_{ij}} - \frac{\cos t_{ik}}{s_{ik}} \;\vdots\; \frac{\sin t_{ij}}{s_{ij}} \;\vdots\; -\frac{\cos t_{ij}}{s_{ij}} \;\vdots\; -\frac{\sin t_{ik}}{s_{ik}} \;\vdots\; \frac{\cos t_{ik}}{s_{ik}} \right| , \tag{2-6}$$

$$\boldsymbol{u}^T = \left| u_{x_i} \quad u_{y_i} \quad u_{x_j} \quad u_{y_j} \quad u_{x_k} \quad u_{y_k} \right| . \tag{2-7}$$

Die funktionale Beziehung zwischen der linearen Extension e und den Straintensor-Komponenten $\boldsymbol{e}_x$ ist gegeben durch:

$$e_{ij} = e_{xx}\cos^2 t_{ij} + e_{xy}\sin 2t_{ij} + e_{yy}\sin^2 t_{ij} \quad , \tag{2-8}$$

die der Winkelverzerrung g und $\boldsymbol{e}_x$ durch

$$g_{jik} = e_{xy}\left(\cos 2t_{ik} - \cos 2t_{ij}\right) + \frac{1}{2}\left(e_{yy} - e_{xx}\right)\left(\sin 2t_{ik} - \sin 2t_{ij}\right) \; . \tag{2-9}$$

3. ÄQUIVALENZ DER METHODEN

Damit stehen die Grundlagen zur Untersuchung der Invarianzeigenschaften der Strainparameter $\boldsymbol{e}_x$ bereit, wobei die x-Methode nach (1-5) und die l-Methode nach (2-1), (2-8) und (2-9) einander äquivalent sind. Hiervon kann ein Vergleich der Methoden leicht überzeugen.

Wird (1-5) in (2-1) eingesetzt, ergibt sich

$$d\boldsymbol{\ell} = \boldsymbol{\mathcal{L}} \cdot \boldsymbol{\mathcal{H}}_1 \cdot \boldsymbol{p}_1 \quad . \tag{3-1}$$

Für $\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{L}}_e$ nach (2-3) ergibt einfaches Nachrechnen

$$\boldsymbol{\mathcal{L}}_e \cdot \boldsymbol{\mathcal{H}}_1 = \left| \cos^2 t_{ij} \quad \sin 2t_{ij} \quad \sin^2 t_{ij} \quad 0 \quad 0 \quad 0 \right| \quad . \tag{3-2}$$

In Verbindung mit $\boldsymbol{p}_1$ wird damit (3-2) unmittelbar in (2-8) übergeführt.

Für $\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{L}}_g$ nach (2-6) ist in analoger Weise

$$\boldsymbol{\mathcal{L}}_g \cdot \boldsymbol{\mathcal{H}}_1 = \left| -\frac{1}{2}\left(\sin 2t_{ij} - \sin 2t_{ij}\right) \;\vdots\; \cos 2t_{ik} - \cos 2t_{ij} \;\vdots\; \frac{1}{2}\left(\sin 2t_{ij} - \sin 2t_{ij}\right) \;\vdots\; 0 \;\vdots\; 0 \;\vdots\; 0 \right| \; , \tag{3-3}$$

was in Verbindung mit $\boldsymbol{p}_1$ der Beziehung (2-9) entspricht.

Auf die Äquivalenz der Methoden hat schon ***BRUNNER et al.*** (1981) hingewiesen.

4. WEITERE STRAINPARAMETER

Aus den Komponenten des Straintensors $\boldsymbol{e}_x$ können weitere Strainparameter

$$\boldsymbol{p}_2^T = | \gamma_1 \quad \gamma_2 \quad \Delta | \tag{4-1}$$

abgeleitet werden. Es ist

$$\begin{aligned} \gamma_1 &= e_{yy} - e_{xx} && \text{reine Scherung (pure shear)} \\ \gamma_2 &= 2e_{xy} && \text{"Ingenieur-Scherung" (engineering shear)} \\ \Delta &= e_{xx} + e_{yy} && \text{Ausdehnung (dilatation) .} \end{aligned} \tag{4-2}$$

Mit

$$\begin{vmatrix} \gamma_1 \\ \gamma_2 \\ \Delta \end{vmatrix} = \begin{vmatrix} -1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} e_{xx} \\ e_{xy} \\ e_{yy} \end{vmatrix} \tag{4-2a}$$

kann (4-2) zusammengefaßt werden zu

$$\boldsymbol{p}_2 = \mathcal{H}_2 \, \boldsymbol{e}_x \quad . \tag{4-2b}$$

Des weiteren werden die nicht-linearen Beziehungen

$$\begin{aligned} \gamma &= \left(\gamma_1^2 + \gamma_2^2\right)^{1/2} && \text{maximale Scherdehnung (maximum shear strain)} \\ e_1 &= \frac{1}{2}\left(\Delta + \gamma\right) && \text{größte Hauptverzerrung (maximum principal strain)} \\ e_2 &= \frac{1}{2}\left(\Delta - \gamma\right) && \text{kleinste Hauptverzerrung (minimum principal strain)} \\ \Theta &= \frac{1}{2}\arctan\left(-\gamma_2/\gamma_1\right) && \text{Richtungswinkel von } e_1 \\ \Psi &= \Theta + 50\,\text{gon} && \text{Richtungswinkel von } \gamma \end{aligned} \tag{4-3}$$

häufig verwendet. Sie seien in

$$\boldsymbol{p}_3^T = | \gamma \quad e_1 \quad e_2 \quad \Theta \quad \Psi | \tag{4-4}$$

zusammengefaßt.

5. AUSWIRKUNG VON TRANSFORMATIONEN DES KOORDINATENSYSTEMS

Die Koordinaten der Punkte eines geodätischen Überwachungsnetzes sind häufig in einem System gegeben, das bezüglich des zu überwachenden Objekts eine beliebige Orientierung aufweist. In manchen Fällen ist es erwünscht, dem Koordinatensystem eine bestimmte Ausrichtung, z.B. parallel zu einer tektonischen Verwerfung, zu geben. Im folgenden soll untersucht werden, auf welche Weise eine Ähnlichkeitstransformation, vornehmlich eine Drehung, die verschiedenen Strainparameter beeinflußt, mit dem Ziel, Invarianzeigenschaften gewisser Parameter zu ermitteln.

Für den Fall der Ermittlung der Strainparameter e_x sowie der Starrkörperrotation ω aus Koordinatenänderungen hat ***DERMANIS*** (1981) eingehende Untersuchungen gemacht.

Im folgenden sollen die Invarianzeigenschaften auf der Grundlage der l-Methode aufgezeigt werden.

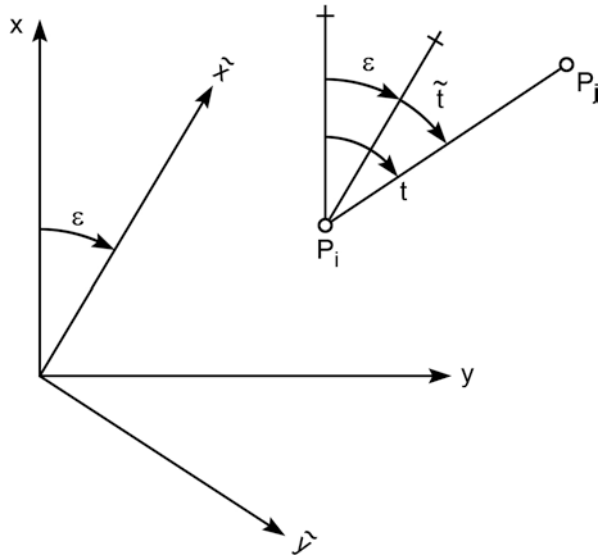


Fig. 2 Drehung des Koordinatensystems

Zur Grundlage der Betrachtungen wird die Beziehung (2-8) gemacht

$$e = e_{xx} \cos^2 t + e_{xy} \sin 2t + e_{yy} \sin^2 t \ . \qquad (2\text{-}8)$$

Hierin werden durch eine sinnvollerweise den Maßstab unverändert belassende Ähnlichkeitstransformation der Richtungswinkel t und, da die lineare Extension e von Translationen und Rotationen unbeeinflußt bleibt, in der Folge die Strainparameter $\boldsymbol{e}_x$ verändert. Die neue Beziehung lautet

$$e = \tilde{e}_{xx}\cos^2\tilde{t} + \tilde{e}_{xy}\sin 2\tilde{t} + \tilde{e}_{yy}\sin^2\tilde{t} \ . \tag{5-1}$$

Nach Ersetzen des neuen Richtungswinkels t durch

$$\tilde{t} = t - \varepsilon \ , \tag{5-2}$$

nach Anwendung trigonometrischer Rechenregeln und Umstellen wird daraus

$$\begin{aligned} e = {} & \cos^2 t\left(\tilde{e}_{xx}\cos^2\varepsilon - \tilde{e}_{xy}\sin 2\varepsilon + \tilde{e}_{yy}\sin^2\varepsilon\right) \\ & + \sin 2t\left(\frac{1}{2}\tilde{e}_{xx}\sin 2\varepsilon + \tilde{e}_{xy}\cos 2\varepsilon - \frac{1}{2}\tilde{e}_{yy}\sin 2\varepsilon\right) \\ & + \sin^2 t\left(\tilde{e}_{xx}\sin^2\varepsilon + \tilde{e}_{xy}\sin 2\varepsilon + \tilde{e}_{yy}\cos^2\varepsilon\right) \ . \end{aligned}$$

Ein Koeffizientenvergleich mit (2-8) führt sofort zu der Beziehung

$$\begin{vmatrix} e_{xx} \\ e_{xy} \\ e_{yy} \end{vmatrix} = \begin{vmatrix} \cos^2\varepsilon & -\sin 2\varepsilon & \sin^2\varepsilon \\ \frac{1}{2}\sin 2\varepsilon & \cos 2\varepsilon & -\frac{1}{2}\sin 2\varepsilon \\ \sin^2\varepsilon & \sin 2\varepsilon & \cos^2\varepsilon \end{vmatrix} \cdot \begin{vmatrix} \tilde{e}_{xx} \\ \tilde{e}_{xy} \\ \tilde{e}_{yy} \end{vmatrix} \tag{5-3}$$

oder

$$\boldsymbol{e}_x = \boldsymbol{\mathcal{T}} \cdot \tilde{\boldsymbol{e}}_x \ . \tag{5-3a}$$

Die inverse Beziehung

$$\tilde{\boldsymbol{e}}_x = \boldsymbol{\mathcal{T}}^{-1}\,\boldsymbol{e}_x \tag{5-4a}$$

lautet infolge einfacher Umkehrung der Rotationsrichtung

$$\begin{vmatrix} \tilde{e}_{xx} \\ \tilde{e}_{xy} \\ \tilde{e}_{yy} \end{vmatrix} = \begin{vmatrix} \cos^2\varepsilon & \sin 2\varepsilon & \sin^2\varepsilon \\ -\frac{1}{2}\sin 2\varepsilon & \cos 2\varepsilon & \frac{1}{2}\sin 2\varepsilon \\ \sin^2\varepsilon & -\sin 2\varepsilon & \cos^2\varepsilon \end{vmatrix} \cdot \begin{vmatrix} e_{xx} \\ e_{xy} \\ e_{yy} \end{vmatrix} \ . \tag{5-4}$$

Ersetzt man weiterhin die Scherkomponente e_{xy} bzw. $\tilde{e}_{xy}$ gemäß (4-2) durch die "Ingenieur-Scherung" $\gamma_2 = e_{xy} + e_{xy}$ bzw. $\tilde{\gamma}_2 = \tilde{e}_{xy} + \tilde{e}_{xy}$ und die Funktionen der doppelten Winkel durch die Summe von Funktionen, so läßt sich (5-4) nach Umstellung auch darstellen als (***DERMANIS***, 1981)

$$\tilde{\mathcal{E}} = \mathcal{R}_\varepsilon \cdot \mathcal{E} \cdot \mathcal{R}_\varepsilon^\top \tag{5-5}$$

mit

$$\mathcal{R}_\varepsilon = \begin{vmatrix} \cos\varepsilon & \sin\varepsilon \\ -\sin\varepsilon & \cos\varepsilon \end{vmatrix} .$$

Untersucht man das Verhalten der Parameter γ_1 und γ_2 , so ergibt sich mit (4-2) aus (5-3) unmittelbar

$$\begin{vmatrix} \gamma_1 \\ \gamma_2 \end{vmatrix} = \begin{vmatrix} \cos 2\varepsilon & \sin 2\varepsilon \\ -\sin 2\varepsilon & \cos 2\varepsilon \end{vmatrix} \cdot \begin{vmatrix} \tilde{\gamma}_1 \\ \tilde{\gamma}_2 \end{vmatrix} \tag{5-6}$$

oder

$$\begin{vmatrix} \gamma_1 \\ \gamma_2 \end{vmatrix} = \mathrm{R}_{2\varepsilon} \begin{vmatrix} \tilde{\gamma}_1 \\ \tilde{\gamma}_2 \end{vmatrix} . \tag{5-6a}$$

Die inverse Beziehung ist infolge der Orthogonalität der Rotationsmatrix $\mathrm{R}_{2\varepsilon}$

$$\begin{vmatrix} \tilde{\gamma}_1 \\ \tilde{\gamma}_2 \end{vmatrix} = \mathcal{R}_{2\varepsilon}^\top \begin{vmatrix} \gamma_1 \\ \gamma_2 \end{vmatrix} . \tag{5-7}$$

Analoges Vorgehen ergibt ohne weiteres

$$\begin{aligned} \Delta &= e_{xx} + e_{yy} = \tilde{e}_{xx} + \tilde{e}_{yy} \\ \gamma &= \left(\gamma_1^2 + \gamma_2^2\right)^{1/2} = \left(\tilde{\gamma}_1^2 + \tilde{\gamma}_2^2\right)^{1/2} \\ e_1 &= \tilde{e}_1 \\ e_2 &= \tilde{e}_2 \quad . \end{aligned} \tag{5-8}$$

Der neue Richtungswinkel $\tilde{\Theta}$ der größten Hauptverzerrung $\tilde{e}_1$ berechnet sich aus

$$\tan 2\Theta = \frac{-\gamma_2}{\gamma_1}$$

$$\tan 2\tilde{\Theta} = \frac{\sin 2\tilde{\Theta}}{\cos 2\tilde{\Theta}} = \frac{-\tilde{\gamma}_2}{\tilde{\gamma}_1} \qquad (4\text{-}3)$$

$$\tilde{\gamma}_1 = \cos 2\tilde{\Theta} \;, \quad \tilde{\gamma}_2 = -\sin 2\tilde{\Theta}$$

$$\tan 2\Theta = \frac{\sin 2\varepsilon\,\tilde{\gamma}_1 - \cos 2\varepsilon\,\tilde{\gamma}_2}{\cos 2\varepsilon\,\tilde{\gamma}_1 + \sin 2\varepsilon\,\tilde{\gamma}_2} = \frac{\sin\left(2\varepsilon + 2\tilde{\Theta}\right)}{\cos\left(2\varepsilon + 2\tilde{\Theta}\right)} \;.$$

Daraus folgt

$$\tilde{\Theta} = \Theta - \varepsilon \;. \qquad (5\text{-}9)$$

Mit den Formeln (5-5) - (5-8) kann die Auswirkung einer Koordinatentransformation auf alle Strainparameter berechnet werden. Es wird deutlich, daß die Komponenten des Straintensors $\boldsymbol{\varepsilon}$ bei Umorientierung des Koordinatensystems den Gesetzen der Kongruenztransformation von Matrizen (***ZURMÜHL***, 1964, S. 55 ff.) unterliegen. Die Komponenten γ_1 und γ_2 erfahren eine Rotation um den doppelten Drehwinkel ε, während sich die übrigen Elemente, die Dilatation Δ, die maximale Scherdehnung γ und die Hauptverzerrungen e_1 und e_2 als Halbachsen der Strainellipse invariant verhalten. Lediglich die Orientierung Θ der Strainellipse folgt einer Drehung des Koordinatensystems so, daß die Lage der Ellipse in Bezug auf das deformierte Objekt erhalten bleibt.

6. AUSWIRKUNG VON DATUMSTRANSFORMATIONEN

Das geodätische Datum wird nicht durch Beobachtungen ermittelt, sondern willkürlich festgelegt. Gründe der Zweckmäßigkeit oder der Anschaulichkeit führen zur Wahl der einen oder anderen Festlegung, ohne damit die im Beobachtungsmaterial enthaltene Relativinformation zu beeinflussen.

In Figur 3 wurde die ursprüngliche Figur eines Quadrates verzerrt. Das Erscheinungsbild der Verschiebungsvektoren $\boldsymbol{u}$ und $\boldsymbol{u}_1$ ist verschieden und hängt nur von der Wahl des geodätischen Datums ab. Bei der Untersuchung der verzerrten Figur und ihrer Beschreibung durch Strainparameter erhebt sich die Frage, in welchem Maße die die Verformung beschreibenden Größen von der

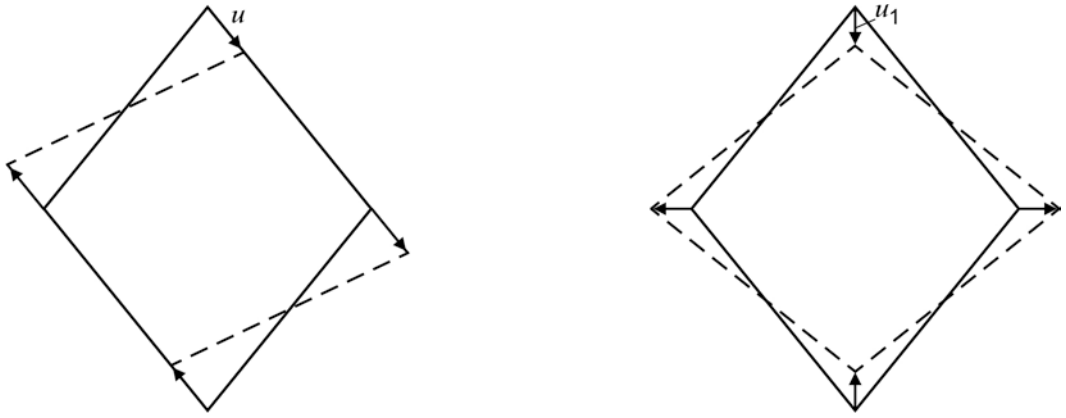


Fig. 3 Auswirkung von Datumsänderungen

Datumswahl abhängig sind. Dies soll im folgenden betrachtet werden.

Es sei $\boldsymbol{u}$ der Vektor der Koordinatenänderungen, bezogen auf ein spezielles Datum. Durch Ähnlichkeitstransformation (S-Transformation) läßt sich daraus eine andere Möglichkeit

$$\boldsymbol{u}_1 = \boldsymbol{u} + \mathcal{H} \cdot \boldsymbol{p} \qquad (6\text{-}1)$$

mit

$$\mathcal{H} = \begin{vmatrix} y & x & 1 & 0 \\ -x & y & 0 & 1 \end{vmatrix}$$

$$\boldsymbol{p}^T = | \; \varepsilon \quad m \quad p_x \quad p_y \; |$$

in Abhängigkeit von der Wahl der Parameter $\boldsymbol{p}$ angeben. So wurde etwa der Vektor $\boldsymbol{u}_1$ in Figur 3 durch die Bedingung $\boldsymbol{u}_1^T \boldsymbol{u}_1 = \min.$ ("innere Lösung") erzeugt.

Die Untersuchungen können nach der x-Methode und nach der l-Methode ausgeführt werden.

Grundlage für die Verwendung der x-Methode ist die Beziehung

$$\boldsymbol{u} = \mathcal{H}_1 \cdot \boldsymbol{p}_1 \qquad (1\text{-}5)$$

bzw. die inverse Beziehung

$$\boldsymbol{p}_1 = \mathcal{H}_1^{-1} \cdot \boldsymbol{u} \quad . \tag{6-2}$$

Wird der Vektor $\boldsymbol{u}$ nach (6-1) in ein anderes Datum transformiert, so geht (6-2) über in

$$\boldsymbol{p}_1^* = \mathcal{H}_1^{-1}(\boldsymbol{u} + \mathcal{H}\boldsymbol{p}) = \mathcal{H}_1^{-1}\boldsymbol{u} + \mathcal{H}_1^{-1}\mathcal{H}\boldsymbol{p} \quad . \tag{6-3}$$

$\boldsymbol{p}_1$ bliebe von dieser Transformation nur dann unbeeinflußt, wenn

$$\mathcal{H}_1^{-1} \cdot \mathcal{H} \stackrel{!}{=} 0 \tag{6-4}$$

gilt.

Nachrechnung (die Inverse $\mathcal{H}_1^{-1}$ ist in Anlage 1 explizit dargestellt) ergibt

$$\boldsymbol{p}_1^* = \begin{vmatrix} e_{xx}^* \\ e_{xy}^* \\ e_{yy}^* \\ \omega^* \\ t_x^* \\ t_y^* \end{vmatrix} = p_1 + \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} \varepsilon \\ m \\ p_x \\ p_y \end{vmatrix} \tag{6-5}$$

oder

$$\begin{aligned} e_{xx}^* &= e_{xx} + m \\ e_{xy}^* &= e_{xy} \\ e_{yy}^* &= e_{yy} + m \\ \omega^* &= \omega + \varepsilon \\ t_x^* &= t_x + p_x \\ t_y^* &= t_y + p_y \quad . \end{aligned} \tag{6-5a}$$

Überträgt man die Änderungen auf die weiteren Strainparameter, so ergibt sich

$$
\begin{aligned}
\gamma_1^* &= \gamma_1 \\
\gamma_2^* &= \gamma_2 \\
\Delta^* &= \Delta + 2\,m \\
\gamma^* &= \gamma \\
e_1^* &= \frac{1}{2}(\Delta + 2\,m + \gamma) = \frac{1}{2}\left(\Delta^* + \gamma\right) \\
e_2^* &= \frac{1}{2}(\Delta + 2\,m - \gamma) = \frac{1}{2}\left(\Delta^* - \gamma\right) \\
\Theta^* &= \Theta \\
\Psi^* &= \Psi \quad .
\end{aligned}
\tag{6-6}
$$

Zu den gleichen Resultaten gelangt man auf der Grundlage der l-Methode (***WELSCH*** und ***ZHANG***, 1983):

Schreibt man nämlich (2-8) in verkürzter Form

$$e = \mathcal{K}_1 \cdot e_x \quad , \tag{2-8}$$

so ergibt sich daraus (etwa für ein Dreieck der Figur 1)

$$e_x = \mathcal{K}_1^{-1}\, e \quad . \tag{6-7}$$

Nun ist aber, vgl. (2-1) und (2-3), auch

$$e = \mathcal{L}_e \cdot u \quad . \tag{6-8}$$

Transformiert man wiederum u gemäß (6-1) in ein anderes Datum, so geht (6-8) und damit (6-7) über in

$$e_x = \mathcal{K}_1^{-1}\mathcal{L}_e\,(u + \mathcal{H}p) = \mathcal{K}_1^{-1}\mathcal{L}_e\,u + \mathcal{K}_1^{-1}\mathcal{L}_e\,\mathcal{H}\cdot p \quad . \tag{6-9}$$

e_x bliebe nur dann unverändert, wenn

$$\mathcal{K}_1^{-1}\mathcal{L}_e\,\mathcal{H} \stackrel{!}{=} \mathcal{O} \tag{6-10}$$

als Bedingung für Datumsinvarianz erfüllt wurde.

Nachrechnung mit der in Anlage 2 explizit dargestellten Inversen $\mathcal{K}_1^{-1}$ ergibt

$$\mathcal{K}_1^{-1}\mathcal{L}_e\,\mathcal{H} = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix} . \qquad (6\text{-}11)$$

Hieraus ergibt sich dieselbe Schlußfolgerung für e_x^* wie in (6-5).

Macht man schließlich die Winkelverzerrung g gemäß (2-9) mit

$$\mathcal{g} = \mathcal{K}_2 \begin{vmatrix} e_{xy} \\ \gamma_1 \end{vmatrix} \qquad (2\text{-}9)$$

zur Grundlage der Betrachtungen, so ergibt sich mit

$$\mathcal{K}_2^{-1}\mathcal{L}_g\,\mathcal{H} = \mathcal{O} \qquad (6\text{-}12)$$

das entsprechende Ergebnis ($\mathcal{K}_2^{-1}$ ist in Anlage 3 aufgeführt), das schon aus den obigen Untersuchungen abgeleitet werden kann.

Als Schlußfolgerung kann man ziehen, daß sich nach (6-5) die durch die Datumstransformation ergebende Starrkörperbewegung (ε, p_x, p_y) unmittelbar auf die entsprechenden Parameter (ω, t_x, t_y), die sich aus der Deformationsanalyse nach der x-Methode ergeben, überträgt.

Der Maßstabsfaktor (der bei den Datumsdefekt der Vergleichsfiguren korrekt berücksichtigenden Vorausgleichungen nicht auftritt) beeinflußt die Parameter des normalen Strains e_x und e_y und die daraus abgeleiteten weiteren Werte der Dilatation Δ und der Hauptverzerrungen e_1 und e_2 .

Abgesehen hiervon ist nachgewiesen, daß Datumstransformationen der Koordinatenänderungen u der Vergleichspunkte die Strainparameter unbeeinflußt lassen. Die Strainparameter sind in dieser Beziehung Invariante.

DANKSAGUNG

Herzlicher Dank gebührt Herrn Dr.-Ing. K. Krack, der die Durchrechnung der Formeln (6-4), (6-11) und (6-12) in allgemeiner Darstellung mit Hilfe des Programms REDUCE 2 (***HEARN***, 1974) durchgeführt und damit auch die Darstellung der Anlagen 1 - 3 ermöglicht hat.

LITERATUR

BRUNNER, F.K., COLEMAN, R., HIRSCH, B., 1981, ***A comparison of computation methods for crustal strains from geodetic measurements***, Tectonophysics 71 (1981), pp. 281-298

DERMANIS, A., 1981, ***Geodetic estimability of crustal deformation parameters***, Quaterniones Geodaesiae 2 (1982) 2, pp. 159-169

HEARN, A.C., 1974, ***REDUCE 2 User's Manual***, University of Utah, Salt Lake City 1974

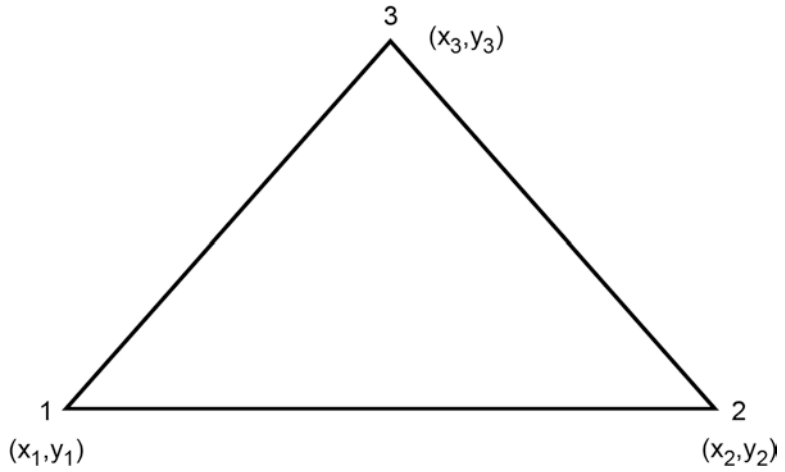
WELSCH, W., 1982, ***Zur Beschreibung homogenen Strains oder Einige Betrachtungen zur affinen Transformation***, Zeitschrift für Vermessungswesen 107 (1982) 5, S. 173-182

WELSCH, W., 1983, ***Finite element analysis of strain patterns from geodetic observations across a plate margin***, Tectonophysics 97 (1983), pp. 57-71

WELSCH, W., ZHANG, Y., 1983, ***Kongruenz und Affinität von geodätischen Netzen und Teilnetzen***, Journal of the Wuhan College of Geodesy, Photogrammetry and Cartography, in Druck

ZURMÜHL, R., 1964, ***Matrizen***, Springer-Verlag, Berlin-Göttingen-Heidelberg 1964

Anhang 1



$$\mathcal{H}_1 = \begin{vmatrix} x_1 & y_1 & 0 & y_1 & 1 & 0 \\ 0 & x_1 & y_1 & -x_1 & 0 & 1 \\ x_2 & y_2 & 0 & y_2 & 1 & 0 \\ 0 & x_2 & y_2 & -x_2 & 0 & 1 \\ x_3 & y_3 & 0 & y_3 & 1 & 0 \\ 0 & x_3 & y_3 & -x_3 & 0 & 1 \end{vmatrix}$$

$$H_1^{-1} = \frac{1}{\mathrm{Det}} \begin{vmatrix} h_{11} & h_{12} & \cdots & h_{16} \\ \vdots & & & \vdots \\ h_{51} & & & h_{56} \\ h_{61} & h_{62} & \cdots & h_{66} \end{vmatrix} ,$$

wobei die Symbole Det, h_{ij} für die folgenden Terme stehen:

$$DET = 2*(- X3^2*Y2^2 + 2*X3^2*Y2*Y1 - X3^2*Y1^2 + 2*X3*Y3*X2*Y2 - 2*X3*Y3*X2*Y1 - 2*X3*Y3*Y2*X1 + 2*X3*Y3*X1*Y1 - 2*X3*X2*Y2*Y1 + 2*X3*X2*Y1^2 + 2*X3*Y2^2*X1 - 2*X3*Y2*X1*Y1 - Y3^2*X2^2 + 2*Y3^2*X2*X1 - Y3^2*X1^2 + 2*Y3*X2^2*Y1 - 2*Y3*X2*Y2*X1 - 2*Y3*X2*X1*Y1 + 2*Y3*Y2*X1^2 - X2^2*Y1^2 + 2*X2*Y2*X1*Y1 - Y2^2*X1^2)$$

$$H(1.1) = 2*(- X3*Y3*Y2 + X3*Y3*Y1 + X3*Y2^2 - X3*Y2*Y1 + Y3^2*X2 - Y3^2*Y1 - Y3*X2*Y2 - Y3*X2*Y1 + 2*Y3*Y2*X1 + X2*Y2*Y1 - Y2^2*X1)$$

$$H(1.2) = 0$$

$$H(1.3) = 2*(X3*Y3*Y2 - X3*Y3*Y1 - X3*Y2*Y1 + X2*Y1^2 - Y3^2*X2 + Y3^2*X1 + 2*Y3*X2*Y1 - Y3*Y2*X1 - Y3*X1*Y1 - X2*Y1^2 + Y2*X1*Y1)$$

$$H(1.4) = 0$$

$$H(1.5) = 2*(- X3*Y2^2 + 2*X3*Y2*Y1 - X3*Y1^2 + Y3*X2*Y2 - Y3*X2*Y1 - Y3*Y2*X1 + Y3*X1*Y1 - X2*Y2*Y1 + X2*Y1^2 + Y2^2*X1 - Y2*X1*Y1)$$

$$H(1.6) = 0$$

$$H(2.1) = X3^2*Y2 - X3^2*Y1 - X3*Y3*X2 + X3*Y3*X1 - X3*X2*Y2 + 2*X3*X2*Y1 - X3*Y2*X1 + Y3*X2^2 - Y3*X2*X1 - X2^2*Y1 + X2*Y2*X1$$

$$H(2.2) = - X3*Y3*Y2 + X3*Y3*Y1 + X3*Y2^2 - X3*Y2*Y1 + Y3^2*X2 - Y3^2*X1 - Y3*X2*Y2 - Y3*X2*Y1 + 2*Y3*Y2*X1 + X2*Y2*Y1 - Y2^2*X1$$

$$H(2.3) = - X3^2*Y2 + X3^2*Y1 + X3*Y3*X2 - X3*Y3*X1 - X3*X2*Y1 + 2*X3*Y2*X1 - X3*X1*Y1 - Y3*X2*X1 + Y3*X1^2 + X2*X1*Y1 - Y2*X1^2$$

$$H(2.4) = X3*Y3*Y2 - X3*Y3*Y1 - X3*Y2*Y1 + X3*Y1^2 - Y3^3*X2 + Y3^2*X1 + 2*Y3*X2*Y1 - Y3*Y2*X1 - Y3*X1*Y1 - X2*Y1^2 + Y2*X1*Y1$$

$$H(2.5) = X3*X2*Y2 - X3*X2*Y1 - X3*Y2*X1 + X3*X1*Y1 - Y3*X2^2 + 2*Y3*X2*X1 - Y3*X1^2 + X2^2*Y1 - X2*Y2*X1 - X2*X1*Y1 + Y2*X1^2$$

$$H(2.6) = - X3*Y2^2 + 2*X3*Y2*Y1 - X3*Y1^2 + Y3*X2*Y2 - Y3*X2*Y1 - Y3*Y2*X1 + Y3*X1*Y1 - X1*Y2*Y1 + X2*Y1^2 + Y2^2*X1 - Y2*X1*Y1$$

$$H(3.1) = 0$$

$$H(3.2) = 2*(X3^2*Y2 - X3^2*Y1 - X3*Y3*X2 + X3*Y3*X1 - X3*X2*Y2 + 2*X3*X2*Y1 - X3*Y2*X1 + Y3*X2^2 - Y3*X2*X1 - X2^2*Y1 + X2*Y2*X1)$$

$$H(3.3) = 0$$

$$H(3.4) = 2*(- X3^2*Y2 + X3^2*Y1 + X3*Y3*X2 - X3*Y3*X1 - X3*X2*Y1 + 2*X3*Y2*X1 - X3*X1*Y1 - Y3*X2*X1 + Y3*X1^2 + X2*X1*Y1 - Y2*X1^2)$$

$$H(3.5) = 0$$

$$H(3.6) = 2*(X3*X2*Y2 - X3*X2*Y1 - X3*Y2*X1 + X3*X1*Y1 - Y3*X2^2 + 2*Y3*X2*X1 - Y3*X1^2 + X2^2*Y1 - X2*Y2*X1 - X2*X1*Y1 + Y2*X1^2)$$

$$H(4.1) = X3^2*Y2 - X3^2*Y1 - X3*Y3*X2 + X3*Y3*X1 - X3*X2*Y2 + 2*X3*X2*Y1 - X3*Y2*X1 - Y3*X2^2 - Y3*X2*X1 - X2^2*Y1 + X2*Y2*X1$$

$$H(4.2) = X3*Y3*Y2 - X3*Y3*Y1 - X3*Y2^2 + X3*Y2*Y1 - Y3^2*X2 + Y3^2*X1 + Y3*X2*Y2 + Y3*X2*Y1 - 2*Y3*Y2*X1 - X2*Y2*Y1 + Y2^2*X1$$

$$H(4.3) = - X3^2*Y2 + X3^2*Y1 + X3*Y3*X2 - X3*Y3*X1 - X3*X2*Y1 + 2*X3*Y2*X1 - X3*X1*Y1 - Y3*X2*X1 + Y3*X1^2 + X2*X1*Y1 - Y2*X1^2$$

$$H(4.4) = - X3*Y3*Y2 + X3*Y3*Y1 + X3*Y2*Y1 - X3*Y1^2 + Y3^2*X2 - Y3^2*X1 - 2*Y3*X2*Y1 + Y3*Y2*X1 + Y3*X1*Y1 + X2*Y1^2 - Y2*X1*Y1$$

$$H(4.5) = X3*X2*Y2 - X3*X2*Y1 - X3*Y2*X1 + X3*X1*Y1 - Y3*X2^2 + 2*Y3*X2*X1 - Y3*X1^2 + X2^2*Y1 - X2*Y2*X1 - X2*X1*Y1 + Y2*X1^2$$

$$H(4.6) = X3*Y2^2 - 2*X3*Y2*Y1 + X3*Y1^2 - Y3*X2*Y2 + Y3*X2*Y1 + Y3*Y2*X1 - Y3*X1*Y1 + X2*Y2*Y1 - X2*Y1^2 - Y2^2*X1 + Y2*X1*Y1$$

$$H(5.1) = 2*(- X3^2*Y2^2 + X3^2*Y2*Y1 + 2*X3*Y3*X2*Y2 - X3*Y3*X2*Y1 - X3*Y3*Y2*X1 - X3*Y2^2*X1 + Y3*X2^2*X1 + Y3*X2^2*Y1 - Y3*X2*Y2*X1)$$

$$H(5.2) = 0$$

$$H(5.3) = 2*(X3^2*Y2*Y1 - X3^2*Y1^2 - X3*Y3*X2*Y1 - X3*Y3*Y2*X1 + 2*X3*Y3*X1*Y1 + X3*X2*Y1^2 - X3*Y2*X1*Y1 + Y3*X2*X1 - Y3*X1^2 - Y3*X2*X1*Y1 + Y3*Y2*X1^2)$$

H (5.4) = 0

$$H(5.5) = 2*(- X3*X2*Y2*Y1 + X3*X2*Y1^2 + X3*Y2^2*X1 - X3*Y2*X1*Y1 + Y3*X2^2*Y1 - Y3*X2*Y2*X1 - Y3*X2*X1*Y1 + Y3*Y2*X1^2 - X2^2*Y1^2 + 2*X2*Y2*X1*Y1 - Y2^2*X1^2)$$

H (5.6) = 0

H (6.1) = 0

$$H(6.2) = 2*(- X3^2*Y2^2 + X3^2*Y2*Y1 + 2*X3*Y3*X2*Y2 - X3*Y3*X2*Y1 - X3*Y3*Y2*X1 - X3*X2*Y2*Y1 + X3*Y2^2*X1 - Y3^2*X2^2 + Y3^2*X2*X1 + Y3*X2^2*Y1 - Y3*X2*Y2*X1)$$

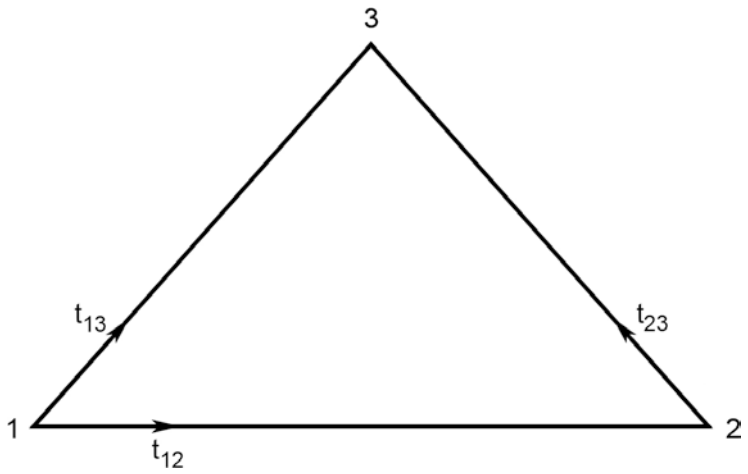
H (6.3) = 0

$$H(6.4) = 2*(X3^2*Y2*Y1 - X3^2*Y1^2 - X3*Y3*X2*Y1 - X3*Y3*Y2*X1 + 2*X3*Y3*X1*Y1 + X3*X2*Y1^2 - X3*Y2*X1*Y1 + Y3^2*X2*X1 - Y3^2*X1^2 - Y3*X2*X1*Y1 + Y3*Y2*X1^2)$$

H (6.5) = 0

$$H(6.6) = 2*(- X3*X2*Y2*Y1 + X3*X2*Y1^2 + X3*Y2^2*X1 - X3*Y2*X1*Y1 + Y3*X2^2*Y1 - Y3*X2*Y2*X1 - Y3*X2*X1*Y1 + Y3*Y2*X1^2 - X2^2*Y1^2 + 2*X2*Y2*X1*Y1 - Y2^2*X1^2)$$

Anhang 2



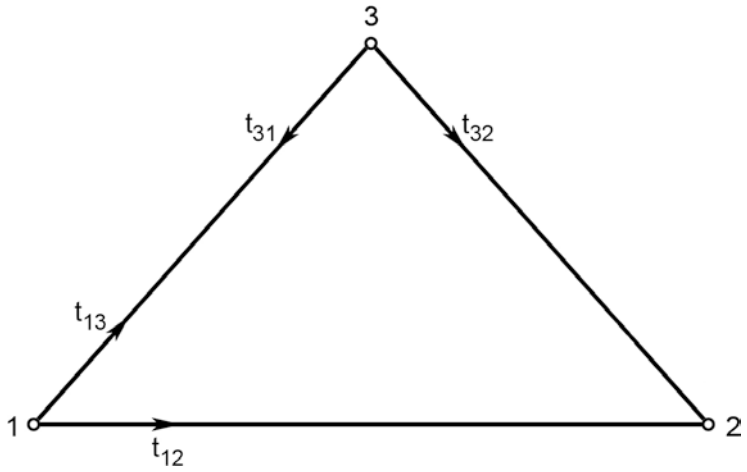
$$\mathcal{K}_1 = \begin{vmatrix} \cos^2 t_{12} & \sin 2t_{12} & \sin^2 t_{12} \\ \cos^2 t_{23} & \sin 2t_{23} & \sin^2 t_{23} \\ \cos^2 t_{13} & \sin 2t_{13} & \sin^2 t_{13} \end{vmatrix}$$

$$\mathcal{K}_1^{-1} = \frac{1}{\mathrm{Det}} \begin{vmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{vmatrix} ,$$

wobei die Symbole Det, k_{ij} für die folgenden Terme stehen:

```
                                2          2                          2          2
DET =  - SIN(2*T13)*SIN(T23) *COS(T12)  + SIN(2*T13)*COS(T23) *SIN(T12)  + SIN

                          2          2                          2          2
      (2*T23)*SIN(T13) *COS(T12)  - SIN(2*T23)*COS(T13) *SIN(T12)  - SIN(2*T12

                   2          2                          2          2
      )*SIN(T13) *COS(T23)  + SIN(2*T12)*COS(T13) *SIN(T23)

                                      2                          2
K  (1.1) =  - SIN(2*T13)*SIN(T23)  + SIN(2*T23)*Sin(T13)

                                   2                          2
K  (1.2) = SIN(2*T13)*SIN(T12)  - SIN(2*T12)*SIN(T13)

                                      2                          2
K  (1.3) =  - SIN(2*T23)*SIN(T12)  + SIN(2*T12)*SIN(T23)

                          2          2             2          2
K  (2.1) =  - SIN(T13) *COS(T23)  + COS(T13) *SIN(T23)

                       2          2             2          2
K  (2.2) = SIN(T13) *COS(T12)  - COS(T13) *SIN(T23)

                          2          2             2          2
K  (2.3) =  - SIN(T23) *COS(T12)  + COS(T23) *SIN(T12)

                                   2                          2
K  (3.1) = SIN(2*T13)*COS(T23)  - SIN(2*T23)*COS(T13)

                                      2                          2
K  (3.2) =  - SIN(2*T13)*COS(T12)  + SIN(2*T12)*COS(T13)

                                   2                          2
K  (3.3) = SIN(2*T23)*COS(T12)  - SIN(2*T12)*COS(T23)
```

Anhang 3



$$\mathcal{K}_2 = \begin{vmatrix} \cos 2t_{12} - \cos 2t_{13} & \frac{1}{2}(\sin 2t_{12} - \sin 2t_{13}) \\ \cos 2t_{13} - \cos 2t_{32} & \frac{1}{2}(\sin 2t_{13} - \sin 2t_{32}) \end{vmatrix}$$

$$\mathcal{K}_2^{-1} = \frac{1}{\mathrm{Det}} \begin{vmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{vmatrix} \quad ,$$

wobei die Symbole Det, k_{ij} für die folgenden Terme stehen:

```
DET = ( - COS(2*T13)*SIN(2*T12) + COS(2*T13)*SIN(2*T32) + COS(2*T12)*SIN(2*T13
      ) - COS(2*T12)*SIN(2*T32) - SIN(2*T13)*COS(2*T32) + SIN(2*T12)*COS(2*
      T32))/2
 K  (1.1) = (SIN(2*T13) - SIN(2*T32))/2
 K  (1.2) = (SIN(2*T13) - SIN(2*T12))/2
 K  (2.1) =  - COS(2*T13) + COS(2*T32)
 K  (2.2) =  - COS(2*T13) + COS(2*T12)
```

BIVARIATE POLYNOME ZUR GENÄHERTEN BESTIMMUNG VON UTM-KOORDINATEN (EUROPA-DATUM 1950) AUS GAUSS-KRÜGER-KOORDINATEN (POTSDAM-DATUM)

von Klaus KRACK und Albert SCHÖDLBAUER

SUMMARY

The transformation of Gauss-Krüger-coordinates (Deutsches Hauptdreiecksnetz, Potsdam datum) into UTM-50-coordinates (Europäisches Dreiecksnetz 1950, European datum 1950) is generally performed by means of meshwise affine transformations. This procedure requires the coordinates of all control points of both primary networks. For an approximate solution a closed system of two bivariate fifth-order-polynomials is proposed. The obtained accuracy with maximum deficiencies of 35 cm is sufficient for many tasks.

ZUSAMMENFASSUNG

Anstelle der maschenweisen affinen Umformung der Gauß-Krüger-Koordinaten (Deutsches Hauptdreiecksnetz, Potsdam-Datum) in UTM-50-Koordinaten (Europäisches Dreiecksnetz 1950, Europa-Datum 1950), die die Kenntnis der Koordinaten aller Festpunkte der zugeordneten Grundlagennetze voraussetzt, wird als Näherungslösung ein geschlossenes System von zwei bivariaten Polynomen fünfter Ordnung angegeben. Die Abweichungen erreichen Maximalwerte von 35 cm, was für zahlreiche Aufgabenstellungen ausreicht.

1. EINFÜHRUNG

Entgegen allen Vereinheitlichungsbemühungen sind im Bereich der Bundesrepublik Deutschland immer noch mehrere Koordinatensysteme nebeneinander in Gebrauch, die - mehr oder weniger unabhängig voneinander - die Lage der Punkte des landesweiten geodätischen Festpunktfeldes beschreiben. An Systemen, die bei praktischen Vermessungen, insbesondere im amtlichen Vermessungswesen eine Rolle spielen, sind hier die "alten" Koordinatensysteme der frühen deutschen Landesvermessungen, das Gauß-Krüger-System (Potsdam-Datum) und das UTM-System 50 (Europa-Datum 1950) zu nennen. Die Systeme stützen sich jeweils auf verschiedene Bezugsflächen und auf spezifisch zugeordnete Grundlagennetze, denen unterschiedliche Ausgleichungsmodelle und zum Teil unterschiedliche Beobachtungsdaten zugrunde liegen.

Das Gauß-Krüger-System ist mit Datum Potsdam (Datum Rauenberg) dem Deutschen Hauptdreiecksnetz (DHDN) und dem Bessel-Ellipsoid zugeordnet (***APEL***, 1975; ***WOLF***, 1981), das UTM-System 50 mit Europa-Datum 1950 dem aus dem Zentraleuropäischen Netz (ZEN) entwickelten Europäischen Dreiecksnetz 1950 (DEN 50) bzw. dem Internationalen Ellipsoid (***WOLF***, 1975, 1981; ***KOBOLD***, 1976).

(Die Begriffe Gauß-Krüger- und UTM-50-System, die im folgenden noch mehrfach benutzt werden, kennzeichnen - wie man erkennt - im hier gegebenen Zusammenhang nicht nur und nicht so sehr die Art der Abbildung des Bezugsellipsoids in die Abbildungsebene, sondern darüber hinaus und vor allem die zugrundeliegenden Dreiecksnetze und ihre Verknüpfungen mit den jeweiligen Bezugsellipsoiden.)

Während die alten Systeme mehr und mehr an Bedeutung verlieren, werden sich das Gauß-Krüger- und das UTM-50-System auch in der nächsten Zukunft konkurrierend gegenüberstehen. (Die Konkurrenz dreier Systeme mit den verschiedenen Entwicklungsstufen des Réseau Européen Trigonométrique - RETrig - hat vorläufig rein wissenschaftlichen Charakter, da die aus der RETrig fließenden Koordinaten nicht als Gebrauchskoordinaten genutzt werden.) Nutzer des Gauß-Krüger-Systems sind in erster Linie zivile Stellen, vor allem die Dienststellen der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland. Der Schwerpunkt der Nutzung des UTM-50-Systems liegt bei den Streitkräften der NATO, in jüngerer Zeit mehren sich aber auch für dieses

System die zivilen Anwendungen, insbesondere bei grenzüberschreitenden Projekten, z.B. im Rahmen der übernationalen Raumplanung, die mehr und mehr an Bedeutung gewinnt. Dabei sind Wirtschaftsdaten, sozioökonomische und sozioökologische Daten sowie Umweltdaten, die aufgrund statistischer Erhebungen, Messungen vor Ort und Satellitenaufnahmen anfallen, mit dem durch die Koordinaten beschriebenen "Raum" (Großraum) zu verknüpfen.
Im Gauß-Krüger-System ist das gesamte, in der Bundesrepublik liegende amtliche Lagefestpunktfeld beschrieben, von den trigonometrischen Punkten des DHDN über alle Netzhierarchien bis hin zu den Punkten des Aufnahmenetzes und des Katasterfestpunktfeldes. Dagegen ist im UTM-50-System nur eine Teilmenge dieser Punkte nach Koordinaten erfaßt, wobei gegenüber dem Gauß-Krüger-System grundlegende Unterschiede in der Art und Weise der Koordinatenberechnung bestehen. Diese erfolgte und erfolgt im Gauß-Krüger-System, ausgehend von den Koordinaten des DHDN, unter Nutzung aller verfügbarer Meßdaten im Wege der Punkteinschaltung in mehreren Hierarchien (Einzelpunkt- oder Mehrpunkteinschaltung nach vermittelnden Beobachtungen, Polygonierung, Orthogonal- und Polaraufnahme, photogrammetrische Aufnahmeverfahren), im UTM-50-System stützt sich nur das Grundlagennetz, das EDN 50, auf originäre Meßdaten, die Koordinaten aller übrigen Netzpunkte wurden und werden durch maschenweise Umbildung im Wege der Affintransformation aus dem Gauß-Krüger-System abgeleitet (***SCHÖDLBAUER***, 1982, S. 258-271); sie können insofern nur als Näherungskoordinaten betrachtet werden. Der Grund für die unterschiedliche Sachbehandlung liegt in den Zugriffsmöglichkeiten auf die ursprünglichen Meßdaten und in der Bedarfslage der Nutzer begründet: Bei den Hauptnutzern des Gauß-Krüger-Systems, den Dienststellen der Vermessungsverwaltung, der Trend zum Kordinatenkataster, in dem im Laufe der Entwicklung jeder Grenzstein und jeder topographische Gegenstand durch Koordinaten hoher Genauigkeit festgehalten werden soll, auf der Seite der Nutzer des UTM-50-Systems aufgabenbedingte Beschränkung auf ein verhältnismäßig grobes Punktraster, das im wesentlichen durch die Trigonometrischen Punkte gebildet wird, mit im übrigen geringeren Genauigkeitsansprüchen. Mittlerweile wird auch im UTM-50-System für verschiedene neue Aufgaben eine wesentlich höhere Punktdichte angestrebt.

2. AFFINTRANSFORMATION

Bei der bereits erwähnten maschenweisen Umbildung im Wege der Affintransformation werden die im Gauß-Krüger-System beschriebenen Maschen des DHDN auf

die Maschen des EDN 50 abgebildet, wobei die Zuordnung über die in der Natur identischen Maschenneckpunkte erfolgt. Die affine Transformation ist eine lineare Transformation, mit der eine Ebene in eine andere abgebildet werden kann, wobei Punkte wieder als Punkte, gerade Linien und parallele Geraden als parallele Geraden dargestellt werden und die auf (geraden) Strecken bestehenden Streckenverhältnisse erhalten bleiben.

Mit den Koordinatenbezeichnungen

R = Rechtswert
H = Hochwert
} im Gauß-Krüger-System

und

E = Easting
N = Northing
} im UTM-50-System

lauten die Transformationsformeln für eine bestimmte Dreiecksmasche, etwa jene mit der laufenden Maschen-Nummer k

$$E_k = a_{1k} \cdot R_k + a_{2k} \cdot H_k + a_{3k}$$
$$N_k = b_{1k} \cdot R_k + b_{2k} \cdot H_k + b_{3k} \qquad (1)$$

Die einzelne Transformation ist - wie man erkennt - bei $a_{1k} \cdot b_{2k} - a_{2k} \cdot b_{1k} \neq 0$ durch sechs Transformationsparameter a_{1k}, a_{2k}, a_{3k}, b_{1k}, b_{2k}, b_{3k} umkehrbar eindeutig bestimmt. Mit ihr läßt sich ein vorgegebenes Dreieck in ein beliebiges anderes überführen.

Diese Eigenschaft der Affintransformation läßt sich nutzen, um - wie bereits erwähnt - die Dreiecke des DHDN und die von den Dreiecksmaschen umschlossenen Ausschnitte des Festpunktfeldes auf der einen Seite und die Dreiecke des EDN 50, die sich aufgrund der geschilderten Unterschiede in den geodätischen Grundlagen im Gauß-Krüger- und im UTM-50-System in geringfügig unterschiedlicher Form und Größe darstellen, auf der anderen Seite, aufeinander zu beziehen (***SCHÖDLBAUER***, 1982, S. 258-271).

Die Auswertung der Formeln (1) für beliebige in der Bundesrepublik liegende Punkte setzt die Kenntnis der Gauß-Krüger- und der UTM-50-Koordinaten aller

identischen Festpunkte des DHDN und des EDN 50 voraus, was die Bereitstellung umfangreicher Koordinatenverzeichnisse des Ausgangspunkte erfordert.

3. TRANSFORMATION DURCH BIVARIATE POLYNOME

Im folgenden wird ein auf Polynome mit zwei Veränderlichen aufgebautes geschlossenes Formelsystem entwickelt, das unabhängig von der Kenntnis der Koordinaten des DHDN und des EDN 50 die Umformung von Gauß-Krüger- in UTM-50-Koordinaten für ein ganzes Meridianstreifensystem in einem Guß bewirkt, wobei gewisse Genauigkeitseinbußen von vorneherein hingenommen werden. Diese Einschränkung ist jedoch für zahlreiche Anwendungen mit Genauigkeitsansprüchen oft nur im Meter-Bereich ohne Belang, zumal auch der maschenweisen Umbildung nur die Qualität einer Näherungslösung zugesprochen werden kann.

Der Lösungsansatz lautet:

$$\Delta E = \sum_{i=0}^{m} \sum_{j=0}^{m} a_{ij} \cdot \Delta R^i \cdot \Delta H^j$$

$$\Delta N = \sum_{i=0}^{m} \sum_{j=0}^{m} b_{ij} \cdot \Delta R^i \cdot \Delta H^j \qquad (2)$$

mit $i + j \leq m$.

Er stellt eine Beziehung her zwischen den auf die jeweiligen Schwerpunkte bezogenen Koordinatenunterschiede im Gauß-Krüger-System

$$\Delta R = R - R_s$$

$$\Delta H = H - H_s$$

und den entsprechenden Koordinatenunterschieden im UTM-System

$$\Delta E = E - E_s$$

$$\Delta N = N - N_s \ .$$

Die $2 \cdot u = (m+1) \cdot (m+2)$ Koeffizienten a_{ij} und b_{ij} werden nun durch Parameterschätzung so bestimmt, daß die Differenzen zwischen den vorgegebenen Koordinaten $\overline{E}$ und $\overline{N}$ der Punkte des EDN 50 und den aus den Koordinaten R und H der Punkte des DHDN abgeleiteten, nach (2) umgeformten Koordinaten E und N, also die Restklaffungen $\overline{E} - E$ und $\overline{N} - N$ über

$$\sum \left\{ (\overline{E} - E)^2 + (\overline{N} - N)^2 \right\} = \text{minimum} \qquad (3)$$

einer plausiblen Minimumsbedingung unterworfen werden. (3) ist dabei über alle in einem bestimmten Meridianstreifensystem liegenden Festpunkte der aufeinander bezogenen Grundlagennetze zu summieren.

Über den Grad der Entwicklung m kann mit Einschränkung noch oben hin zunächst frei verfügt werden. (Die Beschränkung ist wegen der Eigenschaft von Polynomfunktionen geboten, bei zu vielen Stützstellen u.U. zu "schwingen".) Im einfachsten Fall, der für $m = 0$ vorliegt, werden die Koordinatensysteme lediglich durch Parallelverschiebung aneinander herangeführt. $m = 1$ steht für eine Affintransformation nach dem von ***HASSE*** (1963) und ***WOLF*** (1968, S. 444-446) beschriebenen Prinzip, das der Minimierungsvorschrift der "Helmert-Transformation" folgt. Man wird die Entwicklung zweckmäßigerweise dort abbrechen, wo der zusätzliche Rechenaufwand, den die Erweiterung der Polynome zwangsläufig auslöst, im Verhältnis zum Genauigkeitsgewinn nicht mehr sinnvoll erscheint. Kriterien für die Güte der Annäherung sind die aus allen Restklaffungen fließenden

mittleren Restklaffungen

$$r_{oE} = \pm \sqrt{\frac{\Sigma (\overline{E} - E)^2}{n - u}} \qquad \text{und}$$

$$r_{oN} = \pm \sqrt{\frac{\Sigma (\overline{N} - N)^2}{n - u}} \qquad , \qquad (4)$$

die den Grad der Anpassung des Umformungsmodells (2) an das Modell der maschenweisen Affintransformation (1) beschreiben, und die

maximalen Restklaffungen

$$r_{max\,E} = \text{Maximalwert aus allen } |\overline{E} - E| \qquad (5)$$
$$r_{max\,N} = \text{Maximalwert aus allen } |\overline{N} - N|$$

Die Indizes E und N bei r_0 und r_{max} kennzeichnen die beiden Koordinatenrichtungen,

n = Anzahl der Vergleichspunkte und

u = Anzahl der Parameter a_{ij} bzw. b_{ij}, die mit dem Entwicklungsgrad m über $2 \cdot u = (m+1) \cdot (m+2)$ zusammenhängt.

Zu (2) inverse Polynome

$$\Delta R = \sum_{i=0}^{m} \sum_{j=0}^{m} a^*_{ij} \cdot \Delta E^i \cdot \Delta N^j \qquad (6)$$
$$\Delta H = \sum_{i=0}^{m} \sum_{j=0}^{m} b^*_{ij} \cdot \Delta E^i \cdot \Delta N^j$$

mit $i + j \leq m$,

die für die Berechnung von Gauß-Krüger- aus UTM-50-Koordinaten benutzt werden können, lassen sich in analoger Weise finden, wenn in allen Ansätzen (2) bis (5) die Koordinatensysteme R, H auf der einen und E, N auf der anderen Seite miteinander vertauscht werden. Weitere Lösungsansätze ergeben sich durch den Einsatz der von ***GLASMACHER/KRACK*** (1984) angegebenen Verfahren der Reihenumkehr, durch die (6) unmittelbar aus (2) entwickelt werden kann.

4. KOEFFIZIENTENBESTIMMUNG

Zur Bestimmung eines konkreten Koeffizientensatzes a_{ij}, b_{ij} wurden alle in der UTM-Zone 32 (9°-System mit einer Streifenbreite von 6°) liegenden Punkte des DHDN und des EDN 50 herangezogen und gleichgewichtig in den Ausgleichungsalgorithmus eingebracht. Die der Beurteilung der gewonnenen Polynome dienenden Ergebnisse sind in der nachfolgenden Tabelle 1 wiedergegeben.

m	u	r_{oE}	r_{oN}	$r_{max\ E}$	$r_{max\ N}$
0	1	± 50,6 m	± 95,2 m	100,6 m	177,4 m
1	3	± 0,75 m	± 0,68 m	2,6 m	2,1 m
2	6	± 0,20 m	± 0,19 m	0,72 m	0,84 m
3	10	± 0,13 m	± 0,16 m	0,43 m	0,61 m
4	15	± 0,11 m	± 0,13 m	0,34 m	0,46 m
5	21	± 0,10 m	± 0,11 m	0,29 m	0,41 m
6	28	± 0,08 m	± 0,10 m	0,25 m	0,51 m

Tabelle 1

Man erkennt, daß mit steigender Ordnungszahl m des Entwicklungsgrades der eingesetzten Polynome die mittleren Restklaffungen r_{oE} und r_{oN} immer kleiner werden. Dies entspricht dem Anpassungsverhalten und -vermögen der Polynome.

Mehr als die mittleren Restklaffungen sind für den Anwender die maximalen Restklaffungen $r_{max\ E}$ und $r_{max\ N}$ von Bedeutung, die die Grenzen der Polynomanpassung und damit eines auf diesem Wege abgeleiteten Formelsystems erkennen lassen. Auch diese Klaffungen nehmen genähert mit zunehmender Ordnungszahl m ab, erreichen dann aber bei $m = 5$ ein Minimum. Mit Sicht auf die maximalen Restklaffungen liefern offenbar Polynome 5. Ordnung bessere Ergebnisse als die entsprechenden Polynome der 6. Ordnung.

Die in Tabelle 1 ausgewiesenen Zahlenwerte ergaben sich, wie bereits erwähnt, bei gleichgewichtiger Behandlung aller Restklaffungen im Ausgleichungsmodell. Untersuchungen mit verschiedenen anderen Gewichtsmodellen zeigten, daß mit proportional zu den Restklaffungen angesetzten Gewichten noch eine weitere Minimierung der maximalen Restklaffungen erreicht werden kann. Für den Entwicklungsgrad $m = 5$ ergibt sich dabei:

m	u	r_{oE}	r_{oN}	$r_{max\ E}$	$r_{max\ N}$
5	21	± 0,10 m	± 0,11 m	0,25 m	0,35 m

Tabelle 2

Die diesen Ergebnissen entsprechenden Parameter a_{ij} und b_{ij} bzw. a^*_{ij} und b^*_{ij} sind in die nachfolgend angegebenen Rechenformeln (7) und (8) übernommen.

5. RECHENFORMELN

Nachfolgend sind die Formeln zusammengestellt, die für den Bereich der Bundesrepublik Deutschland zur Transformation von UTM-50-Koordinaten der Zone 32 (6°-12° ö.L.) in Gauß-Krüger-Koordinaten des 9°-Meridianstreifensystems sowie für die Umkehrung diesen Vorgangs ermittelt wurden.

Für den Bereich 10°30'-12° ö.L. ist dabei zu beachten, daß die amtlichen Verzeichnisse die hier einschlägigen Gauß-Krüger-Koordinaten nicht im 9°-System (Kennziffer 3), sondern im 12°-System (Kennziffer 4) ausweisen. Gauß-Krüger-Koordinaten des 12°-Systems im Bereich 10°30'-12° (Kennziffer 4, Rechtswert R < 500 km) sind daher vor Anwendung der Formeln (7) mit Hilfe einer Meridianstreifentransformation z.B. nach ***SCHÖDLBAUER*** (1982, S. 137-225) in das 9°-System (Kennziffer 3) überzuführen. Entsprechendes gilt für die Weiterverarbeitung der Ergebnisse aus (8). Fallen die Punkte in den vorerwähnten Bereich, so müssen die nach (8) ermittelten Koordinaten mit Hilfe einer Meridianstreifentransformation aus dem 9°-System (Kennziffer 3) in das 12°-System (Kennziffer 4) umgesetzt werden.

UTM-50-Koordinaten (Zone 32) aus Gauß-Krüger-Koordinaten (9°-System)

$\Delta R = R - R_S$ mit $R_S = 3\,512\,691{,}423$ m

$\Delta H = H - H_S$ $H_S = 5\,658\,061{,}458$ m

$$\Delta E = -0{,}384\ \text{m} + (-0{,}387247\cdot10^{-3} + 1)\cdot\Delta R + 0{,}7141\cdot10^{-5}\cdot\Delta H - 0{,}1650\cdot10^{-10}\ \text{m}^{-1}\cdot\Delta R^2 + 0{,}4884\cdot10^{-11}\ \text{m}^{-1}\cdot\Delta R\cdot\Delta H + 0{,}9376\cdot10^{-11}\ \text{m}^{-1}\cdot\Delta H^2 - 0{,}8298\cdot10^{-16}\ \text{m}^{-2}\cdot\Delta R^3 + 0{,}8897\cdot10^{-17}\ \text{m}^{-2}\cdot\Delta R^2\cdot\Delta H - 0{,}4666\cdot10^{-16}\ \text{m}^{-2}\cdot\Delta R\cdot\Delta H^2 + 0{,}5642\cdot10^{-18}\ \text{m}^{-2}\cdot\Delta H^3 + 0{,}3724\cdot10^{-22}\ \text{m}^{-3}\cdot\Delta R^4 - 0{,}1514\cdot10^{-21}\ \text{m}^{-3}\cdot\Delta R^3\cdot\Delta H + 0{,}2004\cdot10^{-22}\ \text{m}^{-3}\cdot\Delta R^2\cdot\Delta H^2 - 0{,}5215\cdot10^{-22}\ \text{m}^{-3}\cdot\Delta R\cdot\Delta H^3 + 0{,}1679\cdot10^{-22}\ \text{m}^{-3}\cdot\Delta H^4 + 0{,}6216\cdot10^{-27}\ \text{m}^{-4}\cdot\Delta R^5 - 0{,}3037\cdot10^{-27}\ \text{m}^{-4}\cdot\Delta R^4\cdot\Delta H + 0{,}1057\cdot10^{-27}\ \text{m}^{-4}\cdot\Delta R^3\cdot\Delta H^2 + 0{,}1993\cdot10^{-27}\ \text{m}^{-4}\cdot\Delta R^2\cdot\Delta H^3 + 0{,}1343\cdot10^{-27}\ \text{m}^{-4}\cdot\Delta R\cdot\Delta H^4 + 0{,}4290\cdot10^{-28}\ \text{m}^{-4}\cdot\Delta H^5$$

$$\Delta N = -0{,}145\ \text{m} - 0{,}1129\cdot10^{-4}\cdot\Delta R + (-0{,}390609\cdot10^{-3} + 1)\cdot\Delta H + 0{,}6631\cdot10^{-11}\ \text{m}^{-1}\cdot\Delta R^2 - 0{,}2662\cdot10^{-10}\ \text{m}^{-1}\cdot\Delta R\cdot\Delta H + 0{,}7037\cdot10^{-12}\ \text{m}^{-1}\cdot\Delta H^2 + 0{,}5943\cdot10^{-16}\ \text{m}^{-2}\cdot\Delta R^3 + 0{,}7967\cdot10^{-16}\ \text{m}^{-2}\cdot\Delta R^2\cdot\Delta H - 0{,}1064\cdot10^{-16}\ \text{m}^{-2}\cdot\Delta R\cdot\Delta H^2 - 0{,}1210\cdot10^{-16}\ \text{m}^{-2}\cdot\Delta H^3 - 0{,}4731\cdot10^{-22}\ \text{m}^{-3}\cdot\Delta R^4 + 0{,}1727\cdot10^{-21}\ \text{m}^{-3}\cdot\Delta R^3\cdot\Delta H - 0{,}6578\cdot10^{-22}\ \text{m}^{-3}\cdot\Delta R^2\cdot\Delta H^2 - 0{,}2637\cdot10^{-22}\ \text{m}^{-3}\cdot\Delta R\cdot\Delta H^3 + 0{,}1944\cdot10^{-23}\ \text{m}^{-3}\cdot\Delta H^4 - 0{,}1504\cdot10^{-27}\ \text{m}^{-4}\cdot\Delta R^5 - 0{,}6793\cdot10^{-27}\ \text{m}^{-4}\cdot\Delta R^4\cdot\Delta H - 0{,}1498\cdot10^{-27}\ \text{m}^{-4}\cdot\Delta R^3\cdot\Delta H^2 - 0{,}3840\cdot10^{-27}\ \text{m}^{-4}\cdot\Delta R^2\cdot\Delta H^3 + 0{,}6315\cdot10^{-28}\ \text{m}^{-4}\cdot\Delta R\cdot\Delta H^4 + 0{,}9029\cdot10^{-28}\ \text{m}^{-4}\cdot\Delta H^5$$

$E = E_S + \Delta E$ mit $E_S = 512\,693{,}746$ m

$N = N_S + \Delta N$ $N_S = 5\,656\,440{,}634$ m .

(7)

Gauß-Krüger-Koordinaten (9°-System) aus UTM-50-Koordinaten (Zone 32)

$\Delta E = E - E_S$ mit $E_S = 512\,693{,}746$ m

$\Delta N = N - N_S$ $N_S = 5\,656\,440{,}634$ m

$$\begin{aligned}
\Delta R = {} & + 0{,}384 \text{ m} \\
& + (0{,}387397 \cdot 10^{-3} + 1) \cdot \Delta E \\
& - 0{,}7147 \cdot 10^{-5} \cdot \Delta N \\
& + 0{,}1652 \cdot 10^{-10}\ m^{-1} \cdot \Delta E^2 \\
& - 0{,}4891 \cdot 10^{-11}\ m^{-1} \cdot \Delta E \cdot \Delta N \\
& - 0{,}9387 \cdot 10^{-11}\ m^{-1} \cdot \Delta N^2 \\
& + 0{,}8911 \cdot 10^{-16}\ m^{-2} \cdot \Delta E^3 \\
& - 0{,}8911 \cdot 10^{-17}\ m^{-2} \cdot \Delta E^2 \cdot \Delta N \\
& + 0{,}4674 \cdot 10^{-16}\ m^{-2} \cdot \Delta E \cdot \Delta N^2 \\
& - 0{,}5658 \cdot 10^{-18}\ m^{-2} \cdot \Delta N^3 \\
& - 0{,}3731 \cdot 10^{-22}\ m^{-3} \cdot \Delta E^4 \\
& + 0{,}1517 \cdot 10^{-21}\ m^{-3} \cdot \Delta E^3 \cdot \Delta N \\
& + 0{,}2008 \cdot 10^{-22}\ m^{-3} \cdot \Delta E^2 \cdot \Delta N^2 \\
& - 0{,}5225 \cdot 10^{-22}\ m^{-3} \cdot \Delta E \cdot \Delta N^3 \\
& + 0{,}1682 \cdot 10^{-22}\ m^{-3} \cdot \Delta N^4 \\
& + 0{,}6231 \cdot 10^{-27}\ m^{-4} \cdot \Delta E^5 \\
& - 0{,}3045 \cdot 10^{-27}\ m^{-4} \cdot \Delta E^4 \cdot \Delta N \\
& + 0{,}1060 \cdot 10^{-27}\ m^{-4} \cdot \Delta E^3 \cdot \Delta N^2 \\
& + 0{,}1997 \cdot 10^{-27}\ m^{-4} \cdot \Delta E^2 \cdot \Delta N^3 \\
& + 0{,}1346 \cdot 10^{-27}\ m^{-4} \cdot \Delta E \cdot \Delta N^4 \\
& + 0{,}4299 \cdot 10^{-28}\ m^{-4} \cdot \Delta N^5
\end{aligned}$$

$$\begin{aligned}
\Delta H = {} & + 0{,}145 \text{ m} \\
& + 0{,}1130 \cdot 10^{-4} \cdot \Delta E \\
& + (0{,}390761 \cdot 10^{-3} + 1) \cdot \Delta N \\
& - 0{,}6638 \cdot 10^{-11}\ m^{-1} \cdot \Delta E^2 \\
& + 0{,}2666 \cdot 10^{-10}\ m^{-1} \cdot \Delta E \cdot \Delta N \\
& - 0{,}7049 \cdot 10^{-12}\ m^{-1} \cdot \Delta N^2 \\
& - 0{,}5952 \cdot 10^{-16}\ m^{-2} \cdot \Delta E^3 \\
& - 0{,}7979 \cdot 10^{-16}\ m^{-2} \cdot \Delta E^2 \cdot \Delta N \\
& + 0{,}1065 \cdot 10^{-16}\ m^{-2} \cdot \Delta E \cdot \Delta N^2 \\
& + 0{,}1212 \cdot 10^{-16}\ m^{-2} \cdot \Delta N^3 \\
& + 0{,}4739 \cdot 10^{-22}\ m^{-3} \cdot \Delta E^4 \\
& - 0{,}1730 \cdot 10^{-21}\ m^{-3} \cdot \Delta E^3 \cdot \Delta N \\
& + 0{,}6592 \cdot 10^{-22}\ m^{-3} \cdot \Delta E^2 \cdot \Delta N^2 \\
& + 0{,}2642 \cdot 10^{-22}\ m^{-3} \cdot \Delta E \cdot \Delta N^3 \\
& - 0{,}1948 \cdot 10^{-23}\ m^{-3} \cdot \Delta N^4 \\
& + 0{,}1508 \cdot 10^{-27}\ m^{-4} \cdot \Delta E^5 \\
& + 0{,}6809 \cdot 10^{-27}\ m^{-4} \cdot \Delta E^4 \cdot \Delta N \\
& + 0{,}1501 \cdot 10^{-27}\ m^{-4} \cdot \Delta E^3 \cdot \Delta N^2 \\
& + 0{,}3849 \cdot 10^{-27}\ m^{-4} \cdot \Delta E^2 \cdot \Delta N^3 \\
& - 0{,}6332 \cdot 10^{-28}\ m^{-4} \cdot \Delta E \cdot \Delta N^4 \\
& - 0{,}9050 \cdot 10^{-28}\ m^{-4} \cdot \Delta N^5
\end{aligned}$$

$R = R_S + \Delta R$ mit $R_S = 3\,512\,691{,}423$ m

$H = H_S + \Delta H$ $H_S = 5\,658\,061{,}458$ m .

(8)

6. RECHENBEISPIEL

Die folgende Tabelle weist für den TP "Münsing Kirche" die Ergebnisse der Transformation nach den Formeln (7) und (8) aus. Zum Vergleich sind auch die Ergebnisse einer Affintransformation nach (1) angegeben, die sich auf die Masche des DHDN/EDN 50 "München, Frauenkirche - Hohenpeissenberg - Benediktbeuern, Benediktenwand" beziehen.

GK-Ausgangskoordinaten im 12°-System (Kennziffer 4)					
R	44 52 783,94 m	Überführung in das 9°-System			
H	53 07 417,56 m	z.B. nach ***SCHÖDLBAUER*** (1982, S. 137-225)			
GK-Ausgangskoordinaten im 9°-System (Kennziffer 3)					
R	36 77 042,16 m	R_S	35 12 691,42 m	ΔR	164 350,74 m
H	53 09 940,33 m	H_S	56 58 061,46 m	ΔH	- 348 121,13 m
UTM- aus GK-Koordinaten durch Polynomtransformation (7)					
E	676 978,01 m	E_S	512 693,75 m	ΔE	164 284,26 m
N	53 08 454,98 m	N_S	56 56 440,63 m	ΔN	- 347 985,65 m
GK- aus UTM-Koordinaten durch Polynomtransformation (8)					
R	36 77 042,17 m	R_S	512 691,42 m	ΔR	164 350,75 m
H	53 09 940,33 m	H_S	56 58 061,46 m	ΔH	- 348 121,13 m
Vergleichsergebnisse durch Affintransformation (1)					
P_1	München, Fr.Ki.	P_2	Hohenpeißenberg	P_3	Benediktenwand
R_1	36 91 564,20 m	R_2	36 50 592,78 m	R_3	36 85 301,31 m
H_1	53 36 611,13 m	H_2	52 97 872,95 m	H_3	52 82 371,37 m
E_1	691 494,43 m	E_2	650 539,14 m	E_3	685 233,82 m
N_1	53 35 115,08 m	N_2	52 96 392,33 m	N_3	52 80 896,84 m
E	676 978,06 m	Alle UTM-Koordinaten beziehen			
N	53 08 454,88 m	sich auf UTM-Zone 32			

LITERATUR

APEL, H.: ***Bemerkungen über Aufbau und Genauigkeit einiger Teile des Deutschen Hauptdreiecksnetzes***. Zeitschrift für Vermessungswesen 100, S. 149-157, 1975

GLASMACHER, H., KRACK, K.: ***Umkehrung von vollständigen Potenzreihen mit zwei Veränderlichen***. Schriftenreihe des Wissenschaftlichen Studiengangs Vermessungswesen der Hochschule der Bundeswehr München, Heft 10, 10 Jahre HSBwM - Arbeiten aus dem Institut für Geodäsie, S. 49-70, München, 1984

HASSE, D.: ***Affine Transformation analog der Helmerttransformation***. Allgemeine Vermessungsnachrichten 71, S. 456-457, 1963

KOBOLD, F.: ***Das Europäische Dreiecksnetz (RETRIG) – Entstehung, Stand der Arbeiten***. Vermessung, Photogrammetrie, Kulturtechnik 74, S. 333-338, 1976

SCHÖDLBAUER, A.: ***Rechenformeln und Rechenbeispiele zur Landesvermessung Teil 2 – Geodätische Berechnungen im System der Gaußschen konformen Abbildung eines Bezugsellipsoids unter besonderer Berücksichtigung des Gauß-Krüger- und des UTM-Koordinatensystems im Bereich der Bundesrepublik Deutschland***. Herbert Wichmann Verlag, Karlsruhe, 1982

WOLF, H.: ***Ausgleichungsrechnung nach der Methode der kleinsten Quadrate***. Dümmlers Verlag, Bonn, Hannover, München, 1968

WOLF, H.: ***Das Europäische Dreiecksnetz – Ursprung, Entwicklung und künftige Möglichkeiten***. Geodätische Woche Köln, 1975, S. 40-50, Konrad Wittwer Verlag, Stuttgart, 1976

WOLF, H.: ***Die deutsche Landesvermessung in ihrer historischen Entwicklung und ihren künftigen Möglichkeiten***. Festschrift 100 Jahre Geodätische Lehre und Forschung in Hannover; Heft 100 der Wissenschaftlichen Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, S. 47-53, Hannover, 1981

STRENGE, NICHT DIFFERENTIELLE LÖSUNG FÜR DEN ELLIPSOIDÜBERGANG UND DIE KOORDINATEN-TRANSFORMATION NACH HELMERT IM SYSTEM DER GEOGRAPHISCHEN KOORDINATEN

von Albert SCHÖDLBAUER

SUMMARY

Helmert's method for ellipsoid transition and displacement, rotation and scale change of two dimensional horizontal geodetic networks extended on reference ellipsoids uses the total differentials of the corresponding transfer functions (system functions). The method yields correct results only by applying complicated rigorous formulas of the partial derivatives and by restricting its application to small (differential) changes of the parameters.

The solution described in the following introduces the transfer functions themselves. With the aid of modern computing potentials no difficulties are rising in the evaluation of these functions to any accuracy and without restrictions in the magnitude of changes of the parameters.

ZUSAMMENFASSUNG

Das von Helmert angegebene Verfahren für den Ellipsoid-Übergang sowie zur Verschiebung und Drehstreckung geodätischer Lagenetze auf Bezugsellipsoiden stützt sich auf totale Differentiale der entsprechenden Übertragungsfunktionen (Stammfunktionen). Es führt bei ausgedehnten Netzen nur dann zu richtigen Ergebnissen, wenn für die auftretenden partiellen Differentialquotienten strenge Formeln verwendet werden, deren Auswertung jedoch sehr umständlich ist. Das Verfahren gilt im übrigen nur für differentielle Vorgänge.

Der im folgenden dargestellte neue Lösungsweg benutzt die Übertragungsfunktionen selbst, die aufgrund der heute zu Gebote stehenden rechentechnischen Möglichkeiten in aller Strenge ausgewertet werden können. Sie liefern auch bei beliebig ausgedehnten Netzen und bei beliebigen Beträgen der Netzverschiebung, -drehung und -streckung genaue Ergebnisse.

Zu den Grundaufgaben der geodätischen Praxis zählt das Problem, Koordinaten eines Punktfeldes, die durch unterschiedliche Meßdaten, Datumsvereinbarungen und Ausgleichungsverfahren in zwei voneinander unabhängigen geodätischen Systemen bestimmt sind -"ungleichartige Koordinaten" -, miteinander in Beziehung zu bringen. Die Aufgabe stellt sich bei allenthalben notwendigen Koordinatenvergleichen, aber auch, wenn Koordinaten des einen Systems, des "Startsystems", in solche des anderen, des "Zielsystems", überführt werden sollen, z.B. Gauß-Krüger- oder geographischen Koordinaten von Punkten, die dem Deutschen Hauptdreiecksnetz angehören oder daraus abgeleitet sind (Bessel-Ellipsoid, Potsdam-Datum), in UTM- oder geographische Koordinaten im System des Zentraleuropäischen Netzes (Internationales Ellipsoid, Europa-Datum). Aufgrund der fehlenden Verknüpfung der beiden Systeme sind die Koordinaten des einen Systems mit denen des anderen nur beschränkt kompatibel.

Die strenge Lösung der Aufgabe würde erfordern, die Punktverdichtung in beiden Systemen mit identischem Datenmaterial und nach einheitlichen Ausgleichungsregeln nachzuvollziehen. Wegen meist fehlender Zugriffsmöglichkeiten auf die Ursprungsdaten, aber auch wegen des Umfangs der dabei notwendigen Rechenarbeiten, ist dieser unter theoretischen Aspekten wünschenswerte Weg in der Regel nicht gangbar.

Vor diesem Hintergrund vermögen Näherungsverfahren, mit denen ungleichartige Koordinaten miteinander verglichen und verknüpft werden können, wertvolle Dienste zu leisten.

Die folgenden Betrachtungen befassen sich, wie bereits beispielhaft angedeutet, mit Punktfeldern, die durch ungleichartige geographische oder Gaußsche Koordinaten auf Bezugsellipsoiden verankert sind. Es handelt sich dabei vor allem um die nach den Verfahren der Landesvermessung entwickelten weiträumigen Lagenetze, die das betreffende Punktfeld je nach dem jeweils zugeordneten Bezugsellipsoid und in Abhängigkeit vom Datum des Koordinatensystems, vom Beobachtungsmaterial und vom Ausgleichungsverfahren unterschiedlich modellieren und koordinatenmäßig festlegen.

Als wichtigste Näherungsverfahren für den Vergleich und die Verknüpfung ungleichartiger Koordinaten der Landesvermessung stehen die Affintransformation (***SCHÖDLBAUER***, 1982, S. 258-271) sowie verschiedene Verfahren für "Ellip-

soidübergang und Netzverschiebung" (***JORDAN / EGGERT / KNEISSL***, 1959, S. 1247-1282) zur Verfügung. Unter der letztgenannten Gruppe von Transformationen, die eine Reihe von miteinander konkurrierenden Verfahren enthält, ist vor allem der Beitrag von ***HELMERT*** (1880, S. 279-296; 1893) hervorzuheben, der Theorie und Praxis nachhaltig befruchtet hat. Auch die nachfolgenden Betrachtungen heben darauf ab.

Helmert denkt sich bei dem Vorgang die auf einem Bezugsellipsoid I verankerten Punkte des Punktfeldes P_i^I auf einen Zentralpunkt P_o^I bezogen und durch geodätische Polarkoordinaten - Azimute A_{oi}^I, kürzeste Verbindungslinien S_{oi}^I - festgelegt. In einem ersten Schritt ("Ellipsoidübergang") werden ohne Änderung der Polarkoordinaten die Ellipsoidkonstanten so geändert, daß die Dimension des korrespondierenden Ellipsoids II erreicht wird. Der zweite Schritt ("Netzverschiebung", "Drehstreckung") beinhaltet die Verschiebung des Zentralpunktes mit gleichzeitiger Drehung und Maßstabsänderung des polaren Koordinatensystems. Beide Schritte haben Lageänderungen der Punkte P_i^I des Punktfeldes zur Folge, die sich bei differentieller Betrachtung wie folgt darstellen:

$$\begin{aligned} dB_i &= \frac{\partial B_i}{\partial a}\cdot da + \frac{\partial B_i}{\partial f}\cdot df + \frac{\partial B_i}{\partial B_o}\cdot dB_o + \frac{\partial B_i}{\partial L_o}\cdot dL_o + \frac{\partial B_i}{\partial A_{oi}}\cdot dA + \frac{\partial B_i}{\partial m_{oi}}\cdot dm \\ dL_i &= \frac{\partial L_i}{\partial a}\cdot da + \frac{\partial L_i}{\partial f}\cdot df + \frac{\partial L_i}{\partial B_o}\cdot dB_o + \frac{\partial L_i}{\partial L_o}\cdot dL_o + \frac{\partial L_i}{\partial A_{oi}}\cdot dA + \frac{\partial L_i}{\partial m_{oi}}\cdot dm \,. \end{aligned} \qquad (1)$$

Die Ausdrücke sind die totalen Differentiale von Funktionen

$$\begin{aligned} B_i &= B(a,f,B_o,L_o,A_{oi},S_{oi}) \\ L_i &= L(a,f,B_o,L_o,A_{oi},S_{oi}) \,, \end{aligned} \qquad (2)$$

die die sogenannte Erste geodätische Grundaufgabe auf Bezugsellipsoiden beschreiben. Die Variablen dieser Funktionen haben dabei folgende Bedeutung:

a = große Halbachse
b = kleine Halbachse
$f = \frac{a-b}{a}$ = Abplattung
} des Bezugsellipsoids

B_o = geographische Breite
L_o = geographische Länge
} des Zentralpunkts P_o

B_i = geographische Breite
L_i = geographische Länge
} der Feldpunkte P_i

A_{oi} = Azimute der geodätischen Linien P_oP_i in P_o

S_{oi} = Geodätische Linien (Linienabschnitte) P_oP_i

Auf diese Variablen beziehen sich auch die Differentiale von (1), wobei über $dS_{oi} = S_{oi} \cdot dm$ eine Beziehung zwischen den Strecken S_{oi} und einem an alle Strecken anzulegenden Maßstabsfaktor m hergestellt ist.

Für die partiellen Differentialquotienten in (1) bzw. für Differentialquotienten $\frac{\partial \beta_i}{\partial a}, \frac{\partial \beta_i}{\partial e^2}, \frac{\partial L_i}{\partial e^2}, \frac{\partial B_i}{\partial S_{oi}}, \frac{\partial L_i}{\partial S_{oi}}$ (β = "reduzierte Breite", e^2 = Quadrat der "ersten numerischen Exzentrizität"). die - wenn auch mit anderen Parametern - gleichwertig zu den hier benutzten Ausdrücken $\frac{\partial B_i}{\partial a}, \frac{\partial B_i}{\partial f}, \frac{\partial L_i}{\partial f}, \frac{\partial B_i}{\partial m_{oi}}, \frac{\partial L_i}{\partial m_{oi}}$ die Auswirkungen von differentiellen Dimensionsänderungen des Bezugsellipsoids auf die Lage der Punkte P_i darlegen, hat ***HELMERT*** (1880, S. 282, 283, 285, 292, 293, 294, 295) strenge und genäherte Formeln angegeben (siehe hierzu auch ***JORDAN/EGGERT/KNEISSL***, 1959, S. 1253-1260), die die differentiellen Zusammenhänge zwischen den unabhängigen und den abhängigen Veränderlichen von (2) erschöpfend und in aller Strenge beschreiben. Kennzeichnet man die beiden über die Differentialbezeichnungen (1) korrespondierenden Bezugsellipsoide und die jeweils zugeordneten Punktkoordinaten durch die Indizes I und II, so lautet die Lösung des Problems

$$
\begin{aligned}
\text{mit}\quad da &= a^{II} - a^{I} \\
df &= f^{II} - f^{I} \\
dB_o &= B_o^{II} - B_o^{I} \\
dL_o &= L_o^{II} - L_o^{I} \\
dA_{oi} &= dA \left(= A_{oi}^{II} - A_{oi}^{I}\right) \\
dm_{oi} &= dm \left(= \left(S_{oi}^{II} - S_{oi}^{I}\right)/S_{oi}^{I}\right)
\end{aligned}
\qquad (3)
$$

nach Auswertung von (1):

$$
\begin{aligned}
B_i^{II} &= B_i^{I} + dB_i \\
L_i^{II} &= L_i^{I} + dL_i
\end{aligned}
\qquad (4)
$$

Die strengen Ausdrücke von (1), insbesondere diejenigen, die den Ellipsoidübergang beschreiben, waren wegen ihrer komplizierten Struktur ursprünglich kaum auswertbar und schon ***HELMERT*** (1880, S. 285, 295) hat sich mit Hilfe sphärischer Approximationen um Vereinfachungen bemüht (siehe hierzu auch ***JORDAN/EGGERT/KNEISSL***, 1959, S. 1255, 1259), um die praktische Anwendung der Formeln zu erleichtern. Beiträge von ***HRISTOW*** (1942) und ***WELSCH*** (1969) auf der Grundlage von Taylor-Entwicklungen zielten in gleiche Richtung. Alle Vereinfachungen mußten aber mit Einschränkungen in der Reichweite erkauft werden.

Heute bereitet die Auswertung der Helmertschen Formeln keine unüberwindlichen rechentechnischen Schwierigkeiten mehr (***SCHNÄDELBACH***, 1982). Die Struktur des Lösungsalgorithmus bleibt jedoch kompliziert.

Eine Möglichkeit, die Aufgabe "Ellipsoidübergang, Netzverschiebung und Drehstreckung" mit einfacheren Strukturen und ohne Kenntnis des Differentialquotienten in (1) zu lösen, bieten die Funktionen (2), aus denen sie abgeleitet sind (Stammfunktionen) in Verbindung mit ihren Umkehrfunktionen

$$
\begin{aligned}
A_{oi} &= A\left(a, f, B_o, L_o, B_i, L_i\right) \\
S_{oi} &= S\left(a, f, B_o, L_o, B_i, L_i\right)
\end{aligned}
\qquad (5)
$$

Beide Funktionsgruppen sind algorithmisch hinreichend erschlossen, entsprechende Algorithmen stehen in großer Auswahl und allen Genauigkeitsansprüchen genügend zur Verfügung (***SCHÖDLBAUER***, 1981, S 18-71, 72-115). Der Weg über die Stammfunktionen hat darüberhinaus noch den Vorzug, daß die Lösung nicht auf differentielle Vorgänge beschränkt ist. Dieser Lösungsweg, der sich in drei Schritte gliedert, läßt sich wie folgt skizzieren:

Erster Schritt: Berechnung der Polarkoordinaten der Punkte P_i bezüglich eines Zentralpunktes P_o^I auf Ellipsoid I vermittels (5)

$$\begin{aligned} A_{oi}^I &= A\left(a^I, f^I, B_o^I, L_o^I, B_i^I, L_i^I\right) \\ S_{oi}^I &= S\left(a^I, f^I, B_o^I, L_o^I, B_i^I, L_i^I\right) \end{aligned} \tag{6}$$

Zweiter Schritt: Drehstreckung des Punktfeldes (der Punkte P_i^I) um den Punkt P_o^I, wobei die neuen Polarkoordinaten gleichzeitig als Polarkoordinaten des Ellipsoids II verstanden werden

$$\begin{aligned} A_{oi}^{II} &= A_{oi}^I + da \\ S_{oi}^{II} &= S_{oi}^I + S_{oi}^I \cdot dm \ . \end{aligned} \tag{7}$$

Dritter Schritt: Neukoordinierung des Punktfeldes auf Ellipsoid II von einem vorgegebenen Zentralpunkt P_o^{II} aus mit Hilfe von (2)

$$\begin{aligned} B_i^{II} &= B\left(a^{II}, f^{II}, B_o^{II}, L_o^{II}, A_o^{II}, S_o^{II}\right) \\ L_i^{II} &= L\left(a^{II}, f^{II}, B_o^{II}, L_o^{II}, A_o^{II}, S_o^{II}\right) \end{aligned} \tag{8}$$

Nun tritt die Aufgabenstellung in der Regel nicht in der durch die Formeln (3), (1), (4) bzw. (6), (7), (8) beschriebenen elementaren Form auf, sondern als Problem der "Anfelderung" oder im Zusammenhang mit einer Lotabweichungsausgleichung. In diesem Fall sind für bestimmte Punkte P_i des Punktfeldes (Verknüpfungspunkte) die auf unterschiedliche geodätische Grundlagen bezogenen Koordinaten B_i^I, L_i^I (Startsystem) und $\overline{B}_i^{II}, \overline{L}_i^{II}$ (Zielsystem) vorgegeben. Ellipsoidübergang, Netzverschiebung und Drehstreckung, d.h. Überführung der Koordinaten B_i^I, L_i^I des Systems I in solche des Systems II: $\overline{B}_i^{II}, \overline{L}_i^{II}$ sind dabei nach einem von ***HELMERT*** (1893) eingeführten Minimumsprinzip ("Helmert-

Transformation") so vorzunehmen, daß die Restklaffungen

$$v_{B_i} = \overline{B}_i^{II} - B_i^{II}$$
$$v_{L_i} = \overline{L}_i^{II} - L_i^{II} \quad , \tag{9}$$

die bei der Transformation der Koordinaten B_i^I, L_i^I in Koordinaten B_i^{II}, L_i^{II} wegen der Ungleichartigkeit der verglichenen Koordinaten zwangsläufig verbleiben, der Bedingung

$$\sum_{i=1}^{n} \left(v_{B_i}^2 + \cos^2 B_i \cdot v_{L_i}^2 \right) = \text{minimal} \tag{10}$$

genügen. Dies kann unter Zuhilfenahme des Algorithmus der Ausgleichung nach vermittelnden Beobachtungen über die "Fehlergleichungen"

$$v_{B_i} = \frac{\partial B_i}{\partial B_o} \cdot dB + \frac{\partial B_i}{\partial L_o} \cdot dL + \frac{\partial B_i}{\partial A_{oi}} \cdot dA + \frac{\partial B_i}{\partial m_{oi}} \cdot dm - l_{B_i}$$
Gewichte: 1
$$v_{L_i} = \frac{\partial L_i}{\partial B_o} \cdot dB + \frac{\partial L_i}{\partial L_o} \cdot dL + \frac{\partial L_i}{\partial A_{oi}} \cdot dA + \frac{\partial L_i}{\partial m_{oi}} \cdot dm - l_{L_i} \tag{11}$$
Gewichte: $\cos^2 B_i$

mit den "Beobachtungen"

$$-l_{B_i} = \frac{\partial B_i}{\partial a} \cdot da + \frac{\partial B_i}{\partial f} \cdot df + dB_i$$
$$dB_i = \overline{B}_i^{II} - B_i^I$$
$$-l_{L_i} = \frac{\partial L_i}{\partial a} \cdot da + \frac{\partial L_i}{\partial f} \cdot df + dL_i$$
$$dL_i = \overline{L}_i^{II} - L_i^I$$

und den Unbekannten (gekürzten Parametern)

dB_o, dL_o, dA und dm

erfolgen. Nach Ermittlung der Unbekannten auf dem üblichen Weg über Normalgleichungen fließen die gesuchten transformierten Koordinaten B_i^{II}, L_i^{II} der Punkte P_i aus den Gleichungen (1) und (4) ein.

Die Genauigkeit des Verfahrens wird maßgeblich von der Genauigkeit der in (11) und (1) auftretenden Differentialquotienten bestimmt, deren exakte Berechnung - wie bereits dargelegt - Schwierigkeiten bereitet.

Der im folgenden dargestellte neue Lösungsweg für den Ellipsoidübergang und die Helmert-Transformation auf Bezugsellipsoiden unterliegt dieser Einschränkung nicht. Er stützt sich auf den in (6), (7) und (8) niedergelegten strengen Lösungsansatz, der analog zu (9), (10) und (11) in einen Ausgleichungsalgorithmus eingebaut wird. Die Im Rechenablauf in (11) benötigten Differentialquotienten dienen dabei nur als Hilfsmittel in einem Iterationsprozeß, bestimmen aber - anders als bisher - nicht die Genauigkeit der Ergebnisse. Daher werden an ihre Genauigkeit keine hohen Ansprüche gestellt, so daß an Stelle der (komplizierten) ellipsoidischen die (wesentlich einfacheren) sphärischen Ausdrücke treten können.

Der Rechengang wird - analog zur Vorgehensweise nach Helmert - durch Festlegung eines Zentralpunktes P_o^I in der Mitte des Punktfeldes vorbereitet. Dieser Punkt übernimmt die Funktion des Poles eines Polarkoordinatensystems. Eine plausible Punktlage im Startsystem, also im Koordinatensystem des Ellipsoids I. Daher ergibt sich dabei zu

$$B_o^I = \frac{\sum_{i=1}^{n} B_i^I}{n}$$

$$L_o^I = \frac{\sum_{i=1}^{n} L_i^I}{n} \qquad (12)$$

Bezogen auf diesen Pol P_o^I werden nun im Startsystem die Polarkoordinaten A_{oi}^I, S_{oi}^I der Punkte und die entsprechenden Gegenazimute A_{io}^I nach (5) und (6) berechnet:

$$A^I_{oi} = A\left(a^I, f^I, B^I_o, L^I_o, B^I_i, L^I_i\right)$$

$$S^I_{oi} = S\left(a^I, f^I, B^I_o, L^I_o, B^I_i, L^I_i\right) \qquad (13)$$

$$A^I_{io} = A\left(a^I, f^I, B^I_i, L^I_i, B^I_o, L^I_o\right)$$

Diese auf das Ellipsoid I bezogenen Polarkoordinaten (einschließlich der Gegenazimute) werden sodann mit zunächst willkürlich gewähltem Pol, mit willkürlicher Orientierung und ohne Maßstabsänderung auf das Ellipsoid II übertragen:

$$\left(A^{II}_{oi}\right) = A^I_{oi}$$

$$\left(S^{II}_{oi}\right) = S^I_{oi} \qquad (14)$$

$$\left(A^{II}_{io}\right) \approx A^I_{io} \quad .$$

Setzt man den Zentralpunkt $\left(P^{II}_o\right)$ des Polarkoordinatensystems auf dem Ellipsoid II in Analogie zu (12) zunächst zu

$$\left(B^{II}_o\right) = \frac{\sum_{i=1}^{n} \overline{B}^{II}_i}{n}$$

$$\left(L^{II}_o\right) = \frac{\sum_{i=1}^{n} \overline{L}^{II}_i}{n} \qquad (15)$$

fest, so bestimmen (15) und (14) auf Ellipsoid II ein zum Startsystem analoges Punktfeld, dessen geographische Koordinaten sich über (2)/(8) zu

$$\left(B^{II}_i\right) = B\left(a^{II}, f^{II}, \left(B^{II}_o\right), \left(L^{II}_o\right), \left(A^{II}_{oi}\right), \left(S^{II}_o\right)\right)$$

$$\left(L^{II}_i\right) = L\left(a^{II}, f^{II}, \left(B^{II}_o\right), \left(L^{II}_o\right), \left(A^{II}_{oi}\right), \left(S^{II}_o\right)\right) \qquad (16)$$

ergeben. Der veränderten Geometrie des Bezugsellipsoids II wird dabei durch

die entsprechenden Ellipsoidkonstanten a^{II} und f^{II} Rechnung getragen.

Wegen der bei der Wahl des Zentralpunktes (P_o^{II}) sowie der Orientierung und des Maßstabes geübten Willkür, aber auch wegen der vorausgesetzten Ungleichartigkeit der Koordinatensysteme werden die Koordinaten (B_i^{II}), (L_i^{II}) nicht mit den Bezugskoordinaten $\overline{B}_i$, $\overline{L}_i$ übereinstimmen. Man wird zwischen diesen vielmehr die Ausgangsklaffungen

$$l_{B_i}^{II} = -dB_i^{II} = -\left(\overline{B}_i^{II} - (B_i^{II})\right)$$
$$l_{L_i}^{II} = -dL_i^{II} = -\left(\overline{L}_i^{II} - (L_i^{II})\right) \qquad (17)$$

feststellen.

Über den Weg einer Ausgleichung nach vermittelnden Beobachtungen können nun aus den Ausgangsklaffungen, die formal als Beobachtungen betrachtet werden, die Parameter des funktionalen Modells "Helmert-Transformation" abgeleitet werden, wobei die Verbesserungen die verbleibenden Restklaffungen beschreiben. Die entsprechenden Fehlergleichungen lauten:

$$v_{B_i}^{II} = \frac{\partial B_i}{\partial B_o}.dB_o^{II} + \frac{\partial B_i}{\partial L_o}.dL_o^{II} + \frac{\partial B_i}{\partial A_{oi}}.dA^{II} + \frac{\partial B_i}{\partial m_{oi}}.dm^{II} - l_{B_i}^{II}$$
Gewichte: 1

$$v_{L_i}^{II} = \frac{\partial L_i}{\partial B_o}.dB_o^{II} + \frac{\partial L_i}{\partial L_o}.dL_o^{II} + \frac{\partial L_i}{\partial A_{oi}}.dA^{II} + \frac{\partial L_i}{\partial m_{oi}}.dm^{II} - l_{L_i}^{II} \qquad (18)$$
Gewichte: $\cos^2 B_i$

Mit den unterschiedlichen Gewichten 1 und $\cos^2 B_i$ wird dem Umstand Rechnung getragen, daß den Winkelelementen l_{B_i} und l_{L_i} in geographischer Breite und Länge unterschiedliche Linearmaße auf der Ellipsoidoberfläche entsprechen, nämlich $\approx l_{B_i} \cdot R_o$ bzw. $l_{L_i} \cdot R_o$.

Wie bereits erwähnt, können die in (18) enthaltenen partiellen Differentialquotienten durch ihre rein sphärischen Näherungen ersetzt werden, die sich aus der Spezialisierung der Formeln von ***HELMERT*** (1880, S. 282, 283, 292, 293) oder - unabhängig davon - aus rein sphärischen Betrachtungen wie folgt ergeben, ohne daß Fehler im Endergebnis zu befürchten wären:

$$\frac{\partial B_i}{\partial B_o} = -\cos\frac{S_{oi}}{R_o}\cdot\sin A_{oi}\cdot\sin A_{io} - \cos A_{oi}\cdot\cos A_{io} = \cos(L_i - L_o)$$

$$\frac{\partial B_i}{\partial L_o} = 0$$

$$\frac{\partial B_i}{\partial A_{oi}} = \sin\frac{S_{oi}}{R_o}\cdot\sin A_{io} = -\cos B_o\cdot\sin(L_i - L_o)$$

$$\frac{\partial B_i}{\partial m_{oi}} = -\frac{S_{oi}}{R_o}\cdot\cos A_{io}$$

$$\frac{\partial L_i}{\partial B_o} = \tan B_i\cdot\sin(L_i - L_o)$$

$$\frac{\partial L_i}{\partial B_o} = 1$$

$$\frac{\partial L_i}{\partial A_{oi}} = -\frac{\sin(L_i - L_o)\cdot\cos A_{io}}{\sin A_{oi}}$$

$$\frac{\partial L_i}{\partial m_{oi}} = -\frac{S_{oi}}{R_o}\cdot\frac{\sin A\,io}{\cos B_i}$$

(19)

mit R_o = Gaußscher Krümmungsradius des Bezugsellipsoids in P_o.

Die gesuchten gekürzten Parameter dB_o^{II}, dL_o^{II}, dA^{II}, dm^{II} werden - dem Ausgleichungsalgorithmus entsprechend - über die dem System (18) entsprechenden Normalgleichungen gewonnen. Mit diesen Größen lassen sich die durch (14) festgelegten Näherungswerte (B_o^{II}), (L_o^{II}) für den Zentralpunkt (P_o^{II}) und für die Polarkoordinaten $\left(A_{oi}^{II}\right)$ und (S_{oi}^{II}) der Punkte P_i^{II} verbessern.

$$
\begin{aligned}
B_o^{II} &= \left(B_o^{II}\right) + dB_o^{II} \\
L_o^{II} &= \left(L_o^{II}\right) + dL_o^{II} \\
A_{oi}^{II} &= \left(A_{oi}^{II}\right) + dA^{II} \\
S_{oi}^{II} &= \left(S_{oi}^{II}\right) + \left(S_{oi}^{II}\right) \cdot dm^{II}
\end{aligned}
\qquad (20)
$$

Die neue, und in den meisten Fällen endgültige Lage der Punkte P_i^{II} - ausgedrückt im System der geographischen Koordinaten - findet man auf der Grundlage dieser Ergebnisse schließlich durch Auswertung von (8)

$$
\begin{aligned}
B_i^{II} &= B\left(a^{II}, f^{II}, B_o^{II}, L_o^{II}, A_{oi}^{II}, S_{oi}^{II}\right) \\
L_i^{II} &= L\left(a^{II}, f^{II}, B_o^{II}, L_o^{II}, A_{oi}^{II}, S_{oi}^{II}\right)
\end{aligned}
\qquad (21)
$$

Der Vergleich dieser Ergebnisse mit den Bezugskoordinaten $\overline{B}_i^{II}$, $\overline{L}_i^{II}$ der Verknüpfungspunkte $\overline{P}_i$ entsprechend (9) ergibt - konkurrierend mit (18) - die exakten Werte der Restklaffungen

$$
\begin{aligned}
v_{B_i}^{II} &= \overline{B}_i^{II} - B_i^{II} \\
v_{L_i}^{II} &= \overline{L}_i^{II} - L_i^{II} \quad .
\end{aligned}
\qquad (22)
$$

Der Abschluß der Berechnungen an dieser Stelle ist freilich nur möglich, wenn die linearisierten Fehlergleichungen (18) in Verbindung mit den genäherten Differentialquotienten (19) die nichtlinearem funktionalen Zusammenhänge zwischen den Ausgangsklaffungen (17) und den auf dem Weg über (21) ermittelten Restklaffungen (22) hinreichend genau beschreiben, was nur bei kleinen Punktfeldern und kleinen (differentiellen) Beträgen der Netzverschiebung, -drehung und -streckung zu erwarten und an der Identität der aus (18) und aus (22) fließenden Ergebnisse zu erkennen ist.

Bei ausgedehnten Punktfeldern und größeren Netzbewegungen wird diese Bedingung nicht mehr erfüllt sein. In diesem Fall müssen die Ergebnisse aus (21) analog zu denen aus (16) noch als vorläufig angesehen werden, was dazu zwingt, (17) bis (22) mit verbesserten Ausgangswerten $\left(B_i^{II}\right)$, $\left(L_i^{II}\right)$ erneut zu

durchlaufen. Die neuen Ausgangswerte liefern die Algorithmen (21), die damit an die Stelle von (16) treten

$$B_i^{II}(21) \rightarrow \left(B_i^{II}\right)(16)$$
$$L_i^{II}(21) \rightarrow \left(L_i^{II}\right)(16) \qquad (23)$$

Der Vorgang ist gegebenenfalls mehrmals zu wiederholen und zwar so lange, bis die Beträge der gekürzten Transformationsparameter dB_o^{II}, dL_o^{II}, dA^{II}, dm^{II} "verschwinden", d.h. vorgegebene numerischen Schranken unterschreiten:

$$\left|dB_o^{II}\right| < \Delta dB \quad (\Delta dB \text{ z.B. } 10^{-5}")$$

$$\left|dL_o^{II}\right| < \Delta dL \quad (\Delta dL \text{ z.B. } 10^{-5}"/\cos B_i)$$

$$\left|dA^{II}\right| < \Delta dA \quad (\Delta dA \text{ z.B. } R_o \cdot 10^{-5}"/S_{oi\,max})$$

$$\left|dm^{II}\right| < \Delta dm \quad \left(\Delta dm \text{ z.B. } R_o \cdot 10^{-5}"/(S_{oimax} \cdot \rho")\right) .$$

Die durch den Iterationsprozeß bewirkte Konvergenz dieser Werte nach Null, läßt es zu, die Genauigkeitsansprüche an die Differentialquotienten der Fehlergleichungen (18) niedrig zu halten. Die Näherungswerte, die sich aus den sphärischen Ausdrücken (19) ergeben, werden diesen Ansprüchen ohne weiteres gerecht. Die Vernachlässigungen aufgrund der Substitution der strengen ellipsoidischen Differentialquotienten durch (19) wirken sich allenfalls auf die Konvergenz des Iterationsprozesses (ohne sie dabei zu gefährden oder wesentlich zu beeinträchtigen), nicht aber auf die Genauigkeit der Ergebnisse aus, die von den strengen Beziehungen (21) bestimmt wird.

RECHENBEISPIEL

Das folgende Rechenbeispiel stützt sich auf sechs Punkte, die sowohl dem Deutschen Hauptdreiecksnetz als auch dem Europäischen Dreiecksnetz 1950 angehören. Die zugeordneten Koordinaten beziehen sich demnach zum einen auf das Bessel-Ellipsoid mit Datum Potsdam (System I), zum anderen auf das Internationale Ellipsoid mit Europa-Datum 1950 (System II). Die Koordinaten-Transformation erfolgte dabei von System I in System II, wobei zwei Iterationsschritte erforderlich waren.

Die Berechnungen dienten im vorliegenden Fall nicht der Netzanalyse, sondern allein der anschaulichen und beispielhaften Erläuterung des Rechenablaufs. Diese Zielsetzung legte eine Beschränkung der Anzahl der Verknüpfungspunkte auf nur sechs nahe.

Modell-Matrix "Helmerttransformation"
(berechnet aus den Elementen der Tabelle "0. Iteration")

	A				diag P
i	$\left(\frac{\partial B_i}{\partial B_o}\right)$	$\left(\frac{\partial B_i}{\partial L_o}\right)$	$\left(\frac{\partial B_i}{\partial A_{oi}}\right)$	$\left(\frac{\partial B_i}{\partial m_{oi}}\right)$	1
	$\left(\frac{\partial L_i}{\partial B_o}\right)$	$\left(\frac{\partial L_i}{\partial L_o}\right)$	$\left(\frac{\partial L_i}{\partial A_{oi}}\right)$	$\left(\frac{\partial L_i}{\partial m_{oi}}\right)$	$\cos^2 B_i$
1	0,998 404	0	- 0,035 626	- 0 056 136	1
	0,062 066	1	0,083 258	0,053 050	0,452 95
2	0,999 910	0	0,008 487	- 0,052 456	1
	- 0,014 875	1	- 0,078 081	- 0,012 676	0,450 05
3	0,999 936	0	- 0,007 125	0,000 577	1
	0,013 904	1	0,000 913	0,011 317	0,449 18
4	0,999 059	0	0,021 368	- 0,004 215	1
	- 0,052 936	1	0,075 393	- 0,043 242	0,401 74
5	0,999 743	0	0,014 281	0,044 963	1
	- 0,030 556	1	0,075 393	- 0,024 034	0,354 37
6	0,999 916	0	- 0,007 377	0,065 871	1
	0,016 490	1	0,113 640	0,112 779	0,334 74

Ausgangslage System I

i	B_i^I L_i^I	B_o^I L_o^I	A_{oi}^I S_{oi}^I	
1	***47°42'15,568 18"***	50°52'58,152 49"	161,210 671 24 gon	
	12°00'48,912 54"	08°46'33,388 70"	424 620,676 m	
2	***47°52'28,170 07"***		210,879 866 56 gon	
	08°00'18,222 95"		339 183,910 m	
3	***50°55' 3,703 83"***		94,309 096 99 gon	
	09°25'23,295 88"		45 688,309 m	
4	***50°40' 2,593 58"***		292,426 389 20 gon	
	06°17'21,685 12"		176 994,337 m	
5	***53°28'10,795 29"***		381,527 652 30 gon	
	07°28'42,998 45"		301 162,523 m	
6	***54°39'48,084 01"***		6,520 426 47 gon	
	09°26'45,217 24"		423 073,764 m	

0. Iteration

	dA^{II}	dB_o^{II}	(B_o^{II})	
	0	0	50°52'56,496 85"	
	dm^{II}	dL_o^{II}	(L_o^{II})	
	0	0	08°46'33,933 71"	
i	$(A_{oi}^{II}) = A_{oi}^I$ $(S_{oi}^{II}) = S_{oi}^I$	$\overline{B}_i^{II}$ $\overline{L}_i^{II}$	(B_i^{II}) (L_i^{II})	$l_{B_i} = -dB_i^{II}$ $l_{L_i} = -dL_i^{II}$
1	161,210 671 24 gon	***47°42'15,495 77"***	47°42'15,640 80"	0,145 03"
	424 620,676 m	***12° 0'47,418 95"***	12°00'47,497 58"	0,078 63"
2	210,879 866 56 gon	***47°52'28,012 39"***	47°52'28,125 20"	0,112 81"
	339 183,910 m	***08° 0'19,052 68"***	08°00'19,236 34"	0,183 66"
3	94,309 096 99 gon	***50°55' 2,008 97"***	50°55' 2,030 51"	0,021 54"
	45 688,309 m	***09°25'23,392 44"***	09°25'23,421 70"	0,029 26"
4	292,426 389 20 gon	***50°40' 1,120 59"***	50°40' 1,072 23"	- 0,048 36"
	176 994,337 m	***06°17'23,629 93"***	06°17'23,831 51"	0,201 58"
5	381,527 652 30 gon	***53°28' 7,877 97"***	53°28' 7,729 07"	- 0,148 90"
	301 162,523 m	***07°28'44,502 83"***	07°28'44,435 17"	- 0,067 66"
6	6,520 426 47 gon	***54°39'44,465 43"***	54°39'44,351 26"	- 0,114 17"
	423 073,764 m	***09°26'45,605 40"***	09°26'45,287 80"	- 0,317 60"

1. Iteration

	dA^{II}	dB_o^{II}	(B_o^{II})	
	0,644 11 mgon	0,006 15"	50°52'56,503 00"	
	dm^{II}	dL_o^{II}	(L_o^{II})	
	$1{,}0360 \cdot 10^{-5}$	- 0,025 72"	08°46'33,907 99"	
i	(A_{oi}^{II}) (S_{oi}^{II})	$\overline{B}_i^{II}$ $\overline{L}_i^{II}$	(B_i^{II}) (L_i^{II})	$l_{B_i} = -dB_i^{II}$ $l_{L_i} = -dL_i^{II}$
1	161,211 315 35 gon	***47°42'15,495 77"***	47°42'15,452 17"	- 0,043 60"
	424 625,075 m	***12° 0'47,418 95"***	12°00'47,411 72"	- 0,007 23"
2	210,880 gon	***47°52'28,012 39"***	47°52'28,036 82"	0,024 43"
	339 187,424 m	***08° 0'19,052 68"***	08°00'19,020 53"	- 0,032 15"
3	94,309 741 10 gon	***50°55' 2,008 97"***	50°55' 2,022 98"	0,014 01"
	45 688,782 m	***09°25'23,392 44"***	09°25'23,422 12"	0,029 68"
4	292,427 033 31 gon	***50°40' 1,120 59"***	50°40' 1,126 62"	0,006 03"
	176 996,171 m	***06°17'23,629 93"***	06°17'23,699 33"	0,069 40"
5	381,528 296 41 gon	***53°28' 7,877 97"***	53°28' 7,861 25"	- 0,016 72"
	301 165,643 m	***07°28'44,502 83"***	07°28'44,515 33"	0,012 50"
6	6,521 070 58 gon	***54°39'44,465 43"***	54°39'44,482 83"	0,017 40"
	423 078,147 m	***09°26'45,605 40"***	09°26'45,526 56"	- 0,078 84"

2. Iteration (definitive Berechnung)

	dA^{II}	dB_o^{II}	$(B_o^{II}) = dB_o^{II}$	
	- 0,000 50 mgon	0,000 04"	50°52'56,503 04"	
	dm^{II}	dL_o^{II}	$(L_o^{II}) = dL_o^{II}$	
	$-1{,}3 \cdot 10^{-7}$	0,000 00"	08°46'33,907 99"	
i	$(A_{oi}^{II}) = A_{oi}^{I}$ $(S_{oi}^{II}) = S_{oi}^{I}$	$\overline{B}_i^{II}$ $\overline{L}_i^{II}$	$(B_i^{II}) = B_i^{II}$ $(L_i^{II}) = L_i^{II}$	$l_{B_i} = -dB_i^{II} = v_{B_i}$ $l_{L_i} = -dL_i^{II} = v_{L_i}$
1	161,211 314 85 gon	***47°42'15,495 77"***	***47°42'15,452 42"***	- 0,043 35"
	424 625,020 m	***12° 0'47,418 95"***	***12°00'47,411 72"***	- 0,007 23"
2	210,880 510 17 gon	***47°52'28,012 39"***	***47°52'28,036 98"***	0,024 59"
	339 187,380 m	***08° 0'19,052 68"***	***08°00'19,020 69"***	- 0,031 99"
3	94,309 740 60 gon	***50°55' 2,008 97"***	***50°55' 2,023 03"***	0,014 06"
	45 688,776 m	***09°25'23,392 44"***	***09°25'23,427 09"***	0,034 65"
4	292,427 032 81 gon	***50°40' 1,120 59"***	***50°40' 1,126 63"***	0,006 04"
	176 996,148 m	***06°17'23,629 93"***	***06°17'23,699 45"***	0,069 52"
5	381,528 295 91 gon	***53°28' 7,877 97"***	***53°28' 7,861 16"***	- 0,016 81"
	301 165,604 m	***07°28'44,502 83"***	***07°28'44,515 26"***	0,012 43"
6	6,521 070 08 gon	***54°39'44,465 43"***	***54°39'44,482 72"***	0,017 29"
	423 078,092 m	***09°26'45,605 40"***	***09°26'45,526 34"***	- 0,079 06"

LITERATUR

HELMERT, F.R.: ***Die mathematischen und physikalischen Theorien der Höheren Geodäsie, Teil I.*** Teubner Verlag, Leipzig, 1880

HELMERT, F.R.: ***Die Europäische Längengradmessung in 52° Breite von Greenwich bis Warschau, I. Heft, Hauptdreiecke und Grundlinienanschlüsse von England bis Polen.*** Veröffentlichung des Königlich Preußischen Geodätischen Instituts und Centralbureaus der Internationalen Erdmessung, S. 47-50, Verlag Stankiewicz, Berlin, 1893

HRISTOW, W.: ***Änderung der geographischen Koordinaten infolge Umorientierung eines geodätischen Netzes und Übergang zum anderen Referenzellipsoid.*** Zeitschrift für Vermessungswesen 71, S. 132-140, 1942

JORDAN, W./EGGERT, O./KNEISSL, M.: ***Handbuch der Vermessungskunde, Band IV, zweite Hälfte.*** Metzlersche Verlagsbuchhandlung, Stuttgart, 1959

SCHNÄDELBACH, K.: ***On Free Net Adjustment on the Ellipsoid. Proceedings of the International Symposium on Geodetic Networks and Computations of the International Association of Geodesy, Munich, August 31 to September 5, 1981.*** Veröffentlichung der Deutschen Geodätischen Kommission bei der Bayerischen Akademie der Wissenschaften, Reihe B, Heft Nr. 258/VI, S. 75-82, München, 1982

SCHÖDLBAUER, A.: ***Rechenformeln und Rechenbeispiele zur Landesvermessung, Teil 1. Die geodätischen Grundaufgaben auf Bezugsellipsoiden im System der geographischen Koordinaten und die Bewertung ellipsoidischer Dreiecke.*** Wichmann-Skripten, Heft 2, Wichmann Verlag, Karlsruhe, 1981

SCHÖDLBAUER, A.: ***Rechenformeln und Rechenbeispiele zur Landesvermessung, Teil 2. Geodätische Berechnungen im System der Gaußschen konformen Abbildung eines Bezugsellipsoids unter besonderer Berücksichtigung des Gauß-Krüger- und des UTM-Koordinatensystems im Bereich der Bundesrepublik Deutschland.*** Wichmann-Skripten, Heft 2, Wichmann Verlag, Karlsruhe, 1982

WELSCH, W.: ***Beiträge zur Transformation geodätischer geographischer Koordinaten nach Hristow.*** Dissertation (TH München), München, 1969

UNTERSUCHUNG UND EINSATZ VON POTENTIOMETRISCHEN WEGAUFNEHMERN

von Wilfried ELLMER

SUMMARY

Besides other linear transducers precision potentiometers can be employed for the measurement of small linear displacements. The paper deals with the special features of these potentiometers in comparison with other transducers. A series of measurements is analysed to identify advantages and disadvantages of the former and to find the best conditions for the application of such linear potentiometers. A practical example is given.

ZUSAMMENFASSUNG

Zur kontinuierlichen Messung kleiner Wegänderungen können neben anderen Wegaufnehmern auch Präzisionspotentiometer eingesetzt werden. Die Eigenschaften dieser Aufnehmer werden mit denen anderer Systeme verglichen. Aufgrund von Laboruntersuchungen können wesentliche Vor- und Nachteile dargestellt werden, und es wird erläutert, unter welchen Bedingungen mit Potentiometerwegaufnehmern optimale Ergebnisse zu erzielen sind. Der Bericht wird durch ein Anwendungsbeispiel abgeschlossen.

1. EINFÜHRUNG

Verfolgt man die Literatur über geodätische Deformationsmessungen, so fällt auf, daß zur kontinuierlichen Messung kleiner Wegänderungen überwiegend induktive Wegaufnehmer eingesetzt wurden. Andere Wegaufnehmer fanden nur in Einzelfällen Anwendung. In der Meßtechnik, die im Bereich der Elektrotechnik und des Maschinenbaus sehr weit entwickelt ist, findet sich eine große Auswahl von Wegaufnehmern mit sehr unterschiedlichen Eigenschaften. Es ist wichtig, diese Eigenschaften zu kennen, um im konkreten Anwendungsfall den günstigsten Aufnehmertyp einsetzen zu können. Ein Vergleich dieser Charakteristika läßt erwarten, daß Präzisionspotentiometer den induktiven Wegaufnehmern gegenüber einige Vorteile bei der ingenieurgeodätischen Anwendung aufweisen. Am Institut für Geodäsie wurden einige Potentiometer im Labor untersucht und in der Praxis eingesetzt.

2. POTENTIOMETERWEGAUFNEHMER UND IHRE BESONDEREN EIGENSCHAFTEN

2.1 Prinzipielle Wirkungsweise

Potentiometerwegaufnehmer gehören zur Gruppe der analogen Wegaufnehmer. Bei dieser Gruppe wird die Eingangsgröße "Weg" auf eine kontinuierliche Ausgangsgröße abgebildet, während bei digitalen Aufnehmern einem Wertebereich der Meßgröße ein eindeutiger Ausgangswert (etwa eine Zahl) zugeordnet wird. Bei Potentiometerwegaufnehmern wird die Ausgangsgröße mit Hilfe eines ohmschen Widerstandes gebildet. Andere analoge Wegaufnehmer arbeiten mit den Größen Induktivität, Kapazität, Frequenz usw.. Näheres dazu lese man bei ***KAHMEN*** (1978), ***ROESEMANN*** (1976) und ***BOCK*** (1982).

Das Meßprinzip beruht auf der variablen Teilung einer Spannung entlang eines Widerstandes. Dieser Widerstand R_S habe die Länge l (Abb. 1). Die Wegabhängigkeit wird durch einen Schleifer realisiert, der die Leiterbahn in die Stücke l_1 und l_2 teilt. Ist der Widerstandsbelag c konstant, so gilt:

$$\begin{aligned} R_1 &= c \cdot l_1 \\ R_2 &= c \cdot l_2 \\ R_S &= c \cdot l \quad . \end{aligned} \tag{2-1}$$

Ist U die Versorgungsspannung und U_1 die an der Strecke l_1 abgegriffene Spannung, so gilt unter der Voraussetzung, daß der Spannungsteiler unbelastet ist,

$$U_1 = U \frac{R_1}{R_S} = U \frac{l_1}{l} \quad . \tag{2-2}$$

Liegt jedoch an U_1 ein Verbraucher mit relativ niedrigem Widerstand an, so ist wegen des hohen Stromes, der am Abgriff B abfließt, die Beziehung zwischen U_1 und l_1 nicht linear.

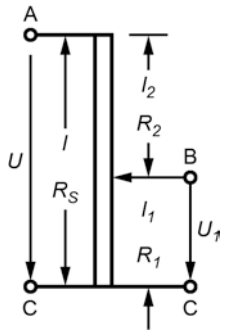


Abb. 1 Unbelasteter Spannungsteiler

2.2 Aufbau

Ihrem äußeren Aufbau nach bestehen die Potentiometerwegaufnehmer ähnlich wie die induktiven Wegaufnehmer (Differentialtransformatoren) aus einem runden oder kantigen Körper und einer Schubstange. Der Meßbereich dieser Wegaufnehmer liegt etwa zwischen 12 mm und 1 m. Der Körper ist etwa 100 bis 200 mm länger als der Meßbereich und 20 bis 35 mm dick. Die Befestigung des Körpers am Meßobjekt erfolgt meist durch Flansche oder Schellen, während die Schubstange an ihrem Ende gewöhnlich ein Schraubgewinde hat. Daneben gibt es noch Ausführungen mit Feder und Tastspitze sowie selbsttragende Wegaufnehmer, die mit ihren Enden an Kugelköpfen aufzuhängen sind.

Im inneren Aufbau unterscheidet man nach dem Material, aus dem die Widerstandsbahn gefertigt ist, zwei Gruppen. Darüberhinaus gibt es noch eine Mischform. Die herkömmliche Art ist die des Drahtpotentiometers, bei dem um einen Kern herum ein feiner Draht gewickelt ist. Anfang der 60-er Jahre begann man, Präzisionspotentiometer zu entwickeln, bei denen die Drahtwick-

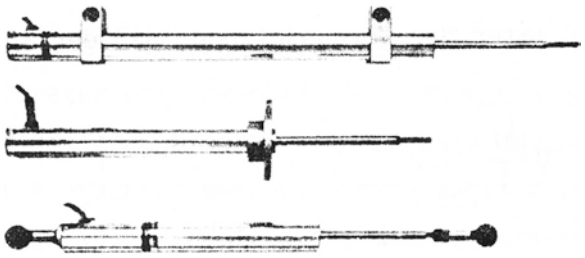


Abb. 2 Ausführungsformen von Potentiometerwegaufnehmern

lung durch eine Schicht aus Leitplastik ersetzt wurde. Dabei wurden je nach Anforderung (hohe Lebensdauer oder rauhe Umgebungsbedingungen) verschiedene Techniken angewandt. In der Praxis wird meistens eine Herz-Kohlenstoff-Mischung auf einem Trägerfilm verwendet. Zwischen diesen beiden Ausführungsformen steht das Hybrid-Potentiometer. Hier ist auf einer Drahtwicklung eine Leitplastik-Schicht aufgebracht (***ROESEMANN*** 1976, ***GASS*** 1978).

2.3 Eigenschaften im Vergleich mit anderen Wegaufnehmersystemen

Bei der Auswahl eines für die jeweilige Meßaufgabe geeigneten Aufnehmertyps sind die besonderen Vor- und Nachteile gegenüber anderen Typen zu berücksichtigen. Dabei können folgende Kriterien in Betracht gezogen werden:

- Auflösungsvermögen
- Meßlänge
- Linearität
- Verfahrgeschwindigkeit
- Rückwirkung auf das Meßobjekt
- Stromversorgung
- Meßwerterfassung
- Umwelteinflüsse
- Empfindlichkeit gegenüber Störimpulsen
- Zuverlässigkeit, Lebensdauer
- Kosten

Einige dieser Kriterien beeinflussen die Genauigkeit der Meßergebnisse. Dabei ist unter Genauigkeit das Zusammenwirken aller Glieder der Meßkette zu

verstehen, angefangen mit dem Meßgrößenaufnehmer über eventuell vorhandene Umformer bis hin zur Meßwertausgabe. Was nun das einzelne Glied der Meßkette betrifft, so ist es auch hier nicht eine einzelne Eigenschaft, wie z.B. das Auflösungsvermögen oder die Temperaturempfindlichkeit, die die Genauigkeit bestimmt, sondern es ist die Gesamtheit aller Einflüsse, die sich je nach den Einsatzbedingungen eines Meßgerätes verschieden stark auswirken. Dementsprechend erhält man aus Firmenprospekten auch keine Angaben über die Genauigkeit eines Gerätes, wie dies bei geodätischen Instrumenten üblich ist.

2.3.1 Auflösungsvermögen

Was das Auflösungsvermögen betrifft, so muß unterschieden werden zwischen endlicher und (zumindest theoretisch) unendlicher Auflösung. Drahtpotentiometer sowie alle digital arbeitenden Aufnehmer (z.B. Inkrementalmaßstäbe) haben eine endliche Auflösung. Sie hängt bei Drahtpotentiometern von der Dicke des Drahtes ab und liegt zwischen 2 und 100 µm. Dagegen haben Leitplastikpotentiometer, induktive und kapazitive Meßsysteme eine theoretisch unendlich feine Auflösung. Hier wird die Genauigkeit allein durch die Linearität und durch äußere Einflüsse bestimmt.

2.3.2 Meßlänge

Potentiometerwegaufnehmer werden in sehr unterschiedlichen Meßbereichen angeboten. Die kleinsten gibt es für 12,5 mm, die längsten liegen bei 1 m. Damit liegen sie etwas über dem Bereich der induktiven Wegaufnehmer, die ihrerseits wiederum ganz kurze Wege bis unter 1 mm sehr genau erfassen können. Längere Wege können durch optische Wegaufnehmer gemessen werden, z.B. werden Inkrementalmaßstäbe angeboten, die aus mehreren Teilstücken zusammengesetzt eine Länge von 20 m abdecken können.

2.3.3 Linearität

Die Linearität eines Aufnehmers hängt zum Teil vom Meßprinzip, zum Teil von der Fertigungsgenauigkeit ab. Potentiometerwegaufnehmer arbeiten, soweit der Schleiferstrom gering genug ist, theoretisch völlig linear. Im Gegensatz dazu haben z.B. induktive Wegaufnehmer nur eine begrenzte Linearität, die umso günstiger wird, je mehr man den Meßbereich einschränkt. Andere Aufnehmer haben von ihrem Meßprinzip her eine nichtlineare Kennlinie, z.B.

Schwingsaitenwegaufnehmer. Aber in allen Fällen ist die Linearität von der Genauigkeit, mit der ein Aufnehmer gefertigt wird, abhängig. Die Grenze wird bei Potentiometerwegaufnehmern mit 0,1 oder 0,2 % des Meßbereiches angegeben.

2.3.4 Verfahrgeschwindigkeit

Die Verfahrgeschwindigkeit ist ein Maß dafür, ob ein Aufnehmer für dynamische Aufgaben geeignet ist. Sie gibt an, bis zu welcher Geschwindigkeit einwandfreie Meßergebnisse erzielt werden können. Beim Präzisionspotentiometer wird sie wesentlich durch den Schleiferdruck mitbestimmt. Je fester der Schleifer auf die Leiterbahn drückt, desto schneller kann noch gearbeitet werden, jedoch wird dadurch die Lebensdauer kürzer und die Meßkräfte größer. Nach Angaben einiger Firmen liegt die Verfahrgeschwindigkeit etwa bei 1 m/s. Andere Firmen geben an, daß Bewegungen bis zu 200 Hz bei 1 mm oder 300 Hz bei 0,5 mm Schwingungsamplitude gemessen werden können. Damit sind schnellere Messungen möglich als z.B. Messungen mit Schwingsaitenaufnehmern, die nur quasistatische Bewegungen von einigen Hz erfassen können. Auch optische Systeme, wie Inkrementalmaßstäbe, arbeiten etwas langsamer als Potentiometer. Für noch schnellere Messungen muß auf induktive oder kapazitive Meßsysteme zurückgegriffen werden. Bei induktiven Aufnehmern kann mit Geschwindigkeiten bis zu maximal einem Zehntel der Trägerfrequenz (bis mehrere kHz) gemessen werden (***HERCEG*** 1976, S. 4-8).

2.3.5 Rückwirkung auf das Meßobjekt

Vollkommen berührungslos arbeitende Meßsysteme beruhen auf induktiven und kapazitiven Meßverfahren. Allerdings sind dem Meßbereich Grenzen gesetzt, und Umwelteinflüsse sind schwierig abzuschirmen. Darüberhinaus gibt es weitere, optische Meßsysteme, die mit positionsempfindlichen Dioden oder Diodenfeldern arbeiten. Verhältnismäßig reibungsarm sind auch Spulen mit Eisenkern und entsprechend gebaute Kondensatoren.

Die meisten der übrigen Meßsysteme verlangen dagegen relativ hohe Verschiebungskräfte, sei es, weil sie stark abgedichtet werden müssen, sei es, um den Schleiferdruck eines Potentiometers zu überwinden. So werden z.B. für Inkrementalmaßstäbe als Meßtaster 0,1 bis 4 N und als seitlich geführte Längenmeßsysteme bis 10 N angegeben. Für Präzisionspotentiometer werden

Kräfte im Bereich von 0,2 bis 2 N angenommen.

2.3.6 Stromversorgung

Die Spannungsversorgung kann einen wesentlichen Einfluß auf die Wirtschaftlichkeit eines Meßsystems haben. Dabei können folgende Kriterien ausschlaggebend sein:

- Gleich- oder Wechselspannung
- frei wählbare oder vorgegebene Spannung
- Einfluß von Spannungsänderungen auf das Meßsignal
- Empfindlichkeit des Systems gegen Spannungsabfall

Präzisionspotentiometer sind in dieser Hinsicht verhältnismäßig flexibel. Sie können sowohl mit Gleich- als auch mit Wechselspannung betrieben werden, eine feste Spannung ist nicht vorgesehen (bis etwa 10 V, teilweise auch über 100 V) und die gemessenen Werte ändern sich nach einem Spannungsausfall nicht. Allerdings hängt das Meßsignal bei den üblicherweise angegebenen Schaltungen linear von der Versorgungsspannung ab. Somit wird eine stabilisierte Spannungsquelle benötigt; doch sind auch Schaltungen denkbar, bei denen dieser Einfluß berücksichtigt werden kann.

Andere Systeme sind weniger flexibel. Induktive und kapazitive Aufnehmer benötigen stets eine Versorgung mit Wechselspannung bei sonst ähnlichen Eigenschaften. Sie werden jedoch auch mit eingebautem Oszillator und Demodulator angeboten, die wiederum den Nachteil haben, nur bei einer bestimmten Gleichspannung, z.B. 5 V, einwandfrei zu arbeiten. Digitale Systeme haben meist eine eingebaute Spannungsversorgung, die mit einer fest vorgegebenen Wechselspannung (z.B. 220 V) zu speisen ist. Nachteilig kann sein, daß diese Systeme oft gegen Spannungsspitzen und Netzausfälle recht empfindlich sind. Insbesondere ist bei inkrementell arbeitenden Maßstäben der Nullpunkt neu zu definieren.

2.3.7 Meßwerterfassung

Die Art der Meßwerterfassung hängt von der Art des ausgegebenen Signals (digital oder analog) ab. Digitale Systeme werden in der Regel mit einem Steuergerät angeboten, welches auch eine direkte Anzeige ermöglicht. Zur Weiterverarbeitung oder Fernübertragung des Meßsignals steht meist ein Datenausgang zur Verfügung, an den auch Geräte anderer Hersteller angeschlos-

sen werden können, unter der Voraussetzung, daß die Schnittstelle einheitlich definiert wurde.

Die Meßsignale analoger Wegaufnehmer haben meist einen so hohen Signalpegel, daß sie gut an einem einfachen Meßinstrument (Voltmeter o.a.) ablesbar sind. Das gilt auch für Präzisionspotentiometer, soweit dabei ein bestimmter Schleiferstrom noch überschritten wird. Eventuell müssen geeignete Anpasser zwischengeschaltet werden. Die Weiterverarbeitung von Meßwerten ist mit Hilfe von Meßwerterfassungsanlagen möglich, soweit geeignete Eingänge vorhanden sind.

2.3.8 Umwelteinflüsse

Die Empfindlichkeit gegen Umwelteinflüsse aller Art ist nicht allein eine Frage des Meßprinzips, sondern auch eine Frage der Ausführung des einzelnen Aufnehmertyps. Gerade Potentiometerwegaufnehmer werden in sehr unterschiedlicher Bauweise geliefert, auch von ein und derselben Firma. Das gilt ganz besonders für kapazitive Wegaufnehmer oder für Inkrementalmaßstäbe, die vom Meßsystem her sehr empfindlich gegen Staubeinwirkung sind, jedoch bei guter Abschirmung ähnlich zuverlässig arbeiten wie andere Aufnehmer auch.

Der wichtigste Einfluß für alle Aufnehmer ist die Umgebungstemperatur. Im Gegensatz zu den digitalen ist bei analogen Systemen davon nicht nur der Aufnehmer selbst, sondern das gesamte Meßsystem betroffen.

2.3.9 Empfindlichkeit gegenüber Störimpulsen

Eine besondere Gruppe der Umwelteinflüsse ist die der elektrischen Störungen. Digitale Meßsysteme haben dabei den Vorteil, daß die Signalverfälschungen in der Regel (außer bei Inkrementalmaßstäben) so deutlich wird, daß sie als grob fehlerhaft erkennbar sind. Das gilt vor allem für absolut arbeitende Meßsysteme wie schrittmotorgesteuerte Meßspindeln (ELMETER) oder codierte Lineale.

Analoge Systeme sind gegen elektrische Störungen meistens weniger empfindlich, dafür ist die Verfälschung nicht so deutlich am Meßsignal zu erkennen. Insbesondere können induktive und kapazitive Meßsysteme auf elektromagnetische Felder empfindlich reagieren. Potentiometerwegaufnehmer, insbesondere Leitplastikpotentiometer, sind dagegen unempfindlich.

2.3.10 Lebensdauer

Durch die vom Meßprinzip vorgegebene mechanische Abnutzung ist die Lebensdauer der Potentiometerwegaufnehmer im Gegensatz zu der von anderen Aufnehmern begrenzt. Sie beträgt bei Drahtpotentiometern je nach den Umgebungsbedingungen und der Dicke des Drahtes 2 bis 10 Millionen Hübe. Hierin liegt ein wesentlicher Vorteil der Leitplastikpotentiometer, deren Lebensdauer je nach Ausführung zwischen 10 und 250 Millionen Hüben liegt. Die der Hybridpotentiometer liegt in ihrer Größenordnung dazwischen. Ein weiterer, die Lebensdauer bestimmender Faktor ist der Schleiferdruck und die Auswirkung der Umgebung, vor allem der Temperatur. Schließlich kann noch die Lebensdauer durch fehlerhafte Bedienung in Form von elektrischer Überlastung eingeschränkt werden. Leitplastikpotentiometer vertragen nur geringe Schleiferströme, meist nur 1 µA, weshalb sie immer als Potentiometer und nicht als variabler Widerstand geschaltet werden.

3. LABORUNTERSUCHUNG VON POTENTIOMETERWEGAUFNEHMERN

Die erwähnten Eigenschaften machen die Anwendung von Präzisionspotentiometern im Bereich geodätischer Deformationsmessungen interessant. Verschiedene Präzisionspotentiometer wurden am Institut für Geodäsie untersucht (***BRAUER*** 1982).

3.1 Ziele der Untersuchung

Bei einem idealen Potentiometerwegaufnehmer lautet die Kennlinie:

$$U = a + b\,x \quad , \tag{3-1}$$

wobei U die Ausgangsspannung, x der Weg und a und b Konstante sind. In der Praxis jedoch treffen diese Annahmen (Linearität und Konstanz der Parameter) aus den oben genannten Gründen nicht immer zu. Ziel der Untersuchung war es dabei, herauszufinden, welche Größen die Meßsignale am meisten beeinflussen, wie stark dieser Einfluß im Einzelfall ist und ob es sinnvoll ist, diesen Einfluß (z.B. durch eine nichtlineare Kennlinie) nachträglich zu berücksichtigen.

3.2 Aufnehmer und Versuchsaufbauten

Für die Untersuchungen standen insgesamt sechs Potentiometeraufnehmer und zum Vergleich ein induktiver Aufnehmer zur Verfügung.

Nr.	Typ	Meßlänge	Anzahl und Art der Spuren
Potentiometer-Wegaufnehmer der Firma Penny & Giles			
1	HLP 129	100 mm	1 Leitplastik-Hybrid
2	HLP 190	250 mm	2 Leitplastik-Hybrid
3	HLP 350	1000 mm	2 Leitplastik-Hybrid
4	LP 21 C/F	914,4 mm	2 Draht
Potentiometer-Wegaufnehmer der Firma Novotechnik			
5	LFW 6-150.1	150 mm	1 Leitplastik-Film
6	LFW 36-900.1	900 mm	1 Leitplastik-Film
Induktiver Wegaufnehmer der Firma Collins			
7	SS 109	102 mm	1 Differentialtransformator

Tab. 1 Untersuchte Wegaufnehmer

Es waren Wegaufnehmer in den Längen zwischen 100 und 1000 m, von denen einer ein Drahtpotentiometer, drei Leitplastik-Hybrid-Potentiometer und zwei Leitplastik-Film-Potentiometer waren. Der induktive Wegaufnehmer (betrieben mit exakt 6 V Gleichspannung) wurde zum Vergleich mit herangezogen. Anzumerken ist noch, daß zwei Hybridpotentiometer und das Drahtpotentiometer je zwei Bahnen haben, die elektrisch völlig unabhängig arbeiten und daher auch getrennt voneinander untersucht werden mußten.

Die Erfassung der Meßwerte wurde mit einer Meßwerterfassungsanlage der Firma COMPULOG durchgeführt. Diese Anlage ist in erster Linie für Widerstandsmessungen konzipiert und in der Lage, die Wegaufnehmer mit einem kurzzeitigen Spannungsimpuls selbst zu versorgen. Lediglich der induktive Wegaufnehmer mußte mit einem externen Spannungskonstanter betrieben werden. Die Meßwerterfassungsanlage ist in Fortran IV frei programmierbar, so daß unmittelbar im Anschluß an die Erfassung der Messungen auch die Auswertung möglich ist.

Die Untersuchungen gliedern sich im wesentlichen in zwei Gruppen:

- Messungen an der Prüfstrecke
- Temperaturuntersuchungen

Die Messungen an der Prüfstrecke dienen zunächst der Ermittlung der Kennlinie eines einzelnen Aufnehmers, dann aber auch der mehr allgemeinen Beobachtung einiger Eigenschaften wie Linearität, Hysterese, Reproduzierbarkeit und langfristiger Änderung der Kennlinie. Für diese Untersuchungen wurde die Prüfstrecke des Geodätischen Instituts benutzt, deren Laser-Interferometer eine Genauigkeit der Messung von besser als ± 1 µm/m zuläßt. Die Messung und Auswertung wurde mit einem Programmpaket in zwei Schritten durchgeführt. Der erste Schritt, die Messung, wird interaktiv gesteuert; die Meßwerte werden auf eine Diskette angelegt. Im zweiten Schritt, der Ausgleichung, werden Funktionen in Form von Polynomen oder trigonometrischen Reihen als Kennlinien bestimmt. Diese liefern im Vergleich untereinander Aussagen über das günstigste funktionale Modell, wonach durch statistische Tests die den Meßwerten am besten angepaßte Funktion bestimmt wird.

Für die Temperaturuntersuchungen stand die Klimakammer des Geodätischen Instituts zur Verfügung, in der Messungen zwischen -20°C und + 40°C durchgeführt wurden. Zur Festlegung bestimmter Meßlängen wurde für jeden Wegaufnehmer eine Befestigungseinrichtung eingesetzt, die für die meisten Fälle drei konstante Meßpunkte vorsah. Mit der Erfassung der Meßwerte wurde gleichzeitig die Temperatur mittels Pt-100-Temperaturaufnehmern beobachtet, so daß die Längenänderung zwischen den Befestigungspunkten (Metallausdehnung) sofort rechnerisch berücksichtigt werden konnte.

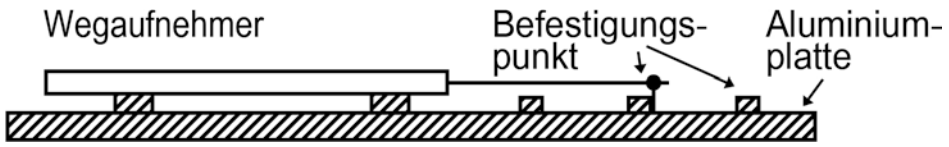


Abb. 3 Meßaufbau in der Klimakammer

3.3 Messungen auf der Prüfstrecke

Ziel der Messungen auf der Prüfstrecke ist es zunächst einmal, die Kennlinie eines jeden Wegaufnehmers zu bestimmen. Diese ist grundsätzlich im Gesamtzusammenhang Spannungsversorgung - Wegaufnehmer - Übertragungsweg - Analog-/Digitalwandler zu sehen und ist bei jeder Änderung dieses Zusammenhangs neu zu bestimmen. Diese Messungen wurden einer eingehenden Analyse unterworfen,

aus der hier einige Schlüsse gezogen werden sollen:

Nr.	Lineare Ausgleichung $\lvert V_{max}\rvert$		s_0		Linearität	Grad	Polynome höheren Grades $\lvert V_{max}\rvert$		s_0	
	[mm]	[%]	[mm]	[%]	[%]		[mm]	[%]	[mm]	[%]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
1	0.06	0.06	0.04	0.04	0.2	–				
2A	0.76	0.30	0.41	0.16	0.2	2	0.40	0.16	0.17	0.07
B	0.20	0.08	0.11	0.04	0.2	–				
3A	2.72	0.27	1.14	0.11	0.2	–				
B	1.13	0.11	0.69	0.07	0.2	2	0.54	0.05	0.26	0.03
4A	0.74	0.08	0.26	0.03	0.2	2	0.27	0.03	0.12	0.01
B	0.37	0.04	0.12	0.01	0.2	2	0.25	0.03	0.10	0.01
5	0.05	0.03	0.02	0.01	0.1	–				
6	0.36	0.04	0.14	0.02	0.1	3	0.27	0.03	0.11	0.01
7	0.20	0.20	0.11	0.11	0.7-1,0	3	0.10	0.10	0.04	0.04

Tab. 2 Ausgleichung der Kalibrierung

Tabelle 2 stellt einige Ergebnisse der Ausgleichung zusammen. Dabei sind die Aufnehmer wie in Tabelle 1 numeriert. A und B bezeichnen die verschiedenen Spuren. Als erstes sind die Ergebnisse der Geradenausgleichung dargestellt. Dabei ist s_0 der mittlere Fehler der Einzelmessung absolut in mm und im Verhältnis zum Meßbereich angegeben. Da die Linearität nach Firmenangabe in Spalte 6 ein Maß für die maximale Abweichung ist, ist in den Spalten 2 und 3 die maximale Verbesserung angegeben (wieder absolut und im Verhältnis zum Meßbereich). Die Spalten 7 bis 11 enthalten die Ergebnisse der Polynomausgleichungen höheren Grades, soweit sie sich als statistisch signifikant erwiesen haben.

Die Tabelle bestätigt, daß die Genauigkeit mit kleiner werdendem Meßbereich zunimmt. Die größten Abweichungen von der Geraden liegen in allen Fällen (außer Nr. 2A und 3A) im Bereich der angegebenen Linearität, teilweise erheblich besser.

In einige Fällen, vor allem bei langen Wegaufnehmern, beschreiben quadratische und kubische Polynome das tatsächliche Verhalten des Instruments besser, so daß eine Berücksichtigung solcher Polynome die Meßgenauigkeit steigert. Dies gilt insbesondere für die Aufnehmer der Firma Penny & Giles. Bei den Potentiometern der Firma Novotechnik führt die Einführung von Gliedern höherer Ordnung zu keiner Genauigkeitssteigerung, da die Kennlinie von Haus aus ausreichend linear ist. Das Vergleichsinstrument, der induktive Wegauf-

nehmer, zeigt die für diesen Typ ausgeprägte Kurve dritten Grades (***KAHMEN*** 1978, S. 218). Entsprechend ist auch hier die Genauigkeit zu steigern, indem Glieder höherer Ordnung eingeführt werden oder der Meßbereich eingeschränkt wird.

Nr.	$\lvert b_2 - b_1 \rvert$ [1/100 mm/digit]	$\lvert (b_2 - b_1)/b_1 \rvert$ [%]
1	0.00047	0.10
2A	0.00020	0.02
B	0	0
3A	0.00150	0.04
B	0	0
4A	0.00184	0.06
B	0.00020	0.01

Tab. 3 Kalibrierungen vor und nach den Temperaturuntersuchungen

Um die Reproduzierbarkeit der Kennlinien zu untersuchen, wurden in Tabelle 3 die linearen Parameter (b_1 und b_2) aus den Kalibrierungen vor und nach den Temperaturuntersuchungen miteinander verglichen. Für die Wegaufnehmer Nr. 1 bis 4 (siehe Tabelle 1) ergeben sich keine deutlichen Änderungen. Die Daten der Aufnehmer 5 und 6 lassen keine Schlüsse zu, da im Datensatz der jeweils ersten Messung größere Fehler durch zu starke Reibungskräfte (Hysterese) auftraten. Erst nach der ersten Messung konnten diese Probleme behoben werden, so daß keine echten Vergleichswerte vorliegen.

3.4 Messungen in der Klimakammer

Zur Untersuchung des Temperatureinflusses auf die Wegaufnehmer wurden Messungen in der Klimakammer, wie schon vorher angegeben, im Bereich von -20°C bis 40°C durchgeführt. Dazu wurde jeder Aufnehmer an zwei bis vier fest definierten Stellen in 10°C-Schritten ausgemessen und an jeder dieser Stellen eine ausgleichende Gerade in Abhängigkeit von der Temperatur bestimmt. Die Ergebnisse dieser Ausgleichungen sind in Tabelle 4 dargestellt.

Unter Nr. stehen die in Tabelle 1 genannten Gerätenummern, s_0 sind die mittleren Fehler aus Tabelle 2 zum Vergleich in 1/100 mm und l die Entfernung zwischen dem Endpunkt der Auszugsstange und dem vorderen Befestigungspunkt des Wegaufnehmers in mm. Δl ist die ausgeglichene Differenz des Meßsignals bei -20°C und bei 40°C, umgerechnet in Wegeinheiten (1/100 mm). Ein

Nr.	l	Δl	l	Δl	l	Δl	l	Δl	s_0
1	45		110						
		- 20.0		- 22.2					4
2	28		158		222				
A		+ 1.9		- 7.6		- 1.4			17
B		- 17.7		- 23.0		- 18.2			11
3	51		303		555		807		
A		+ 22.6		- 6.9		- 17.4		- 34.7	114
B		+ 3.5		- 13.8		- 34.6		- 55.4	26
4	67		319		823				
A		0		- 12.8		- 51.0			12
B		- 22.3		- 44.7		- 81.3			10
5	92		157		222				
		+ 73.4		- 28.7		- 11.7			2
6	116		368		872				
		+ 405.7		+ 264.6		- 35.3			11
7	27		92						
		+ 327.1		- 145.1					4

Tab. 4 Ausgleichung der Temperaturmessungen

positiver Wert besagt, daß am Meßwert bei höherer Temperatur eine größere Verbesserung anzubringen ist.

Es zeigt sich, daß die Meßwerte bei größeren Temperaturunterschieden vom Sollwert stark abweichen. Meist sind bei hohen Temperaturen negative Korrekturen anzubringen, die, außer bei Nr. 2 und 5, umso kleiner werden, je weiter der Meßfühler ausgezogen ist. Auffallend ist, daß die Draht- und Hybridpotentiometer (Nr. 1 bis 4) die geringsten Korrekturen haben, während die beiden Leitplastik-Film-Potentiometer (Nr. 5 und 6) wesentlich größere Verbesserungen zeigen. Der induktive Wegaufnehmer liegt mit seinem Wert nochmal um den Faktor 7 höher. Das stimmt auch mit den Angaben bei ***GASS*** (1978) überein.

Natürlich ist die Temperaturdifferenz von 60°C sehr groß, will man aber mit den Temperaturfehlern im Bereich des mittleren Fehlers s_0 bleiben, so muß sichergestellt sein, daß die Temperaturen bei den Aufnehmern Nr. 1 bis 4 um höchstens 10°C bis 20°C und bei den Nr. 5 und 7 um weniger als 1°C von der Bezugstemperatur der Kalibrierung abweichen.

Vergleicht man jedoch die hier ermittelten Werte mit den Angaben der Firmen (± 125 ppm/°C für die Aufnehmer Nr. 1 bis 4 und ± 350 ppm/°C für die Nr. 5

und 6), so liegen alle gemessenen Werte innerhalb dieser Schranken.

Ein weiterer Aspekt, der in diesem Zusammenhang von Interesse ist, ist das zeitliche Verhalten der Wegaufnehmer bei Temperaturänderungen. Bei stärkeren Temperaturschwankungen reagieren die verschiedenen Aufnehmer unterschiedlich. Nur bei wenigen Geräten folgen die Meßwerte unmittelbar den Temperaturänderungen (Nr. 1, 2A und 6). Die Signale der Aufnehmer 2B, 4A und B zeigen zunächst für eine gewisse Zeit eine entgegengesetzte Tendenz, bevor sie sich langsam dem endgültigen Wert annähern. Der Verlauf der übrigen Aufnehmersignale unterscheidet sich wiederum davon und ist eher einem langsamen Einschwingvorgang zu vergleichen. Aus diesen Beobachtungen ist der Schluß zu ziehen, daß während starker und plötzlicher Temperaturschwankungen sehr präzise Messungen nicht möglich sind.

3.5 Ergebnisse der Laboruntersuchung

Die Untersuchungen zeigen deutlich, daß mit den Potentiometerwegaufnehmern durchaus präzise Messungen möglich sind, die einem Vergleich mit induktiven Wegaufnehmern gut standhalten können. Dabei sind, wie bei allen Aufnehmern, bestimmte Anforderungen zu berücksichtigen. Eine Voraussetzung ist zunächst eine hohe Konstanz der Eingangsspannung, da sich diese in vollem Umfang auf das Meßergebnis auswirkt. Die relative Genauigkeit des Aufnehmers kann nie höher sein als die relative Genauigkeit der Spannung. Der am schwierigsten zu beherrschende Einfluß ist der der Temperatur. Dazu muß für genaue Messungen die Kennlinie unbedingt im Bereich der Betriebstemperatur bestimmt werden; unter Umständen sind mehrere Kennlinien notwendig. Außerdem ist zu berücksichtigen, welche Teile der Meßkette (Aufnehmer, Kabel, A/D-Wandler) eventuellen Temperaturschwankungen ausgesetzt sind. Weiter können hochpräzise Messungen die Berücksichtigung einer nichtlinearen Kennlinie notwendig machen, ein Umstand, der bei programmierbaren Meßwerterfassungsanlagen keine Probleme aufwirft. Schließlich hat es sich bei der Untersuchung noch als bedeutsam herausgestellt, daß wegen der teilweise hohen Reibungskräfte auf eine gute Befestigung der Aufnehmer geachtet werden muß, um Hystereseeffekte zu vermeiden.

Dabei ist es sinnvoll, zu beachten, daß Leitplastik-Hybrid-Potentiometer die besseren Temperatureigenschaften, dafür aber oft erkennbare Nichtlinearitäten und geringere Lebensdauer haben.

4. POTENTIOMETERSYSTEM ZUR MESSUNG VON BRÜCKENTRÄGERVERSCHIEBUNGEN

Als ein Anwendungsbeispiel soll im folgenden der Aufbau und die Funktion eines kleinen Meßsystems dargestellt werden, welches zur Ausmessung von Brückenträgerverschiebungen konzipiert wurde.

4.1 Aufgabenstellung

Für den Bau eines elektromagnetischen Schnellbahnsystems ist es notwendig, die 25 m oder 37 m langen Beton-Fahrwegträger so genau auf den Stützen zu positionieren, daß die eigentliche Fahrwegausrüstung auf ihnen mit höchster Präzision justiert werden kann. Dazu wird der Träger mit einer Genauigkeit von ± 30 mm grob positioniert. Diese Istlage wird mit geodätischen Methoden eingemessen, um anzugeben, um welche Beträge der Träger in die Sollposition verschoben werden muß. Vom Bedienungsstand hydraulischer Pressen aus sollte danach die Möglichkeit bestehen, die Bewegung in Fahrbahnrichtung, quer zur Trasse und in der Höhe zu verfolgen.

Es war nun die Aufgabe gestellt, ein Meßsystem zu entwickeln, mit dem die Verschiebung in Fahrbahnrichtung (x-Richtung) gemessen werden konnte. Gefordert war eine Genauigkeit von ± 1 mm über einen Meßbereich von mindestens 100 mm. Darüberhinaus sollte in 30 m Entfernung vom Meßort (Spalt zwischen zwei Brückenträgern) eine Anzeige installiert werden, an der der Verschiebungswert kontinuierlich und direkt in einer Wegeinheit abgelesen werden kann.

4.2 Aufbau des Systems

Wegen der Kosten, der Zuverlässigkeit und der günstigsten Lieferzeit wurde zur Lösung dieser Aufgabe auf ein Präzisionspotentiometer als Wegaufnehmer zurückgegriffen. Die auch in Erwägung gezogenen Meßsysteme mit inkremental geteilten Glasmaßstäben wurden nicht eingesetzt, da ihr Anschluß auch mit Verlängerungskabel eine Fernübertragung der Meßsignale nur bis 17 m zuließ.

Der gewählt Wegaufnehmer ist ein Leitplastik-Hybrid-Potentiometer, der etwa mit dem Aufnehmer Nr. 2 in Tabelle 1 verglichen werden kann, allerdings mit nur einer Spur. Um für den Betrieb genügend Sicherheit zu geben, wurde ein Meßbereich von 125 mm gewählt, so daß mit einer Genauigkeit von ± 0,3 mm ge-

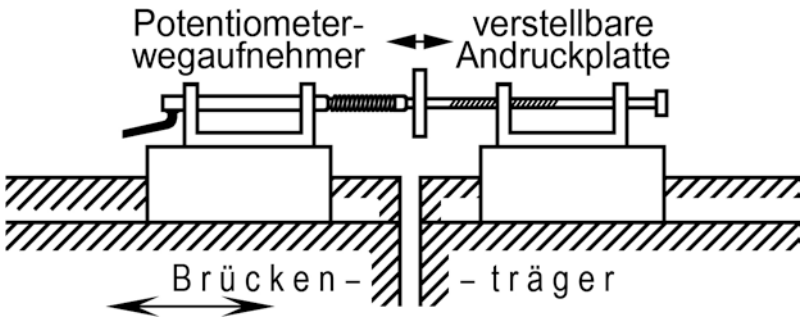


Abb. 4 Das Meßsystem auf dem Brückenträger

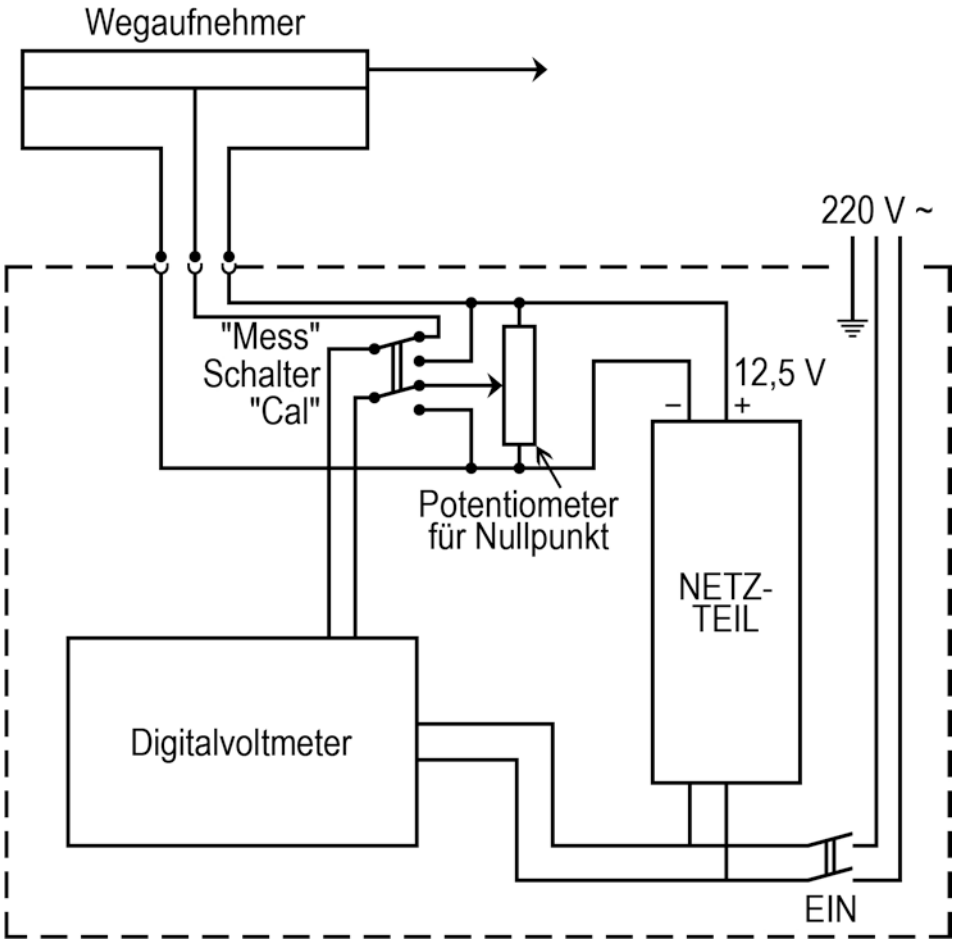


Abb. 5 Innerer Aufbau des Wegaufnehmersystems

rechnet werden konnte. Dieser Aufnehmer wurde mit einer Feder versehen und auf einer Platte montiert, so daß er auf dem Fahrweg befestigt werden konnte. Auf der gegenüberliegenden Seite wurde eine verstellbare Platte als Anschlag angebracht (Abb. 4).

Über ein Kabel ist der Aufnehmer mit einer Anzeigeeinheit verbunden, die für den Bedienungsstand der Pressen vorgesehen ist. Die Schaltung dieser Anzeigeeinheit ist in Abbildung 5 dargestellt.

Die Speisung des Aufnehmers erfolgt über ein stabilisiertes Netzteil, welches eine konstante Spannung zwischen 7 und 15 Volt liefert. Diese Spannung wurde auf 12,5 V eingestellt, kann aber jederzeit nachgeprüft und über ein Potentiometer justiert werden. Ein weiteres Potentiometer dient der Festlegung des Nullpunktes in der Mitte des Meßbereichs. Die Abtastung erfolgt über in Digitalvoltmeter, dessen Dezimalpunkt so verschoben wurde, daß durch den Maßstab 0,1 V $\triangleq$ 1 mm genau der Verschiebeweg in mm angezeigt wird.

4.3 Testmessungen

Das System wurde vor seinem Einsatz umfangreichen Testmessungen unterworfen. Diese unterscheiden sich von den in Kapitel 3 genannten vor allem dadurch, daß hier das gesamte System mit Spannungsversorgung und Kabel untersucht und kalibriert wurde. Im folgenden sollen einige Ergebnisse zusammengefaßt werden.

Der Maßstab des Systems wird wesentlich durch die Speisespannung bestimmt. Diese kann so genau eingestellt werden, daß keine signifikanten Verfälschungen des Maßstabs festzustellen sind.

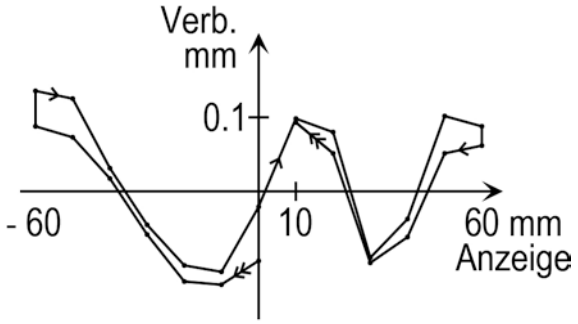


Abb. 6 Verbesserungskurve

Abbildung 6 zeigt eine typische Verbesserungskurve, die auch über andere Meßreihen reproduzierbar ist. Auffallend ist dabei die Hysterese, die sich bei Hin- und Rückgang aus der relativ hohen Reibung und im Verhältnis dazu geringem Andruck der Feder ergibt. In einer linearen Ausgleichung ergibt sich ein mittlerer Fehler einer Anzeige von ± 0,1 mm.

Bei den Temperaturuntersuchungen wurde das Anzeigegerät teils in der Klimakammer und teils außerhalb untergebracht. Dabei ist auffallend, daß die Meßergebnisse sehr stark von der Temperatur des Anzeigegeräts abhängen. Für einen Weg von 60 mm ergibt sich eine Temperaturverbesserung von ± 0,01 mm/°C.

Insgesamt erwies sich das Meßsystem als sehr zuverlässig und den Anforderungen entsprechend.

5. ZUSAMMENFASSUNG

Aus diesen Erkenntnissen heraus läßt sich feststellen, daß Präzisionspotentiometer für ingenieurgeodätische Anwendungen einige Vorteile gegenüber anderen Wegaufnehmern haben. Als wichtig wäre zu nennen, daß sie flexibel im Meßbereich sind, vor allem in den Bereichen, die für Deformationsmessungen von Bedeutung sind. Weiter ist die hohe Reproduzierbarkeit der Meßwerte zu erwähnen. Darüberhinaus sind sie einfach in der Handhabung und relativ unempfindlich gegenüber äußeren Einflüssen. Ein wichtiger Nachteil ist für manche Anwendungen sicher der, daß die Messung nicht berührungsfrei ist, da zur Verschiebung des Gebers ein gewisser Kraftaufwand nötig ist. Um gute Ergebnisse zu erzielen, ist es notwendig, systematische Fehlereinflüsse der Speisespannung, Temperatur und der Nichtlinearität der Kennlinie zu berücksichtigen.

Bei normalem, einigermaßen sorgfältigem Einsatz lassen sich ohne weiteres Genauigkeiten von 0,2 % des Meßbereichs erzielen. Werden jedoch die genannten Einflüsse durch Abschirmung reduziert oder durch eine entsprechende Auswertung berücksichtigt und ständig kontrolliert, läßt sich die Genauigkeit auf 0,05 % oder besser steigern.

LITERATUR

BOCK, D.: ***Auswahlkriterien für Wegmeßsysteme.*** Feingerätetechnik 31, S. 398-404, 443-447, 1982

BRAUER, R.: ***Untersuchung von potentiometrischen Wegaufnehmern.*** Diplomarbeit am Lehrstuhl für Vermessungskunde an der Hochschule der Bundeswehr München 1982 (unveröffentlicht)

GASS, E.: ***Präzisionspotentiometer – zuverlässige Winkel- und Wegaufnehmer für industrielle Anwendungen.*** Messen + Prüfen / Automatik 14, S. 839-842, 1978

HERCEG, E.E.: ***Handbook of Measurement and Control.*** Schaevitz, Pennsauken N.J. (USA) 1976

JÜTTEMANN, H.: ***Grundlagen des elektrischen Messens nichtelektrischer Größen.*** VDI-Verlag, Düsseldorf 1974

KAHMEN, H.: ***Elektronische Meßverfahren in der Geodäsie.*** Wichmann-Verlag, Karlsruhe 1978

MÖLZER, P., und ZUREK, R.: ***Elektromagnetische Schnellbahnsysteme (EMS).*** Allgemeine Vermessungsnachrichten 87, S. 154-165, 1980

ROESEMANN, R.: ***Übersicht elektronischer Verfahren zur Längen- und Winkelmessung.*** Messen + Prüfen / Automatik 12, S. 33-44, 1976

AUFBAU UND FUNKTIONSWEISE VON KOMPENSATOREN IN OPTISCHEN SEKUNDENTHEODOLITEN UND ELEKTRONISCHEN THEODOLITEN

von Axel GEIGER

SUMMARY

The principle of vertical circle compensators and their classification are treated in this paper. Especially the principle and the function of the compensators of the mostly used optical one-second-theodolites as well as the electronical theodolites, including systematical drawings are described in detail. The most important technical data are put together in a table.

ZUSAMMENFASSUNG

In dem vorliegenden Beitrag wird die Wirkungsweise von Kompensatoren und ihre Einteilung behandelt. Insbesondere wird der Aufbau und die Funktionsweise sowohl der gebräuchlichsten optischen Sekundentheodolite als auch der elektronischen Theodolite, jeweils mit einer Systemskizze versehen, beschrieben. Die wichtigsten technischen Daten sind in einer Tabelle zusammengestellt.

1. EINLEITUNG

Als Neigungskompensatoren werden im geodätischen Instrumentenbau jene Einrichtungen bezeichnet, die bei automatischen Nivellieren die Ziellinienhorizontierung und bei Theodoliten mit automatischem Höhenindex die Indexstabilisierung bewirken.

Die Verwendung des Begriffs "Neigungskompensator" ist in der Literatur nicht einheitlich, da verschiedene Auffassungen darüber bestehen, welche Bauteile zum Neigungskompensator gehören. Teilweise verstehen die Autoren unter diesem Begriff nur das Kompensatorpendel, teilweise zählen aber auch fernrohrfeste optische Elemente dazu. Da letztere keinen Einfluß auf die eigentliche Kompensation haben, zählen hier nur das Pendel und das von diesem gesteuerte Element zum Neigungskompensator.

Für den Begriff Neigungskompensator gibt es in der Literatur mehrere Synonyme: automatischer Höhenindexkompensator, automatischer Kompensator, Kompensationsvorrichtung etc. In dieser Abhandlung wird einheitlich der Begriff "Kompensator" verwendet.

Der vorliegende Bericht soll einen Überblick über den Aufbau und die Funktionsweise jener Kompensatoren vermitteln, die in den gebräuchlichsten optischen Sekundentheodoliten und elektronischen Theodoliten eingebaut sind.

2. WIRKUNGSWEISE EINES KOMPENSATORS

Unter dem Einfluß der Schwerkraft bezieht der Kompensator die Ablesung am Vertikalkreis auf die Lotrichtung am Beobachtungsort, indem er die in der Zielebene liegende Komponente der restlichen Stehachsneigung automatisch berücksichtigt. Die Wirkungsweise eines Neigungskompensators wird an Abbildung 1 erläutert.

Bei streng lotrechter Stehachse VV und genauer Horizontierung der Zielachse ZZ des Fernrohrs steht der Index I der Ablesestelle A am Vertikalkreis gegenüber. Die gemessene Zenitdistanz beträgt 100 gon. Wird ein Stehachsenfehler α erzeugt, weicht die Stehachse V'V' jedoch vom Lot ab. Der Index befindet sich dann in I'. Da der Teilkreis diese Bewegung mitmacht, gelangt A nach A', womit die Ablesung noch nicht fehlerhaft wäre.

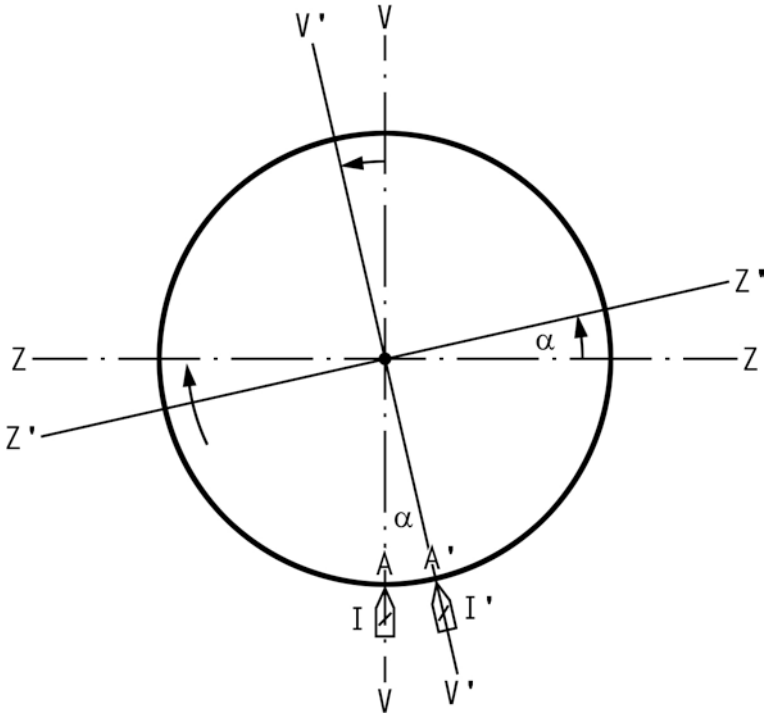


Abb. 1 Wirkungsweise eines Neigungskompensators

Erst die Rückdrehung der Zielachse von Z'Z' nach ZZ beim Wiedereinstellen des Zieles bewirkt eine falsche Messung, weil dabei die dem Index I' gegenüberliegende Ablesestelle A' wieder nach A gelangt. Der am Index abzulesende Wert weicht also um α von Sollablesung 100 gon ab.

Die Aufgabe eines Kompensators besteht nun darin, die Ablesestelle A so auf den Index der Strichplatte des Ablesemikroskops zu projizieren, daß stets die vom Neigungsfehler α befreite Ablesung A erhalten wird. Dazu sind verschiedene konstruktive Möglichkeiten entwickelt worden. Als wichtigste seien genannt:

a) Einbau eines Kompensators in den Strahlengang des Ablesemikroskops

b) Automatische Verschiebung des Indexstriches

c) Automatische Verschiebung des Mikroskopobjektives

d) Erfassen der Lotabweichung durch Sensoren und rechnerische Verbesserung (elektronische Theodolite).

Bei der unter a) genannten Konstruktion wird in das Mikroskop ein Kompensator eingeschaltet (Abbildung 2), der den Abbildungsstrahlengang in der Ebe-

ne durch D um einen Winkel β, der von der Stehachsneigung α abhängig ist, abknickt. Der Abbildung ist die Justierbedingung

$$\frac{\beta}{\alpha} = \frac{a}{b} \cdot \frac{r}{s} = n = \text{const.}$$

zu entnehmen.

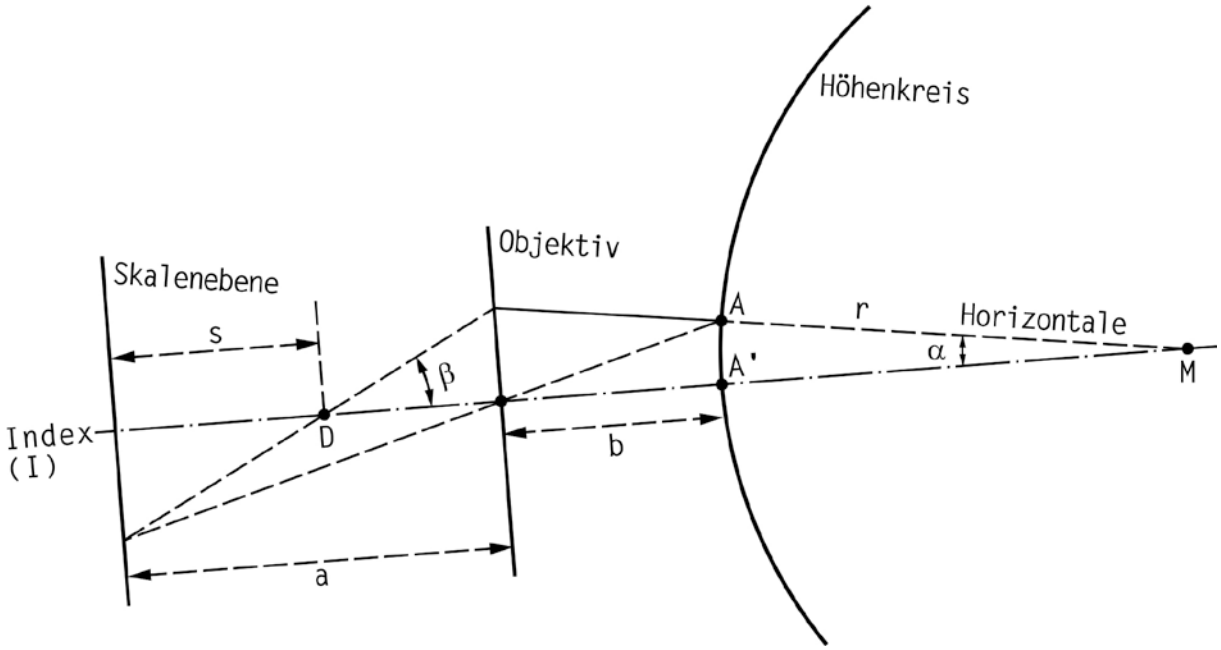


Abb. 2 Schematische Strahlengang im Ablesemikroskop für den Höhenkreis

3. EINTEILUNG VON KOMPENSATOREN

Zur Korrektur der Stehachsneigung sind neben den verschiedenen Anordnungen der Kompensatoren im Strahlengang auch für die Kompensatoren selbst verschiedene konstruktive Lösungen entwickelt worden. Allen Lösungswegen ist gemeinsam, daß ein Teil des Kompensators fest mit dem Instrument verbunden ist, während ein anderes Teil beweglich ist und der Schwerkraft unterliegt.

3.1 Kompensatoren optischer Theodolite

Die allgemein in Theodoliten eingebauten Kompensatoren lassen sich in drei Gruppen einteilen:

a) das freischwingende Pendel

b) Pendel mit erzwungenen Schwingungen -
   auch astasiertes Pendel genannt

c) Flüssigkeitskompensatoren.

Einen guten Überblick gibt die von ***GRÄBER*** (1980) aufgestellte Einteilung der Kompensatoren mit Angabe der Theodolite, die diese Kompensatoren eingebaut haben (Abbildung 3). Besonders gekennzeichnet sind die Kompensatoren, die im Rahmen dieser Arbeit beschrieben werden.

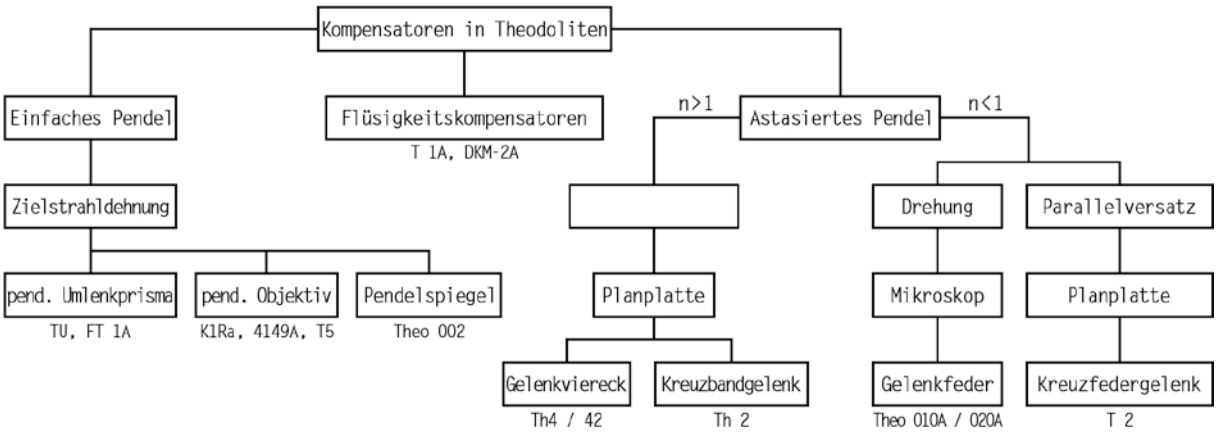


Abb. 3 Übersicht der wichtigsten Kompensatorprinzipien in Theodoliten

Die Theorie zu den Pendelkompensatoren wird ausführlich in ***HOFMANN*** (1958) behandelt.

3.2 Kompensatoren elektronischer Theodolite

Bezeichnend für elektronische Theodolite ist die elektronische Kreisablesung. Bei der Konstruktion dieser Instrumente gilt die Bestrebung, den Stehachsfehler dahingehend zu eliminieren, daß an den Ablesungen keinerlei Korrekturen angebracht werden müssen. Dazu werden von der Schwerkraft gesteuerte Sensoren verwendet. Diese Bauteile sind mit den Kompensatoren in Nivellieren und Theodoliten vergleichbar. Sie können in Flüssigkeitssensoren und Pendelsensoren eingeteilt werden.

4. ANFORDERUNGEN AN KOMPENSATOREN

Prinzipiell können alle Kompensatoren als Pendel betrachtet werden. Das Pendel, das ein optisches Element steuert, nimmt gegenüber einer Basis eine bestimmte Stellung ein, die der jeweiligen Basisneigung entspricht. Das Pendel kann durch ein einfaches Pendel, ein astasiertes Pendel, eine Flüssigkeit oder eine Gasblase dargestellt werden. Unabhängig von der Konstruktion muß ein Kompensator verschiedene Bedingungen erfüllen:

- die Kompensatorfehler müssen kleiner sein als die anderen auf die Messung wirkenden Einflüsse (Zielfehler, Koinzidenzfehler, Kreisteilungsfehler)

- der Vergrößerungsfaktor des Pendels

  $$n = \frac{\alpha}{\beta} = \frac{\text{Neigung des Pendels relativ zur Basis}}{\text{Neigung der Basis relativ zur Horizontalen}}$$

  muß im Arbeitsbereich konstant sein

- der Nullpunkt muß zumindest für die Dauer der Messung konstant sein.

5. KOMPENSATORFEHLER

Die Arbeitsweise der Kompensatoren ist nicht fehlerfrei. ***HOFMANN*** (1958) und ***WENZEL*** (1969) unterscheiden entsprechend der wichtigsten Fehlerursachen Kompensationsfehler, Umkehrspanne, Einspielfehler mit Temperatureinflüssen. Die Instrumentenhersteller geben nur für den "mittleren Einspielfehler" einen Wert an, der bei durchschnittlich ± 0,1 mgon liegt. Eine ausführliche Beschreibung von Kompensationsfehlern in optischen Sekundentheodoliten und deren Bestimmung befindet sich in ***GEIGER*** (1982). Untersuchungen zur Genauigkeit der Kompensatoren in elektronischen Theodoliten gibt es bisher nur wenige.

6. AUFBAU UND FUNKTIONSWEISE VON KOMPENSATOREN IN OPTISCHEN SEKUNDENTHEODOLITEN

6.1 Jenoptik Theo 010A

Beim Theo 010A ist das gesamte Mikroskop mit Ausnahme der Strichplatte und des Höhenindex an einem federgehängten Pendel befestigt. Der Vergrößerungsfaktor n ist dabei < 1.

In Abb. 4 ist die Stehachse um den Winkel α gegen das Lot LL geneigt. Die Pendelaufhängung ist aus LL nach VV ausgewandert. Da diametrale Kreisstellen zusammen abgebildet werden müssen, wird das Objektivsystem an dem Hebelarm l_o durch den Einfluß der Schwerkraft um den Winkel **ω** gedreht. Die verbesserte Abbildung erfolgt in der Weise, daß die Teilkreisstelle A_1 auf die Teilkreisstelle A_2 abgebildet wird.

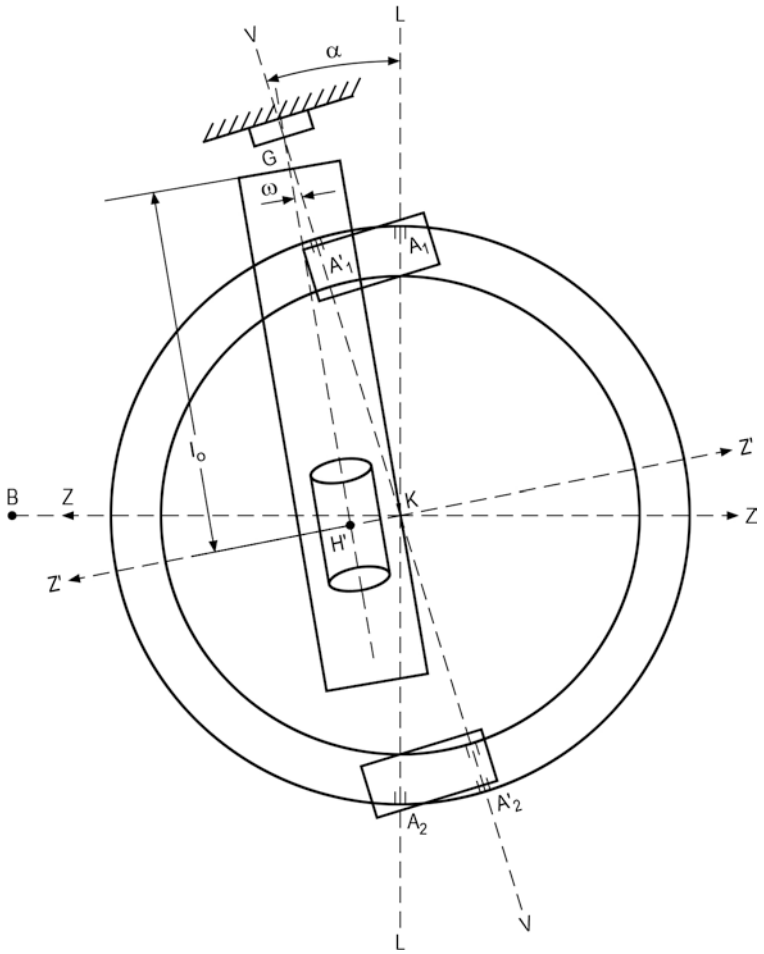


Abb. 4 Kompensator des Jenoptik Theo 010A

6.2 Kern DKM2-A

Beim Kompensator des DKM2-A (Abbildung 5) wird die Totalreflexion, die ein Lichtstrahl beim Austritt aus einem optisch dichteren Medium in ein optisch dünneres Medium erfährt, ausgenutzt. Ein Flüssigkeitsbehälter (1) ist instrumentenfest montiert. An seiner Bodenfläche befindet sich ein symmetrisches, trapezförmiges Prisma (2). Der Brechungsindex der Flüssigkeit (Silikonöl) hat nahezu die gleiche Größe wie der von Glas. Bei lotrechter Aufstellung (LL) werden die diametralen Kreisstellen A_1 und A_2 in A_2 abgebildet. Bei einem Stehachsfehler (VV) sind die instrumentenfesten Teile um den Winkel ν geneigt. Der von A_1 kommende Lichtstrahl wird an der horizontalen Flüssigkeitsoberfläche um 100 gon + 2 ν nach A'_1 reflektiert. Das Mittel aus A'_1 und A'_2 liefert die vom Stehachsfehler befreite Ablesung. Der Kompensator spiegelt also die diametralen Kreisstellen zusammen und eliminiert somit die Stehachsschiefe.

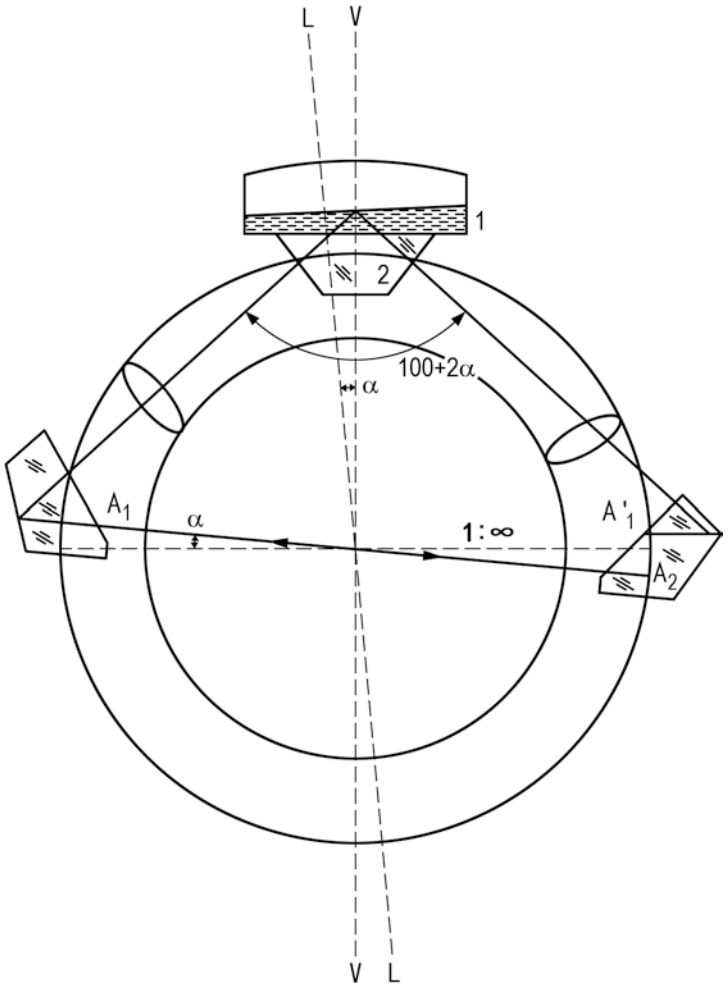


Abb. 5 Kompensator des Kern DKM2-A

6.3 Wild T2

Bei dem Kompensator (1) des Wild T2 (Abbildung 6) handelt es sich um ein mechanisches Kreuzfedergelenk, das sich durch je zwei sich kreuzende Bänder (2) auszeichnet. Der Winkelvergrößerungsfaktor n ist dabei < 1. Das hängende Pendel ist sowohl magnetisch (3) als auch luftgedämpft (4). Daher reagiert der Kompensator auf Störungen relativ unempfindlich. Das Pendel steuert die beiden Planplatten (5) und (6). Nach der Kompensation der Stehachsneigung fallen die Ablesestrahlen auf die Ableseprismen (7) und (8).

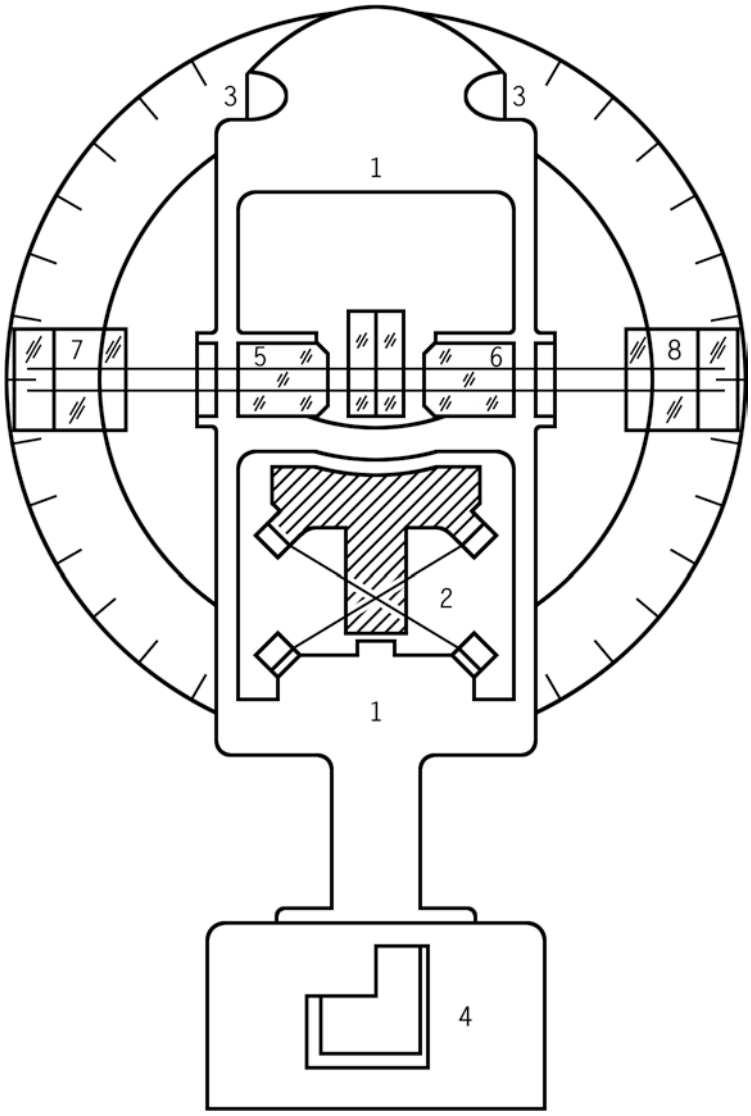


Abb. 6 Kompensator des Wild T2

6.4 Zeiss Th2

Der Kompensator des Zeiss Th2 (Abb. 7) ist ein X-förmiges Gelenkviereck, das wie ein mechanisches Hebelgetriebe wirkt. Der Winkelvergrößerungsfaktor n ist > 1. Der Schwerpunkt der beweglichen Kompensatorteile liegt über dem Drehpunkt. Durch ein längs der Pendelstange verschiebbares Gewicht kann die Schwerpunktlage und somit der Vergrößerungsfaktor reguliert werden. Die gesteuerten optischen Bauelemente sind zwei planparallele Platten, die das Bild der Kreisteilung soweit versetzen, daß eine Stehachsneigung (VV) kompensiert wird. Der Kompensator wirkt auf beide diametrale Kreisstellen. Die fiktive Pendellänge des Kompensators, berechnet aus der Schwingungsdauer, beträgt 1,7 m. Dadurch ist der Kompensator gegenüber von außen aufgezwungenen Schwingungen weitgehend unempfindlich. Der Kompensator ist luftgedämpft. Über einen Hebelarm wird ein Kolben in einen Zylinder eingeführt. Durch die kreisförmige Bewegung des Kolbens ist seine Eindringtiefe in den Zylinder begrenzt. Dadurch wird der Arbeitsbereich des Kompensators eingeschränkt.

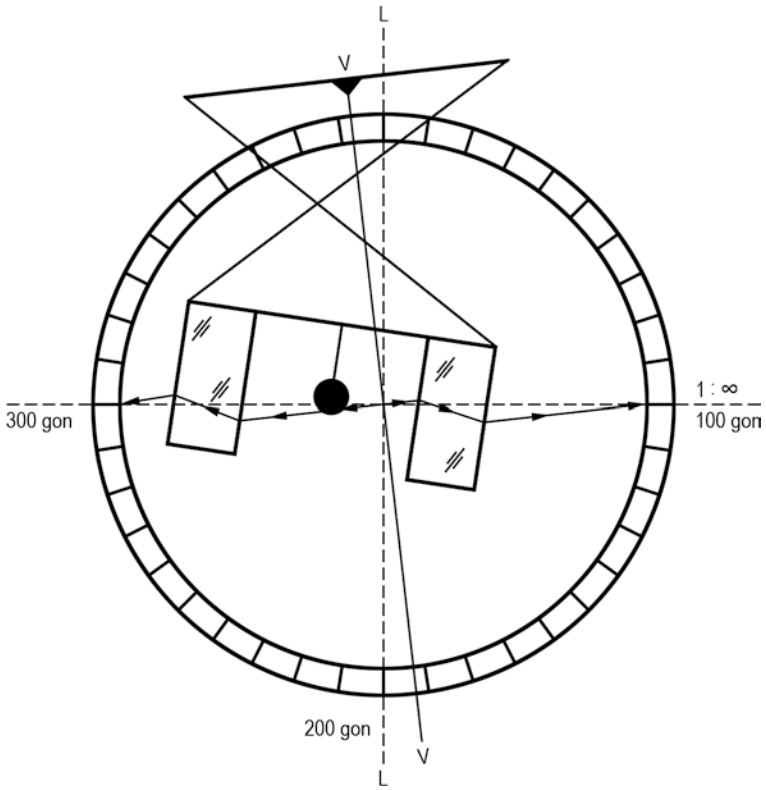


Abb. 7 Kompensator des Zeiss Th2

7. AUFBAU UND FUNKTIONSWEISE VON KOMPENSATOREN IN ELEKTRONISCHEN THEODOLITEN

7.1 AGA Geodimeter 140

Auf eine Beschreibung des im Geodimeter 140 eingebauten Kompensators muß leider verzichtet werden, da vom Hersteller AGA keine exakten Angaben erhältlich waren. Vermutlich ist der Kompensator des Geodimeter 140 ähnlich dem der Semi-Totalstation Geodimeter 120. Hier wird zur Streckenreduktion die Zenitdistanz mit einem mechanischen Pendel als Neigungssensor durch einen Nullabgleich ("Strapdown-Technik") bestimmt. Eine genaue Beschreibung dieses Kompensators befindet sich in ***SCHERER*** (1982).

7.2 Hewlett-Packard 3820 A

Hewlett-Packard verwendet in seiner Totalstation 3820 A einen elektrooptischen Flüssigkeitssensor (Abb. 8). Das System arbeitet mit einem Quecksilberhorizont. Die Stehachsschiefe in zwei zueinander senkrechten Richtungen wird in elektronische Signale umgewandelt. Die Quecksilberoberfläche dient als horizontaler Spiegel für ein Autokollimationsfernrohr mit vertikaler Achse. Durch einen Kondensor (2) wird das Licht einer Leuchtdiode (1) auf einen Sinusschlitz (3) gebracht. Das Licht durchdringt die beiden Objektivglieder, wird am Quecksilber (4) reflektiert und gelangt über die beiden Objektivglieder zu den Sensoren (5). Dort wird die Stehachsschiefe mit zwei Phasenmessern bestimmt, deren Auflösung 0,1 mgon beträgt. Der zweite Sensor liegt senkrecht zu dem in der Zeichnung dargestellten Sensor. Die Oxydation des Quecksilbers wird durch eine Schicht Silikonöl (6) verhindert.

Der Kompensationsbereich ist auf 45 mgon beschränkt. Ein Überschreiten wird durch ein Blinkzeichen angezeigt. Mit den vom Schwerkraftsensor gelieferten Werten werden in einem Rechner Zenitdistanz und Horizontalwinkel korrigiert.

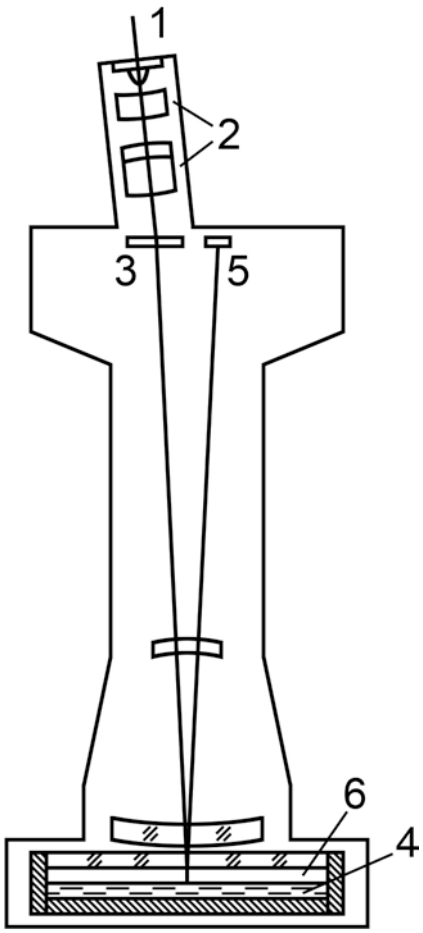


Abb. 8 Flüssigkeitskompensator Hewlett-Packard 3820 A

7.3 Kern E2

Im Gegensatz zum DKM2-A befindet sich der Kompensator des E2 nicht im Ablesestrahlengang der Teilkreise, sondern ist ein getrenntes Bauteil (Abb. 9). Die Kompensation beruht auf der Spiegelung an einer horizontalen Flüssigkeitsoberfläche (1). Der Lichtpunkt einer Beleuchtungsdiode (3) wird nach der Reflektion an der Flüssigkeit auf einer Photodiode (4) abgebildet. Diese Diode funktioniert als zweidimensionaler Positionsdetektor. Trifft der Lichtpunkt auf die Fläche des Detektors, ergeben sich an den vier seitlich angebrachten Elektroden Photoströme. Da diese Photoströme umgekehrt proportional zum Abstand des Lichtpunktes von den Elektroden sind, kann die Lage des Lichtpunktes koordinatenmäßig festgelegt werden. Die eine Koordinate ist ein Maß für die Neigung des Theodolits in Zielachsrichtung, die andere für die Neigung in Kippachsrichtung. Die Kompensationswerte las-

sen sich von der Ablesung trennen und können separat angezeigt werden.

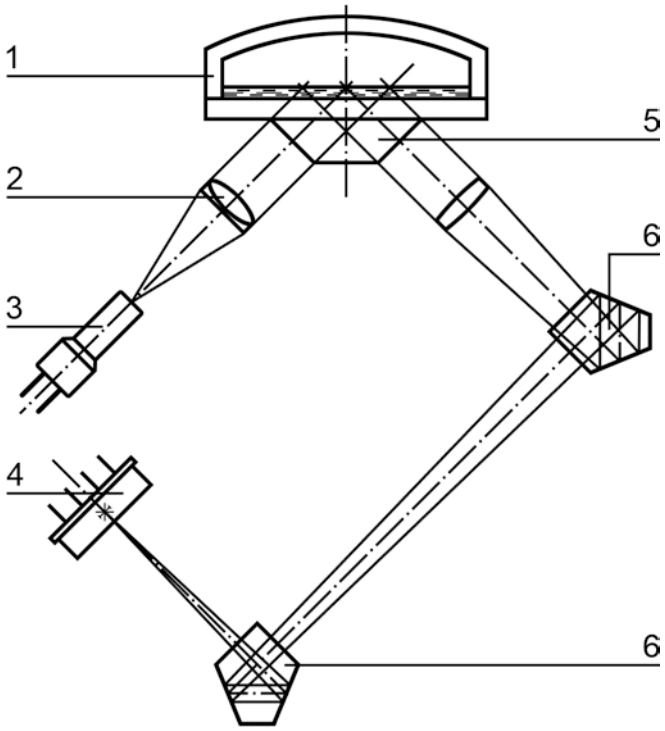


Abb. 9 Kompensator des Kern E2

7.4 Keuffel und Esser Vectron

Beim Vectron befindet sich der elektronische Neigungsmesser (Abb. 10) in der dem Vertikalkreis gegenüberliegenden Fernrohrstütze. Der Kompensator besteht aus einem luftgedämpften Pendel, das in einem geschützten Gehäuse eingebaut ist. Die Richtung der Stehachse wird in Bezug auf die Schwerkraft digital angezeigt. Der Korrekturbereich beträgt maximal ± 65 mgon. Der Neigungssensor kann manuell exakt justiert werden. Die Lage des Pendels in Bezug auf die Stehachse wird mit Hilfe eines Differentialkondensators bestimmt, der die Meßdaten an die Auswerteelektronik weitergibt. Mit den sich ergebenden Neigungen werden die gemessenen Richtungen automatisch verbessert.

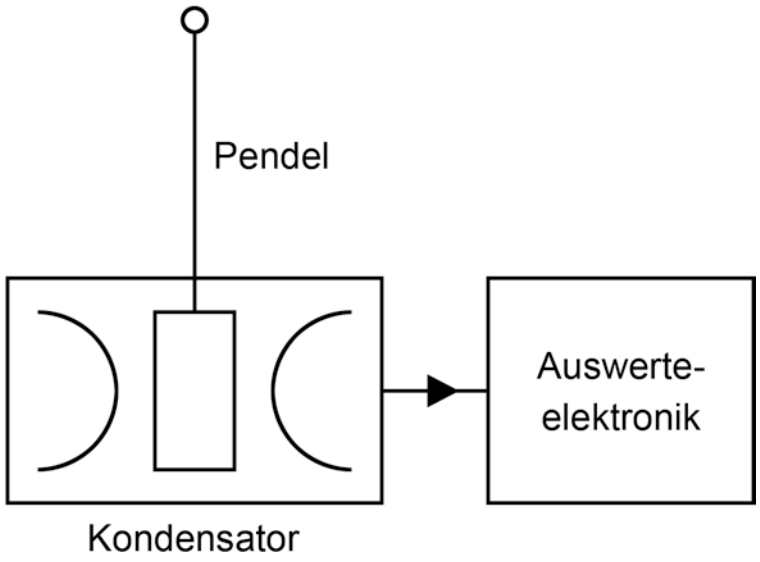


Abb. 10 Pendelsensor des Keuffel und Esser Vectron

7.5 Wild Tachymat TC1

Beim Wild Tachymat TC1 besteht der Kompensator aus einem Kreuzbandpendel (Abb. 11), das dem im Wild T2 (siehe Abbildung 6) eingebauten entspricht. Das Pendel ist an der dem Vertikalkreis gegenüberliegenden Fernrohrstütze angebracht. Es schwingt in einer zum Fernrohr parallelen Vertikalebene und definiert so die Lotlinie. Durch ein Abtastsystem, das aus einem kreisförmigen Ausschnitt mit einer Rasterteilung besteht, wird die Lage des Pendels relativ zum Theodoliten gemessen. Vier Photodetektoren sind so angeordnet, daß sie vier um 100 gon verschobene Sinussignale liefern. Durch Addition und Subtraktion entstehen dadurch vier um 50 gon gegeneinander verschobene Signale mit acht gleichmäßig über die Moiréperiode verteilten Nulldurchgängen. So wird der Einfluß restlicher Horizontalfehler wie bei einem aufgesetzten Index am Vertikalkreis kompensiert. Da die Winkel- und die Sensormessungen erst im Rechner verknüpft werden, kann der Sensor bei Bedarf (z.B. bei Wind oder Erschütterungen) ausgeschaltet werden.

Ein neu auf den Markt kommender elektronischer Theodolit ist der Wild T 2000, der einen Flüssigkeitskompensator besitzt.

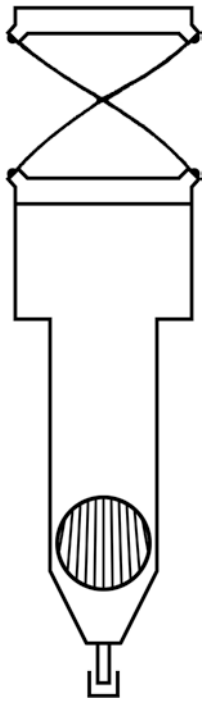


Abb. 11 Pendelsensor des Wild Tachymat TC1

7.6 Zeiss Elta 2

Der im Elta 2 eingebaute Neigungskompensator ist weitgehend identisch mit dem des optischen Theodolits Th 4. Es handelt sich um ein mechanisches V-förmiges Gelenkviereck mit Planplatte (Abb. 12), das die Kompensation der Stehachsschiefe bewirkt. Der Winkelvergrößerungsfaktor ist > 1.
Der Schwerpunkt S der beweglichen Teile des luftgedämpften Kompensators liegt über dem Drehpunkt. Durch ein längs der Pendelstange verschiebbares Gewicht kann die Schwerpunktlage und somit der Vergrößerungsfaktor reguliert werden. Das Code- und Koinzidenzbild des Vertikalkreises wird über die Planplatte des Kompensators zweifach auf die Ablese- und Koinzidenzelektronik übertragen. Bei einer Stehachsschiefe wird diese Abbildung entsprechend der Planplattenkippung auf der Empfängereinheit verschoben.

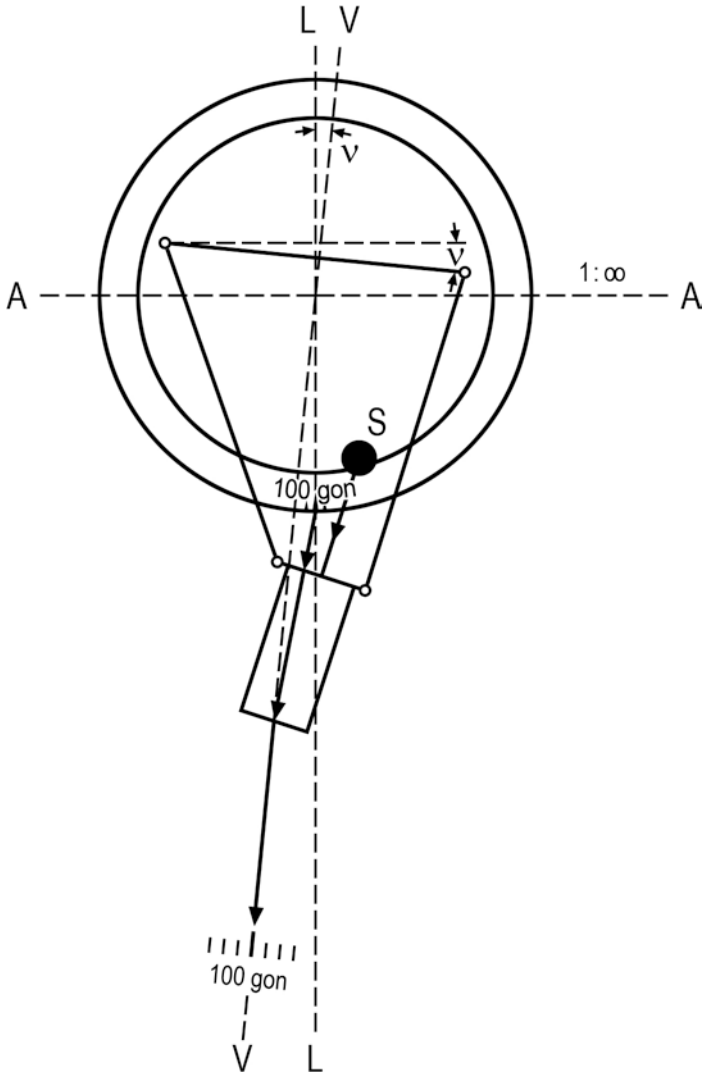


Abb. 12 Kompensator des Zeiss Elta 2

8. ZUSAMMENSTELLUNG TECHNISCHER DATEN

In den folgenden Tabellen 1 und 2 sind die Kompensatorprinzipien, die Arbeitsbereiche sowie die mittleren Einspielfehler, getrennt nach optischen Sekundentheodoliten und elektronischen Theodoliten, zusammengestellt. Die Zahlenwerte sind Herstellerangaben.

Theodolit	Jenoptik Theo 010A	Kern DKM 2-A	Wild T2	Zeiss Th2
Kompensator-prinzip	Gelenkfeder	Flüssigkeit	Kreuzfedergelenk	Kreuzbandgelenk
Arbeitsbereich	± 4' ≙ ± 74 mgon	± 2' ≙ ± 37 mgon	± 5' ≙ ± 93 mgon	± 2' ≙ ± 37 mgon
Mittlerer Einspielfehler	<±0,3" ≙ ±0,1 mgon	<±0,3" ≙ ±0,1 mgon	±0,3" ≙ ±0,1 mgon	±0,2" ≙ ±0,06 mgon

Tab. 1: Technische Daten der Kompensatoren von optischen Sekundentheodoliten

Theodolit	AGA Geodimeter 140	Hewlett Packard 3820 A	Kern E2	Keuffel und Esser Vectron	Wild TC1	Zeiss Elta 2
Kompensator-prinzip	Pendel-sensor	Flüssigkeits-sensor	Flüssigkeits-sensor	Pendel-sensor	Pendel-sensor	Pendel-sensor
Arbeits-bereich	?	± 45 mgon	± 50 mgon	± 65 mgon	± 61 mgon	± 50 mgon
Mittlerer Einspielfehler	± 0,1 mgon	± 0,1 mgon	± 0,1 mgon	± 0,3 mgon	± 0,2 mgon	± 0,2 mgon

Tab. 2: Technische Daten der Kompensatoren von elektronischen Theodoliten

9. ZUSAMMENFASSUNG

Die zur Indexstabilisierung verwendeten Kompensatoren in Theodoliten haben einen hohen technischen Stand erreicht und arbeiten äußerst zuverlässig und präzise. Die Kompensatoren erleichtern und beschleunigen die Arbeit des Vermessungsingenieurs beträchtlich. Sie sind zu einem unentbehrlichen Bestandteil eines jeden Theodolits geworden.

LITERATUR

GIGER, A.: ***Entwicklung und Erprobung eines Präzisionsneigungstisches zur Kalibrierung geodätischer Instrumente***. Schriftenreihe des wiss. Studiengangs Vermessungswesen der HSBw München, Heft 8, Dezember 1982.

GORT, A.F.: ***Die Hewlett-Packard 3820A integrierte elektronische Meßstation***. VR 40/2, März 1978, S. 61-75.

GRÄBER, R.: ***Systematische Darstellung aller Verfahren zur Ziellinien- und Indexstabilisierung***. Unveröffentlichte Diplomarbeit des wiss. Studiengangs an der HSBw München, 1980.

LEITZ, H.: ***Zehn Jahre elektronische Tachymetrie – Zeiss: 1968 Reg Elta 14 - Elta 2 und Elta 4***. ZfV (103), Heft 9, 1978.

MÜNCH, K.H.: ***Der Elektronische Präzisions-Theodolit Kern E.2***. Presented Paper zum 17. FIG-Kongreß 1983, Sofia, Bulgaria, Kommission 5.

SCHERER, M.: ***Untersuchung von Distanzmeßteil und Reduktionssensor der Semi-Totalstation AGA Geodimeter 120***. Forum Nr. 2 II/82.

SCHWENDENER, H.R.: ***Meßsysteme und Registrierung des Wild Tachymat TC1***. Wild A.G., Heerbrugg, Schweiz.

EINIGE MODERNE VERFAHREN ZUR KALIBRIERUNG VON NIVELLIERLATTEN

von Roland TILLE

SUMMARY

The during the last ten years developed comparator equipments for the calibration of invar tape levelling rods are brought into comparison. It proves that at present three different procedures have been applied which are characterized as statical-visual, dynamical-analogous and dynamical-digital. The most important components of the comparators are analyzed regarding their influence on the accuracy and some results obtained from comparison calibrations are represented. Finally the impact of calibrations on precision levelling is shown.

ZUSAMMENFASSUNG

Die im letzten Jahrzehnt neu entwickelten Komparatoreinrichtungen zur Kalibrierung von Invarbandnivellierlatten werden gegenübergestellt. Es zeigt sich dabei, daß derzeit drei unterschiedliche Verfahren im Einsatz sind, die als statisch-visuell, dynamisch-analog und dynamisch-digital bezeichnet werden. Einige wichtige Komponenten der Komparatoren werden hinsichtlich ihres Einflusses auf die Genauigkeit analysiert, und die bei Vergleichskalibrierungen erzielten Resultate werden dargestellt. Schließlich wird die Bedeutung der Kalibrierung für das Präzisionsnivellement aufgezeigt.

1. EINFÜHRUNG

Während in der Lagevermessung heute weitgehend elektromagnetische Wellen den Maßstab der Strecken bestimmen, wird bei der geometrischen Höhenmessung noch ein materieller Maßstab, die Nivellierlatte, für die Bestimmung von Höhenunterschieden benutzt. Die auf einem Invarband aufgetragene Strichteilung verkörpert den Maßstab der Höhenunterschiede.

Bis vor wenigen Jahren wurden die Latten ausschließlich mit Normalmetern auf Komparatoren kalibriert. Als Ergebnis erhielt man ein mittleres Lattenmeter, das zur Korrektur der Höhenunterschiede verwendet wurde. Da man die Lagefehler einzelner Teilstriche des Invarbandes auf diesem Wege nicht bestimmen konnte, wurde unterstellt, daß ihre Größe vernachlässigbar klein ist.

Genauere Untersuchungen haben aber gezeigt, daß die Abweichungen der Teilstriche von ihrer Sollage bis zu 50 µm betragen und daß bei Berücksichtigung eines mittleren Lattenmeters noch individuelle Strichfehler von mehr als 10 µm übrigbleiben.

Da diese Teilungsfehler systematisch wirken und dieselbe Größe annehmen können wie die Standardabweichung der Ablesung mit einem Feinnivellier mit Planplatte, konnten diese Kalibrierverfahren nicht als befriedigend angesehen werden.

Um bei den Höhenbestimmungen in den Genauigkeitsbereich von 10^{-4} bis 10^{-5} m zu gelangen - diese Anforderungen fallen bei folgenden Beobachtungen an

- Erdkrustenbewegungen: 0,1 - 0,5 mm/a
- Fundamentsetzungen von Präzisionsmaschinen: < 1 mm
- Höhengrundlagenetze: systematische Fehler < 0,3 mm/km -

wurden im vergangenen Jahrzehnt neue Kalibrierverfahren entwickelt.

Sind auf herkömmliche Weise einzelne Abschnitte einer Latte stichprobenartig visuell mit einem Meternormal verglichen worden, so werden jetzt die individuellen Fehler aller Teilstriche weitgehend automatisch mit einem Laserinterferometer als Sekundärnormal bestimmt. Die innere Genauigkeit ist bei den meisten Verfahren dabei deutlich besser als 10 µm.

Der Aufbau EDV-gestützter, automatisch ablaufender Kalibrierverfahren gelang

zuerst beim Geodätischen Institut der Universität Karlsruhe (***SCHLEMMER*** 1975) und bei der Physikalisch-Technischen Bundesanstalt in Braunschweig (***GRABE***, ***LEBOWSKY*** 1979). Im deutschsprachigen Raum arbeiten heute fünf weitere Institute mit oder an ähnlichen Verfahren.

2. ÜBERSICHT ÜBER DIE BERÜCKSICHTIGTEN KOMPARATOREN

Am hiesigen Institut wurden 1982/83 drei Maßstabsproben kompariert, die anschließend auf den Komparatoren von vier anderen Instituten untersucht wurden. Die Messungsergebnisse wurden analysiert und anläßlich des "Workshop on Precise Levelling" im März 1983 in Hannover vorgestellt (***HEISTER***, ***CASPARY***, ***TILLE*** 1983). Die dabei erarbeitete Zusammenstellung der eingesetzten Komparatoren wird im folgenden um die Angaben zweier weiterer Institute ergänzt. Damit sind, soweit dem Autor bekannt, alle im deutschsprachigen Raum entwickelten Komparatoren, die zur rationellen Ermittlung von Strichverbesserungen geeignet sind, erfaßt (siehe Tabelle 1).

In den folgenden Kapiteln werden sodann die wesentlichen Komponenten der Komparatoren verglichen und diskutiert.

3. DAS LÄNGENNORMAL

Alle Institute verwenden als Längennormal einen stabilisierten Helium-Neon-Laser, der sich wegen der folgenden Eigenschaften bestens bewährt hat (***KUNZMANN*** 1979).

Die Laserstrahlung kann mit optischen Bauteilen flexibel in beliebige Richtungen gelenkt oder in zwei bis drei Strahlen aufgeteilt werden. Auf dem Komparator sind daher mit einer Lichtquelle mehrere Meßachsen zu verwirklichen. Die kompakte Bauweise (Laserkopflänge ca. 300 mm), die Leistungsaufnahme von etwa 15 Watt und die für das menschliche Auge unschädliche Strahlungsleistungen von etwa 1 mW sind weitere Vorzüge für ein Laborgerät.

Um aus der Laserstrahlung Längen interferometrisch abzuleiten, gibt es zwei Möglichkeiten: die Überlagerung von Wellenamplituden (Michelson-Prinzip) und die Überlagerung von Frequenzen (Doppler-Prinzip). Die Umsetzung der Interferenzen in Meßwerte geschieht jeweils mit Photoempfängern und nachgeschalteter Verstärker- und Zählelektronik.

Institute	RWTH Aachen	PTB Braunschweig	Univ. Karlsruhe
Längennormal	He-N-Laser	He-N-Laser	Zweimoden He-N-Laser
Hersteller	Meßsystem M 100 SORO METRILAS	Stabilisierung und Interferometersystem von PTB entwickelt	Interferometer Eigenbau, Frequenzstabilisierung nach PTB
Frequenzstabilität	$\pm \left(10^{-7} - 10^{-8}\right)$	$\geq 10^{-7}$	ca. 10^{-7}
Atmosphärische Korrektur	manuell automat. möglich	Wellenlänge nach Edlen Temperatur ± 0,1°C Druck ± 0,5 mbar	Erfassung und rechnerische Korrektion
Laborbedingungen	temperaturstabil auf ± 1°C nicht klimatisiert	temperiert ± 0,2°C	thermisch isoliert ± 3°C nicht klimatisiert
Auflösung	0,1 µm	$\lambda/1$, $\lambda/2$, $\lambda/4$, $\lambda/8$	0,3 µm
Stricherfassung	Photoelektrisches Mikroskop 'Schlemmer' (s.a. Uni Karlsruhe), zusätzlich: Optisch-visuelles Verfahren mit Projektionsmikroskop ISOMA P 123 (s.a. ETH Zürich)	Photoelektrisches Mikroskop, Vergrößerung 5x, Spalt 10 - 100 µm Beleuchtung mit frequenzstabilisiertem He-Ne-Laser	Photoelektrisches Mikroskop 'Schlemmer', Vergrößerung 1x, Auflichtbeleuchtung mit Hg-Lampe über Lichtleiter und Ringlicht
Koinzidenz	s. Univ. Karlsruhe zus. manuell/visuell	Auswertung digitaler Strichprofile mit ≤ 5000 Wertepaaren/s	Symmetrievergleich der Lichtintensität auf einer Duo-Diode
Methode	dynamisch-analog statisch-visuell	dynamisch-digital	dynamisch-analog
Geschwindigkeit	0,08 mm/s - 0,6 m/min	0,03 mm/s - 6 mm/s	2 mm/s - 10 mm/s
Meßdauer (3-m-Latte, 1 cm Doppeltlg. Hin-Rück)	phot. 1,5 h vis. 8,5 h	1 h	2 h
Steuerung des Meßablaufs	voll- oder teilautomat. mit cbm 3032	Interaktive Steuerung mit Prozeßrechner Siemens R 30	automatisch
Lagerung des Prüflings	horizontal nach Abbé	horizontal auf zwei Unterstützungen	horizontal auf Meßwagen in Besselpunkten, Korrektionen wegen Nichtparallelität
Abstand Laser-Teilungsachse	< 1 mm	50 mm	95 mm
Genauigkeit			
Standardabweichung einer Stricherfassung (Reproduzierbarkeit)	s_0 (phot.) = 2-10 µm s_0 (vis.) = 2 µm	< 5 µm	2 µm
Standardabweichung einer Strichverbesserung	s_0 (phot.) = 10 µm s_0 (vis.) = 5 µm	10 µm/m	10 µm
Komparatorbank	Holzkastenkonstruktion, Meßschlitten in Kugelbüchsenführung auf zwei Stahlwellen 25,2 m	Gußeiserne Schiene, Meßwagen auf Wälzlagern und Führungsschiene 6 m	Doppel-T-Träger, Meßwagen auf Lauf- und Führungsfläche aus gezogenem Flachstahl 5 m

Tabelle 1 Zusammenstellung der berücksichtigten Komparatorsysteme (ausführliche Beschreibungen folgen als Anhang diesem Beitrag)

HSBw München	Techn. Univ. München	Techn. Univ. Wien	ETH Zürich
Zweifrequenz He-Ne-Laser	Zweifrequenz He-Ne-Laser	Zweifrequenz He-Ne-Laser	Zweifrequenz He-Ne-Laser
Laser-Doppler-Interferometersystem 5501 A Hewlett-Packard	Laser-Doppler-Interferometersystem 5526 A Hewlett-Packard	Laser-Doppler-Interferometersystem 5526 A Hewlett-Packard	Laser-Doppler-Interferometersystem 5526 A Hewlett-Packard
$\pm 5 \cdot 10^{-7}$	$\pm 1 \cdot 10^{-8} - 10^{-9}$	$\pm 1 \cdot 10^{-7} - 10^{-8}$	$\pm 5 \cdot 10^{-7}$
Automatisch mit 5510 A Temperatur ± 0,1 °C Druck ± 0,75 mbar Feuchte ± 10 %	mehrere Thermoelemente d. Fa. SIKA, ± 0,1°C Druck Paulin Barometer ± 0,1 Torr mit rechnerischer Korrektur	automat. Kompensation über hp-System	Manuelle Kompensation Temperatur ± 0,1°C Druck ± 0,4 mbar Feuchte ± 10 %
klimatisiert ± 1°C	nicht klimatisiert	klimatisiert	klimatisiert ± 0,5°C
z.Z. 0,16 µm (0,016 µm mögl.)	0,1, 0,03, 0,01 µm	0,1 µm	0,1 µm
Photoelektrisches Mikroskop LEITZ MPV-compact, Vergrößerung 4-20x, variable Spaltbreite, Auflichtbeleuchtung durch Hg-Lampe, Lichtleiter und Ringkondensor	Photoelektrisches Mikroskop 'Schlemmer' Auflichtbeleuchtung durch Hg-Lampe, Glasfaserleitung und Ringkondensor	Videokamera und Monitor, variable Meßmarken durch Ausblenden von Monitorzellen, Vergrößerung 10x	Optisch-visuell mit Projektionsmikroskop ISOMA, Vergrößerung 10x, 20x, individuelle Strichmaske, Auflichtbeleuchtung durch Hg-Lampe mit 2 Lichtleitern
Auswertung digitaler Intensitätsprofile mit 200 Wertepaaren/mm	Symmetrievergleich der Lichtintensität auf der Duo-Diode	visuell/manuell	visuell/manuell
dynamisch-digital	dynamisch-analog	statisch-visuell	statisch-visuell
0,005 mm/s - 10 m/min	bis 5 mm/s	-	-
3 h	1,5 h	5 h	2,5 h
Automatische Steuerung über hp 9825 A	Vollautomatisch über Steuerung der Antriebsmöglichkeiten	manuelle Auslösung der Datenübernahme PET 2000	manuell im Dialog mit cbm 8032
horizontal auf Meßschlitten, andere Anordnungen möglich	Horizontalkomparator: verschieden Vertikalkomparator: nach Abbé	horizontal Korrektur wegen Nichtparallelität	horizontal verschiedene Anordnungen möglich
< 1 mm	< 1 mm	keine Angaben	< 10 mm
1 µm	1 - 2 µm	5 µm	3 - 5 µm
5 µm	1 - 2 µm	10 µm	keine Angaben
Feingeschliffene Granitblöcke, Meßschlitten auf Luftlager 30 m	Stahlgußbank, Meßschlitten mit Kugellager auf zwei feingeschliffenen Führungsschienen 3 m horizontal Schienensystem mit 2 Meßschlitten 7 m vertikal	Aluminium-T-Träger, Meßschlitten mit Kugellagerbuchsen auf 2 Stahlwellen 4,5 m	IPE-400-Träger, Meßwagen auf zwei geschliffenen Rund- und Flachführungsschienen 9 m

Untersuchungen haben gezeigt, daß die Wellenlänge kommerziell angebotener He-Ne-Laser bei ordnungsgemäßem Betrieb eine Stabilität von 10^{-7} erreicht (***KUNZMANN*** 1979).

Das Problem liegt darin, diese Genauigkeit auf den Meßvorgang zu übertragen.

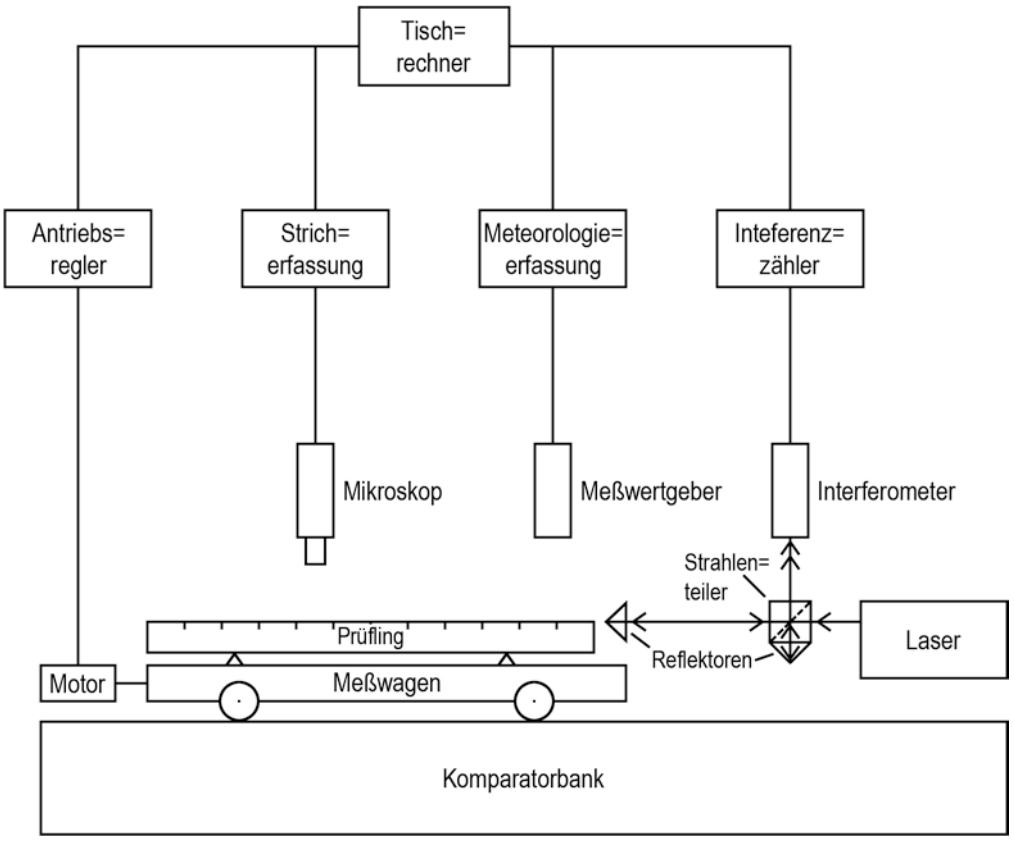


Abb. 1 Prinzipskizze eines computergesteuerten Nivellierlatten-komparators

Der Laserstrahl ändert seine Ausbreitungsgeschwindigkeit, wenn sich der Brechungsindex n der Atmosphäre ändert. Der Brechungsindex kann nach verschiedene Formeln unter Berücksichtigung von Lufttemperatur T, Luftdruck P, relativer Feuchte RF oder Wasserdampfdruck berechnet werden. Abschätzungen zeigen, daß Temperatur, Druck und Feuchte den Brechungsindex unter normalen Bedingungen (20°C, 1013 mbar, 50 % RF) wie folgt ändern:

$$\frac{dn}{dT} = -1 \cdot 10^{-6}\ °C^{-1}$$

$$\frac{dn}{dP} = 3 \cdot 10^{-7}\ mbar^{-1}$$

$$\frac{dn}{dF} = -3 \cdot 10^{-9}\ (\%\ RF)^{-1}$$

Alle Einflüsse sind längenproportional. Soll bei der Kalibrierung von Nivellierlatten der meteorologische Einfluß nicht mehr als 3 µm auf 3 m = $1 \cdot 10^{-6}$ ausmachen, so muß die Temperatur auf 1°C und der Druck auf 3 mbar erfaßt werden. Der Einfluß der Feuchte kann vernachlässigt werden.

Die einzelnen Institute haben unterschiedliche Wege beschritten, um die meteorologischen Daten während der Kalibrierung zu erfassen. Da die Komparatorräume jedoch fast alle klimatisiert oder klimastabil sind, ist dies von untergeordneter Bedeutung. Probleme sind nur dann zu erwarten, wenn sich die Temperatur während der Kalibrierung wesentlich ändert.

4. DIE STRICHERFASSUNG

Die Komparierung aller Teilstriche einer Nivellierlatte erfordert 600 bzw. 1200 Beobachtungen. Wird dazu ein visuelles Verfahren benutzt, so bedeutet dies für den Beobachter mindestens vier Stunden konzentrierter Arbeit. Es ist daher wichtig, ergonomisch günstige visuelle Verfahren einzusetzen. Bewährt haben sich Projektionsmikroskope und Bildschirme.

Als Alternative ist ein beobachterfreies photoelektrisches Verfahren zu sehen, das außerdem eine vollständige Automatisierung des Kalibriervorganges ermöglicht. Die bei den in der Tabelle zusammengestellten Komparatoren eingesetzten Verfahren zur Stricherfassung lassen sich in drei Gruppen einteilen:

1. Statisch-visuelle Verfahren

   Der Teilstrich wird automatisch oder manuell grob angefahren und visuell koinzidiert. Danach wird seine Istlage registriert (TU Wien, ETH Zürich, RWTH Aachen).

2. Dynamisch-analoge Verfahren

   Durch Latten- oder Mikroskopvorschub wird die Teilung ohne Unterbrechung abgefahren. Die Teilstriche werden mit einem lichtempfindlichen Sensor erfaßt, der Licht in elektrische Spannungssignale umwandelt. Durch die Anordnung von zwei Sensoren wird die Strichmitte über eine Symmetrieeinstellung erhalten. Gleichzeitig wird die Strichlage registriert (RWTH Aachen, Uni Karlsruhe, TU München).

3. Dynamisch-digitale Verfahren

   Die Teilung wird wie bei den dynamisch-analogen Verfahren abgefahren und abgetastet. Das Spannungssignal wird in seinem Verlauf digitalisiert. Es entsteht ein Intensitätsprofil des Striches. Durch ein mathematisches Modell wird aus dem Profil die Strichmitte abgeleitet (siehe Abbildung 2).

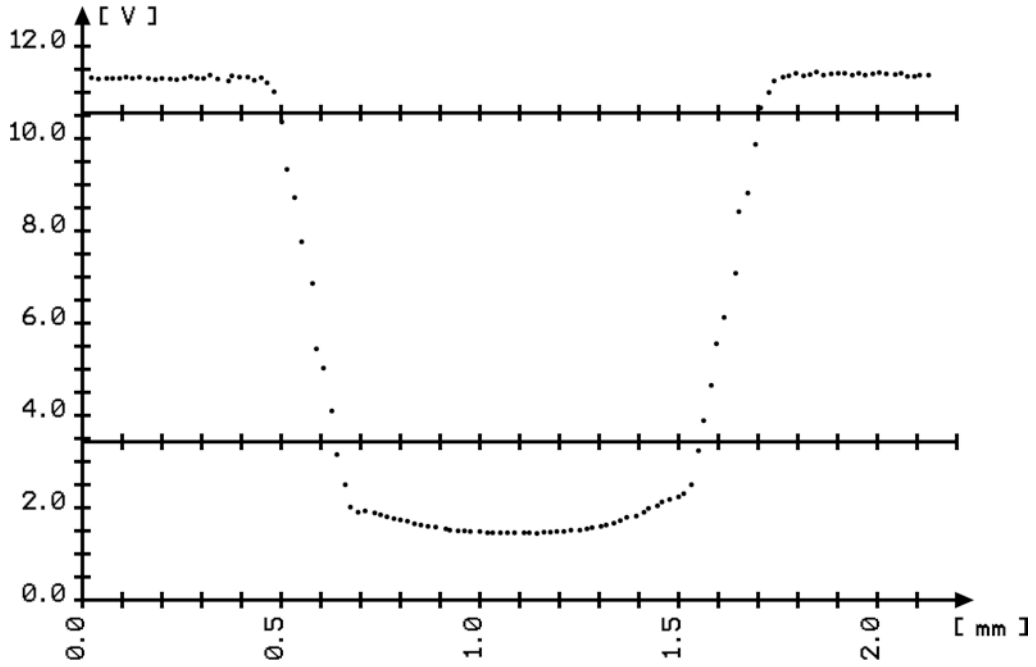


Abb. 2 Intensitätsprofil eines Striches

Das statisch-visuelle Verfahren erfordert den geringsten Aufwand. Ein einfaches Projektionsmikroskop und auf Folie gezeichnete Strichmasken reichen zum Stricheinfang aus. Auf diesem Wege wird im Labor eine Stricheinteilung möglich, die direkt mit der Feldbeobachtung vergleichbar ist. Als Nachteil seien der hohe Zeitaufwand und die damit verbundene Ermüdung des Beobachters genannt, die das Kalibrierungsergebnis beeinflussen können.

Bei den dynamisch-analogen Verfahren wird das von ***SCHLEMMER*** (1975) entwikkelte photoelektrische Mikroskop benutzt. Dieses Mikroskop löst im Moment des Strichmittensignals der Duo-Diode eine Lasermeßwertaufnahme aus. Dieses Verfahren erlaubt eine hohe Abtastgeschwindigkeit (eine Lattenteilung im Hin- und Rückgang ca. 0,5 h) bei geringem Datenanfall (z.B. Strichnummer, Lage, atmosphärische Daten). Besondere Sorgfalt erfordert das Ausleuchten der Teilung, um sicherzustellen, daß der für die Messung notwendige Kontrast gewährleistet ist. Das Abtasten von Strichen, die durch Beschädigung asymmetrisch ausgebildet sind, führt zu Werten, die nicht unbedingt mit der visuellen Mitteneinteilung übereinstimmen müssen.

Bei dem dynamisch-digitalen Verfahren wird das Spannungssignal in mehrere diskrete Meßwerte aufgelöst. Dadurch erhält man zusätzlich eine Information über die Strichqualität. Durch geeignete mathematische Modelle wird die Strichmitte festgelegt, wobei Unterschiede in der Ausbildung und Qualität der Striche berücksichtigt werden können. Um die große Menge anfallender Daten möglichst in Echtzeit zu bewältigen, sind schnelle Interfaces, schnelle Digitalvoltmeter und Assemblerprogrammierung notwendig.

Die am hiesigen Institut durch Vergleichsmessungen an drei Maßstabsproben gewonnenen Erkenntnisse über die Genauigkeit der Stricherfassung seien kurz zusammengefaßt.

Die Standardabweichungen der Strichverbesserungen aus 16 Kalibrierungen von drei Maßstäben (innere Genauigkeit) lagen zwischen 2 und 9 µm. Bei einigen Meßreihen traten aber noch ein oder zwei Ausreißer auf ($< 3\sigma$), deren Ursache ungeklärt ist (***HEISTER***, ***CASPARY***, ***TILLE*** 1983).

Die mittleren Fehler der drei Kalibriermethoden, bezogen auf eine visuelle Referenzkalibrierung (Mittel aus 17 Beobachtungsreihen), nahmen folgende Werte an:

statisch-visuell	± 1 - 5 µm
dynamisch-analog	± 6 - 8 µm
dynamisch-digital	± 4 - 10 µm

Die photoelektrischen Methoden weisen noch systematische Differenzen zu den visuellen Verfahren auf, die bei extrem gestörten Strichen bis zu 20 µm betragen. Bei der dynamisch-digitalen Methode kann die Profilanpassung an die visuelle Strichlage noch verbessert werden.

5. DATENERFASSUNG UND DATENVERARBEITUNG

Bei allen angegebenen Kalibrierverfahren wird die Bestimmung der Strichverbesserungen durch leistungsfähige Tischrechner unterstützt. Wie die Gerätekonfiguration am Institut für Geodäsie der Hochschule der Bundeswehr München zeigt, reicht schon ein Rechner mit 16 kByte Kapazität aus, um den Kalibriervorgang vollständig zu automatisieren. Dazu sind in Echtzeit folgende Daten zu verarbeiten:

- Laserinterferometerwerte
- analog/digital gewandelte Spannungswerte (Mikroskop)
- digital/analog gewandelte Spannungswerte (Motorsteuerung)
- Brechungsindex aus meteorologischen Daten für die Lichtgeschwindigkeit
- Daten an Anzeigegeräten und Druckern

Dabei wird ein Datenbussystem eingesetzt, das Übertragungsraten von 40 kByte/s und Zyklen von Rechnerroutinen von 50 Hz aufweist. Je höher diese Werte liegen, desto schneller kann der automatisch gesteuerte Kalibriervorgang ablaufen.

Die als Ergebnis der Kalibrierung anfallenden Strichverbesserungen sind so abzuspeichern, daß sie von dem Rechner gelesen werden können, mit dem die Feldbeobachtungen korrigiert und die Ausgleichung durchgeführt werden. Nur dann ist die Strichkalibrierung wirtschaftlich zu verwerten (***GESSLER*** 1979).

6. DIE LAGERUNG DES PRÜFLINGS

Bei einem idealen Längenkomparator ist das ***ABBÉ***sche Komparatorprinzip verwirklicht, d.h. die Meßachse und die Prüflingsachse liegen auf einer Geraden (***ABBÉ*** 1890). Diese Idealanordnung ist beim Nivellierlattenkomparator aus mehreren Gründen nicht realisierbar.

Unter dem Einfluß der Strahlenbrechung verläuft der Laserstrahl nicht als Gerade, sondern als eine sich kontinuierlich verändernde Kurvenlinie. Diese Veränderungen können am Reflektor beobachtet werden, auf dem sich der Laserpunkt mit einem Radius von wenigen Millimetern bewegt.
Auch die Achse der auf dem Invarband aufgetragenen Teilung wird keine ideale Gerade bilden, denn das gespannte Band wird zwischen den Nuten oder Halterungen durchhängen. Dazu kommen Durchbiegungen des Lattenrahmens, die nachweislich der Untersuchungen von ***GOTTWALD*** und ***WITTE*** (1983) bei unzureichender Unterstützung zu einem Banddurchhang von 2 mm und mehr führen können. Diese Abweichungen der Teilungsachse von einer Geraden bewirken einerseits einen Maßstabsfehler, andererseits aber auch unterschiedliche Objektabstände zum Mikroskop, wodurch die Abbildungsschärfe verloren geht und negative Auswirkungen auf die Genauigkeit der Stricherfassung zu erwarten sind. Diesen Problemen kann dadurch entgegengewirkt werden, daß man bei liegender Latte den Durchhang meßtechnisch erfaßt und in einem funktionalen Modell seinen Einfluß berücksichtigt. Eine weitere Möglichkeit besteht darin, die Latte in vertikaler Stellung zu kalibrieren, wie es derzeit nur an der Technischen Universität München möglich ist.

Bei angenäherter Anordnung nach dem ***ABBÉ***schen Prinzip ist der Einfluß der Lattenführung von untergeordneter Bedeutung. Bei allen anderen Anordnungen kann dieser Einfluß ausreichend genau erfaßt und berücksichtigt werden. Es wird daher in diesem Beitrag verzichtet, darauf näher einzugehen, statt dessen wird auf ***BRANDSTÄTTER*** (1982) verwiesen.

7. DIE TEMPERATURERFASSUNG AM PRÜFLING

Ein immer noch aktuelles und oft vernachlässigtes Problem sowohl im Labor wie im Feld ist die Bestimmung der Temperatur des Invarbandes. Während es in einem temperaturstabilen Meßraum möglich ist, durch rechtzeitige Lagerung zu erreichen, daß sich die Temperatur des Prüflings an die Umgebungs-

temperatur angleicht, ist beim Feldeinsatz mit unkontrollierten Temperaturschwankungen zu rechnen. Neuere Untersuchungen zeigen (***ZIPPELT*** 1983), daß die Invarbänder durch Sonneneinstrahlung bis zu 5°C höhere Temperaturen aufweisen als die umgebende Luft und daß außerdem kurzzeitige Temperaturschankungen von der gleichen Größenordnung beim Wechsel von Sonne in Schatten und umgekehrt auftreten.

In Zukunft wird daher in ganz besonderem Maße das Augenmerk auf die Erfassung der aktuellen Invarbandtemperatur im Felde zu richten sein.

8. DIE GENAUIGKEIT DER STRICHVERBESSERUNGEN UND IHRE BERÜCKSICHTIGUNG BEIM NIVELLEMENT

Auf dem Workshop on Precise Levelling in Hannover 1983 wurde die Empfehlung ausgesprochen, Strichverbesserungen mit einer Genauigkeit von 10 µm pro Lattenmeter zu bestimmen. Während die Standardabweichungen der Strichverbesserungen, ermittelt aus Wiederholungsmessungen, deutlich unter 10 µm liegen, sind über die absolute Kalibriergenauigkeit nur sehr schwer zuverlässige Aussagen zu gewinnen. Nach einer Abschätzung von ***GRABE*** und ***LEBOWSKY*** (1979) ist unter Einschluß von Führungsfehlern, Strichauffassungs- und -auswertefehlern sowie Fehlern in der Lattenlagerung mit einer Gesamtstandardabweichung von etwa 30 µm für die Strichverbesserungen zu rechnen. Andere Abschätzungen (siehe Tabelle 1) sind bedeutend optimistischer.

Beim Wiederholungsnivellement des Deutschen Haupthöhennetzes 1980 - 1985 wurden erstmalig Einzelstrichverbesserungen berücksichtigt. Auf Empfehlung der Arbeitsgemeinschaft der Vermessungsverwaltungen wurden bei der Physikalisch-Technischen Bundesanstalt 94 Nivellierlatten kalibriert.

Die Standardabweichungen im bis jetzt gemessenen ersten Drittel des Netzes, berechnet aus Differenzen zwischen Hin- und Rückweg, liegen zwischen 0,1 - 0,4 mm/km (***STÜRZER*** 1983). In dem Nivellementsnetz 1960 lagen die vergleichbaren Fehler bei 0,33 - 0,41 mm/km. Die Schleifenschlußfehler von umgerechnet 0,7 mm/km lassen aber vermuten, daß sich damals systematische Fehler in den Messungen summiert haben (***DRAHEIM*** 1979). Vergleichswerte des neuen Netzes liegen derzeit nicht vor.

Von den neuen Lattenkalibrierungen wird erhofft, daß sie zur Minderung dieser systematischen Fehler beitragen.

9. SCHLUSSBEMERKUNGEN

Die vorgestellten Entwicklungen informieren über den derzeitigen Stand der Kalibriertechnik für Nivellierlatten. Aber auch die Entwicklung der Nivellierlatten selbst ist nicht stehengeblieben. Als Teilungsträger werden Carbonbänder getestet, die einen sehr geringen Ausdehnungskoeffizienten von etwa $1 \cdot 10^{-6}$ haben, aber empfindlich auf Feuchtigkeit reagieren (***FISCHER*** 1979). ***SCHLEMMER*** (1983) berichtete von einer neuen Teilungsanlage, die die Teilungsstriche interferometrisch gesteuert auf den Teilungsträger aus Invar mit Laserenergie einbrennt. Nach diesem Verfahren soll der mittlere Teilstrichfehler kleiner als 5 µm sein.

DANKSAGUNG

Alle genannten Institute, die durch ihre freundliche Unterstützung bei der Zusammenstellung der technischen Angaben behilflich waren, sei an dieser Stelle herzlich gedankt.

Es sei noch darauf hingewiesen, daß am Institut für Geodäsie der Hochschule der Bundeswehr München das Forschungsprojekt "Optimaler Einfang von Strichlagen durch photoelektrische Messungen" von der Deutschen Forschungsgemeinschaft gefördert wird.

Geodätisches Institut
der Rheinisch-Westfälischen Technischen Hochschule Aachen

Längennormal: He-Ne-Laser M 100 der Fa. SORO METRILAS, Michelson-Interferometer, Frequenzstabilität ca.$\pm 10^{-7}$ bis 10^{-8} (Werksangaben). Kompensation meteorologischer Einflüsse z.Z. manuell, automatische Kompensation möglich. Auflösung des Längennormals 0,1 µm. Meßlabor auf ± 1°C temperaturstabil, nicht klimatisiert.

Komparatorbahn: 25,2 m, justierbare Holzkastenkonstruktion.
Meßschlitten in Kugelbüchsenführung zwangsfrei verschiebbar.
Antrieb des Meßschlittens je nach Aufgabe über Seilzug oder Gewindespindel mittels Gleichstrommotor.

Stricherfassung: Photoelektrisch = 'Schlemmer'-Mikroskop.
Visuell = Projektionsmikroskop ISOMA P123. Strichplatten können auf die Strichbreite des Prüflings abgestimmt werden.
Beleuchtung: Hg-Lampe mit Lichtleiter und Ringkondensor oder zweifach Punktlicht. Fahrgeschwindigkeit: Seilzug (für EDM) bis 15 m/min, Spindel (für Niv.kal.) 0,05 bis 0,6 m/min.

Meßdauer: 3m-Latte photoelektr. ca. 1,5 h, visuell ca. 8,5 h.

Datenerfassung und -verarbeitung: Commodore CBM 3032/4032, 32 kByte, Floppy-Disk 2x170 kByte, Matrix-Drucker, Plotter hp 9872 C.
Interface: IEEE 488-Bus, 12-bit-Parallel-Bus (Eigenentwicklung).
Programmiersprachen: MICROSOFT-Basic, Assembler.

Steuerung des Meßablauf: Voll- und teilautomatisch mit CBM 3032.
Peripherie (nur für Niv.kal.): Schaltnetzteil (Motorsteuerung), Interferometer (Display), photoelektrisches Mikroskop (Auswerte-Elektronik).

Anordnung des Prüflings: nach *ABBÉ*

Temperaturerfassung am Prüfling: über Thermistoren möglich (± 0,1°C).

Genauigkeitsmaße: Stricherfassung visuell ca. ± 2 µm (Standardabweichung des Mittels aus sechs Koinzidenzen), photoelektrisch ± (2 - 10) µm in Abhängigkeit von der Verfahrgeschwindigkeit.
Reproduzierbarkeit: (Standardabweichung einer Strichverbesserung aus Doppelmessungen) visuell = ± 5 µm, photoelektrisch = ± 10 µm.

Physikalisch-Technische Bundesanstalt Braunschweig
- Industrielle Längenmeßtechnik -

Längennormal: He-Ne-Laser, Frequenzstabilität $< \pm 10^{-7}$, Zählerschritte $\lambda/1$, $\lambda/2$, $\lambda/4$, $\lambda/8$, Ausgabe in Dualzahlen, Berechnung der Wellenlänge nach Edlén unter Berücksichtigung der Lufttemperatur (± 0,1°C) und des Luftdrucks (± 0,5 mbar). Temperierter Meßraum (± 0,2°C).

Komparatorbank: Sechsmeter-Schiene aus Gußeisen, zweifach unterstützt. Meßwagen auf Wälzlagern, Antrieb durch Motor mit Tachogenerator über Schneckenrad auf Kugellager. Fahrgeschwindigkeit stufenlos regelbar von 0,03 bis 6 mm/s.

Stricherfassung: Rechnergestütztes dynamisches Meßsystem. Objektivvergrößerung 5x, Auflichtbeleuchtung mit stabilisiertem He-Ne-Laser, Registrierung des rückgestreuten Lichtes über Spaltblende (10 - 100 µm)-Fotoelement-Meßverstärker-Interface und Prozeßrechner, Datendichte der Intensitätsprofile $I(n\lambda) \leq 5000$ Wertepaare/s.

Meßdauer: für eine 3m-Latte 1 h.

Datenerfassung und -verarbeitung: Siemens R30 Prozeßrechner mit 128 kWorten (à 16 bit) Arbeitsspeicher, zwei Plattenspeicher mit 5 und 25 MWorten, Tektronix 4051 Terminal. Einzug der Längenkoordinaten und der Lichtintensitäten über Interface FMRB 578. Programmiersprachen Fortran/Assembler.

Steuerung des Meßablauf: Software-Paket mit Dialogbetrieb unter Einschluß der Motorsteuerung (Lernphase, Messung, Auswertung, Ausgabe von Strichlagen, Teilungsfehlern und Strichbreiten, ferner graphische und numerische Darstellung einzelner Profile, Rettungsdialog).

Anordnung des Prüflings: Horizontal auf zwei Unterstützungen, Normal und Prüfling in 50 mm Abstand übereinander. Messungen in zwei um 180° gegeneinander gedrehten Lagen, sofern möglich.

Temperaturerfassung am Prüfling: Hg-Thermometer (Ablesung ± .02°C).

Genauigkeitsmaße: Reproduzierbarkeit der Stricherfassung $< \pm 5$ µm, Unsicherheit gemessener Teilstrichkorrektionen in Abhängigkeit von der Lattenqualität $\geq \pm 10$ µm/m.

Geodätisches Institut der Universität Karlsruhe

Längennormal: Zweimoden He-Ne-Laser, Modell 132 der Fa. SPECTRA PHYSICS, Frequenzstabilisierung auf ca.$\pm 6 \cdot 10^{-7}$ mit dem Verfahren "Kunzmann, Lebrowsky, PTB". Interferometer Eigenbau. Meteorologische Einflüsse werden erfaßt und bei der Auswertung berücksichtigt. Labor nicht klimatisiert, aber thermisch isoliert auf ca. ± 3°C.
Auflösung des Längenmeßsystems 0,3 µm.

Komparatorbank: Fünfmeter-Doppel-T-Träger auf zwei Betonsockeln.
Die Laufflächen des Meßwagens bestehen aus gezogenem Flachstahl. Der Meßwagen besteht aus zwei je zweirädrigen Wagenteilen, die in einem Abstand von 1,46 m durch zwei Stahlstangen verbunden sind.
Antrieb des Meßwagens über Stahlseil mit einem gleichstromgesteuerten Motor.

Stricherfassung: Photoelektrisches Mikroskop Eigenbau.
Abbildung des Striches im Maßstab 1 : 1 auf einer Duodiode.
Auflichtbeleuchtung mit Hg-Lampe über Lichtleiter und Ringlicht.
Vergleich der Symmetrielage des Striches durch Intensitätsvergleich aus der Duodiode. Dynamische Registrierung mit 2 mm/s Meßgeschwindigkeit (möglich bis 10 mm/s).

Meßdauer: für 3m-Latte Doppelteilung Hin- und Rückmessung ca. 2 h.

Datenerfassung und -verarbeitung: Minc-11 Laborcomputer (PDP-11/03) mit RX 02-Floppy-Doppellaufwerk, VT 105, 64 kByte, Epson MX 80 Printer.
Programmiersprachen: Basic, Fortran IV.

Steuerung des Meßablauf: k.A.

Anordnung des Prüflings: Horizontal auf dem Meßwagen, in den Besselpunkten unterstützt. Longitudinalkomparator, wobei das *ABBÉ*sche Prinzip nicht eingehalten ist. Korrekturwerte werden ermittelt und angebracht.

Temperaturerfassung am Prüfling: ---

Genauigkeitsmaße:

Standardabweichung einer Stricherfassung	s_0	± 2 µm
Standardabweichung einer Strichverbesserung (Reproduzierbarkeit)	s	± 10 µm

Institut für Geodäsie
der Hochschule der Bundeswehr München

Längennormal: Zweifrequenz He-Ne-Laser, Teil des Laser-Doppler-Interferometers 5501 A der Fa. Hewlett Packard.
Frequenzstabilität lt. Hersteller $\pm 0{,}5 \cdot 10^{-6}$ bzw. $\pm 1{,}5 \cdot 10^{-6}$ mit automatischer Kompensation meteorologischer Einflüsse durch hp 5510 A Kompensator (Temperatur ± 0,1°C, Druck ± 0,75 mbar, Feuchte ± 10 %).
Labor klimatisiert ± 1°C. Auflösung des Längenmeßsystems 0,16 - 0,016 µm.

Komparatorbank: 30 m Granitbank aus sechs Blöcken. Jeder Block auf acht Pfeilern justierbar gelagert. Meßschlitten durch Luftlager auf geschliffenen Granitflächen geführt. Antrieb des Meßschlittens mit gleichstromgesteuertem Motorschlitten mit Winkelschrittgeber und zwangsfreier Koppelung.

Stricherfassung: Photoelektrisches Mikroskop MPV-compact der Fa. Leitz. 5 bis 20fache Vergrößerung, variable Spaltblende, Auflichtbeleuchtung durch Hg-Lampe über Lichtleiter und Ringkondensor. Digitales Intensitätsprofil mit 200 - 300 Wertepaaren/mm. Dynamische Registrierung mit ca. 0,5 mm/s Meßgeschwindigkeit und höherer Fahrgeschwindigkeit (10 m/min für EDM-Kalibrierung).

Meßdauer: ca. 3 h für 3m-Latte.

Datenerfassung und -verarbeitung: hp 9825 A mit 64 kByte, Kassette mit ca. 100 kByte, hp Interface-Bus (IEEE 488-1975). Plotter hp 9872B, Printer 9871 A. Programmiersprache 'hpl'.

Steuerung des Meßablauf: Vollautomatisch mit Controller hp 9825 A. Digitalvoltmeter hp 3456A, D/A-Wandler hp 59501A für Motorschlitten-Antrieb, Quarzthermometer hp 28041 als Temperaturreferenz.

Anordnung des Prüflings: in *ABBÉ*scher Anordnung Abstand Teilungsachse - mittlere Interferometerachse < 1 mm.
Verschiedene Anordnungen möglich.

Temperaturerfassung am Prüfling: Thermistor hp 10563 A (± 0,1°C bei 13 - 40°C). Automatische Kompensation mit hp 5510A Kompens.

Genauigkeitsmaße: Standardabweichung einer Stricherfassung ± 1 µm.
Standardabweichung einer Strichverbesserung ≥ ± 5 µm.
Individuelle Genauigkeitsanalyse ist vorgesehen.

Geodätisches Institut der Technischen Universität München

Längennormal: Zweifrequenz He-Ne-Laser. Laser-Doppler-Interferometer hp 5526 A von Hewlett Packard. Langzeitstabilität ± $1 \cdot 10^{-8}$, Kurzzeitstabilität ± $1 \cdot 10^{-9}$, Wellenlänge von 632.8 nm ≈ ± $2 \cdot 10^{-9}$ genau (PTB). Meßgenauigkeit abhängig von der Erfassung der atmosphärischen Bedingungen (Druck = ± 0,1 Torr, Trocken-, Feuchttemp. = ± 0,1°C). Temperaturerfassung über Thermoelemente der Fa. SIKA an mehreren Stellen. Druckerfassung an Paulin-Barometern, die am Standbarometer kalibriert werden. Nachträgliche Berücksichtigung im Verlauf der Auswertung, soweit nicht durch Meßaufbau kompensiert. Auflösung des Längenmeßsystems 0,1 µm, 0,01 µm, 0,003 µm.

Komparatorbank Horizontal: Dreimeter-Stahlgußbank der Fa. Hommel. Meßschlitten mit Kugellager geführt. Führungsgenauigkeit < 0,3 mm. Antrieb des Meßschlittens über Gleichstrommotor. Geschwindigkeit regelbar.
Vertikal: Siebenmeter-Schienensystem mit zwei Meßschlitten, verbunden durch 3,40 m lange Aluminiumschiene zur Aufnahme von Invarlatten. Antrieb über Spindel und Schrittmotor. Geschwindigkeit regelbar.

Stricherfassung: Photoel. Mikroskop 'Schlemmer'. Auflichtbeleuchtung durch Hg-Lampe mit Glasfaserltg. und Ringkondensor (Kaltlicht). Stricherfassung mit Doppeldiode (Symmetrieabfrage). Dynamische Registrierung mit 5 mm/s.

Meßdauer: 3m-Latte 0,5 cm Doppelteilung Hin und Rück: 1,5 h.

Datenerfassung und -verarbeitung: Meßwerterfassungssystem mit CBM 8032, Floppy-Disk 8250 2x1 MByte, Plotter Adcomp X100S (32 kByte), Printer Epson XM100, Interface Eigenbau, ab 1983/84

Steuerung des Meßablauf: Vollautomatisch über Steuerung des Antriebs.

Anordnung des Prüflings: Horizontal verschiedene Anordnungen möglich. Vertikal ***ABBÉ***sche Anordnung. Abstand Teil.Achse - Interferom.Achse max. 1 mm.

Temperaturerfassung am Prüfling: Akklimatisierungszeit ca. 12 h. Lufttemperatur = Materialtemperatur. Invartemperatur über Thermoelement, jedoch zusätzlich meßbar. Sonderanfertigung der Fa. SIKA.

Genauigkeitsmaße: Reproduzierbarkeit der Stricherfassung 1 - 2 µm (aus Differenzen zwischen Hin- und Rückmessungen). Abhängig von der Erfassung der atmosphärischen Bedingungen. Mittlere Strichgenauigkeit aus drei Meßreihen mit je Hin- und Rückmessung 1 - 2 µm.

Institut für Landesvermessung und Ingenieurgeodäsie
der Technischen Universität Wien

Längennormal: Zweifrequenz He-Ne-Laser, Teil des Laser-Doppler-Interferometers 5526 A der Fa. Hewlett Packard.
Frequenzstabilität lt. Hersteller $5 \cdot 10^{-8}$ bei automatischer Kompensation meteorologischer Einflüsse durch 5526 A Meßsystem.
Labor klimatisiert.

Komparatorbank: T-Träger aus Aluminium zur Lagerung des Prüflings. Meßwagen auf Führungs- und Entlastungswelle aus Stahl mit zylindrischen Kugellagerbuchsen.

Stricherfassung: z.Z. mit Videokamera und Monitor. Vergrößerung 10 : 1, variable Meßmarken am Monitor durch Ausblenden von Zeilen, visuelle Koinzidenz, manuelle Auslösung der Meßwertübernahme. Geplant: kontinuierliche Erfassung des Hell-Dunkelprofils mit Fairchild-Line-scan-Kamera CCD 1100 C, Pixel-Locator und Controller 1320.

Meßdauer: z.Z. 5 h (3m-Latte).

Datenerfassung und -verarbeitung: Commodore PET 2000, eigenes Interface BCD-seriell, Aufzeichnung auf Datenkassette, Verarbeitung auf Commodore CBM 3000. Programmiersprache Basic.

Steuerung des Meßablauf: manuelle Auslösung der Meßwertübernahme, Übermittlung der Meßwerte vom Interferometer zum Tischrechner mit Interface-Eigenbau. Neigungsmessungen der Führungsbahn vor und nach jeder Kalibrierung.

Anordnung des Prüflings: keine ***ABBÉ***sche Anordnung. Korrektur der Meßwerte durch Kontrollmessungen..

Temperaturerfassung am Prüfling: Meßfühler an drei Stellen der Meßstrecke. Automatische Kompensation über hp 5526 A Meßsystem.

Genauigkeitsmaße: Reproduzierbarkeit
Standardabweichung einer Stricherfassung ± 5 µm,
Standardabweichung einer Strichverbesserung derzeit ± 10 µm.

Institut für Geodäsie und Photogrammetrie
der Eidgenössischen Technischen Hochschule Zürich

Längennormal: Zweifrequenz He-Ne-Laser. Meßsystem 5526 A der Fa. Hewlett Packard. Frequenzstabilität $5 \cdot 10^{-7}$.
Manuelle Kompensation der meteorologischen Einflüsse (Temperatur ± 0,1°C, Druck ± 0,4 mbar, Feuchte ± 10 % bei 20°C).
Labor klimatisiert, Temperatur ± 0,5°C.
Auflösung des Längenmeßsystems 0,1 µm.

Komparatorbank: Neunmeter-IPE 400 Träger in den Besselpunkten unterstützt. Maßstabswagen läuft auf geschliffener Rund- und Flachführung. Antrieb des Maßstabswagens manuell.

Stricherfassung: Auflichtprojektor der Fa. ISOMA in Biel.
Vergrößerungen 10x und 20x. Individuelle Strichmaske, Auflichtbeleuchtung durch Hg-Lampe mit zwei Lichtleitern. Manuelle Einstellung des Strichbildes auf Auflichtprojektor durch Verschieben des Maßstabswagens.

Meßdauer: für 3m-Latte (1 cm Doppelteilung, zwei Beobachter) ca. 2,5 h.

Datenerfassung und -verarbeitung: Commodore CBM 8032 mit 32 kByte,
Dual-Drive Floppy Disk CBM 8050, 512 kByte/Disk.
Printer Centronics 770, Programmiersprache Basic.

Steuerung des Meßablauf: manuell im Dialog mit Rechner.

Anordnung des Prüflings: Auflichtprojektor fest, Verschieben der Latte.
Abstand Teilungsachse - Interferometerachse 1 mm.
Oder Latte fest, Verschieben des Mikroskopwagens, Abstand Teilungsachse - Interferometerachse 1 cm.

Temperaturerfassung am Prüfling: zwei Meßfühler Cuproswem mit ± 0,1°C (bei -20 bis +50°C). Ab Sommer 1983.

Genauigkeitsmaße:
Standardabweichung einer Stricherfassung 3 - 5 µm.

LITERATUR

ABBÉ'sches Komparatorprinzip: Zeitschrift für Instrumentenkunde, unter ***"Kleinere Mitteilungen"***, S. 446-447, 1890

BRANDSTÄTTER, G.: ***Theoretisches, meßtechnisches und datentechnisches Konzept eines Interferometerkomparators***. Festschrift Karl Rinner, Mitteilungen des Geodätischen Instituts der Technischen Universität Graz, Folge 40, S. 68-77, 1982

DRAHEIM, H.: ***Das Nivellementnetz 1960***. Allgemeine Vermessungsnachrichten, S. 171-174, 1979

FISCHER, W.: ***Probleme bei der Herstellung von Präzisionsnivellierlatten***. Physikalisch-Technische Bundesanstalt, Bericht Me-25, S. 52-54, Braunschweig 1979

GESSLER, J.: ***Wirtschaftliche Nutzung von Teilstrichverbesserungen durch automatisierte Nivellementsauswertung***. Physikalisch-Technische Bundesanstalt, Bericht Me-25, S. 55-58, Braunschweig 1979

GOTTWALD, R., und WITTE, B.: ***A Microcomputer-Controlled Comparator for Visual and Photoelectric Calibration of Levelling Rods***. In PELZER, H., und NIEMEIER, W. (Herausgeber): ***Precise Levelling***. Dümmler's Verlag, Bonn 1983, S. 127-140

GRABE, M., und LEBOWSKY, F.: ***Prozeßrechnergesteuerte Prüfung von Präzisionsnivellierlatten***. Physikalisch-Technische Bundesanstalt, Bericht Me-25, S. 24-51, Braunschweig 1979

HEISTER, H., CASPARY, W., und TILLE, R.: ***Problems in Reading the Scale Marks on Levelling Rods by Photo-Electric Methods***. In PELZER, H., und NIEMEIER, W. (Herausgeber): ***Precise Levelling***. Dümmler's Verlag, Bonn 1983, S. 141-153

HEWLETT PACKARD COMPANY: ***5501 A Laser Transducer System Operating and Service Manual***. Santa Clara, California 1976

KUNZMANN, H.: ***Anwendung des Laserinterferometers in der Fertigungsmeßtechnik***. Annals of the International Institution for Production Engineering Research (CIRP), S. 311-316, 1979

MAURER, W., und SCHNÄDELBACH, K.: ***First Experiences with a Vertical Comparator for the Calibration of Invar Rods***. In PELZER, H., und NIEMEIER, W. (Herausgeber): ***Precise Levelling***. Dümmler's Verlag, Bonn 1983, S. 155-163

SCHLEMMER, H.: ***Laser-Interferenzkomparator zur Prüfung von Präzisionsnivellierlatten***. Deutsche Geodätische Kommission, Reihe C, Heft Nr. 210, München 1975

SCHLEMMER, H.: ***A new Technique to Produce Precise Graduations on Invar Tape***. In PELZER, H., und NIEMEIER, W. (Herausgeber): ***Precise Levelling***. Dümmler's Verlag, Bonn 1983, S. 119-125

STÜRZER, R.: ***Präzisionsnivellement***. Im Landesbericht der Bundesrepublik Deutschland über die in den Jahren 1979 bis 1983 ausgeführten geodätischen Arbeiten. Deutsche Geodätische Kommission, Reihe B, Heft Nr. 265, München 1983

ZIPPELT, K.: Measurement of Level Rods Temperature and Effects on Precise Levelling. In PELZER, H., und NIEMEIER, W. (Herausgeber): ***Precise Levelling***. Dümmler's Verlag, Bonn 1983, S. 165-177

Schriftenreihe

Wissenschaftlicher Studiengang Vermessungswesen

Hochschule der Bundeswehr München

Bisher erschienene Hefte:

(Die Hefte erscheinen in zwangloser Folge)

Nr. 1/78 A. Schödlbauer :

Curriculum für den wissenschaftlichen Studiengang Vermessungswesen der Hochschule der Bundeswehr München

Nr. 2/78 A. Chrzanowski and E. Dorrer :

Proceedings "Standards and Specifications for Integrated Surveying and Mapping Systems", Workshop held in Munich, 1-2 June 1977

Nr. 3/78 W. Caspary und A. Geiger :

Untersuchungen zur Leistungsfähigkeit elektronischer Neigungsmesser

Nr. 4/79 E. Baumann, W. Caspary, H. Dupraz / W. Niemeier / H. Pelzer, E. Kuntz / G. Schmitt, W. Welsch :

Seminar über Deformationsmessungen

Nr. 5/81 K. Torlegård :

Accuracy Improvement in Close Range Photogrammetry

Nr. 6/82 W. Caspary und W. Welsch :

Beiträge zur großräumigen Neutrassierung

Nr. 7/82 K. Borre and W.M. Welsch :

Proceedings "Survey Control Networks", Meeting of FIG-Study Group 5B, Aalborg, 7 - 9 July 1982

Nr. 8/82 A. Geiger :

Entwicklung und Erprobung eines Präzisionsneigungstisches zur Kalibrierung geodätischer Instrumente

Nr. 9/83 W. Welsch :

Deformationsanalysen '83

...

Schriftenreihe

Wissenschaftlicher Studiengang Vermessungswesen
Hochschule der Bundeswehr München

Nr. 10/84

W. Caspary, A. Schödlbauer und W. Welsch (Hrsg.) :
Beiträge aus dem Institut für Geodäsie