

**Steffen Schön**

**Analyse und Optimierung geodätischer Messanordnungen  
unter besonderer Berücksichtigung des Intervallansatzes**

**München 2003**

---

Verlag der Bayerischen Akademie der Wissenschaften  
in Kommission beim Verlag C. H. Beck

Analyse und Optimierung geodätischer Messanordnungen  
unter besonderer Berücksichtigung des Intervallansatzes

Zur Erlangung des akademischen Grades eines  
Doktor-Ingenieurs (Dr.-Ing.)  
von der Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften  
der Universität Fridericiana zu Karlsruhe (TH)  
genehmigte Dissertation

von

Dipl.-Ing. Steffen Schön

aus Karlsruhe

München 2003

---

Verlag der Bayerischen Akademie der Wissenschaften  
in Kommission beim Verlag C. H. Beck

Adresse der Deutschen Geodätischen Kommission:

**Deutsche Geodätische Kommission**

Marstallplatz 8 • D – 80 539 München

Telefon (089) 23 031 113 • Telefax (089) 23 031 – 283/– 100

E-mail hornik@dgfi.badw.de • <http://www.dgfi.badw.de/dgfi/DGK/dgk.html>

Hauptreferent: Prof. Dr.-Ing. Dr.-Ing.E.h. Günter Schmitt

Korreferenten: Prof. Dr.rer.nat. Götz Alefeld

Priv.-Doz. Dr.-Ing. Hansjörg Kutterer

Tag der mündlichen Prüfung: 17.7.2003

---

© 2003 Deutsche Geodätische Kommission, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,  
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen

## Kurzfassung

In der Geodäsie wird im Allgemeinen der funktionale Zusammenhang zwischen Beobachtung und gesuchten Parametern, beispielsweise Punktkoordinaten geodätischer Netze, durch ein nichtlineares Modell beschrieben. Werden normalverteilte zufällige Messabweichungen vorausgesetzt, so liefert eine Ausgleichung nach der Methode der kleinsten Quadrate beste Schätzwerte für die unbekannt Parameter. Diese Annahmen sind typisch für das Vorgehen in der modernen geodätischen Datenanalyse. Neben zufälliger Variabilität (Stochastizität) der Messwerte kann jedoch auch Impräzision auftreten, die durch unbekannt systematische Abweichungen zwischen Modell und Messwerten hervorgerufen wird. Das Verhältnis der Impräzision zur Stochastizität hängt von den verwendeten Mess- und Auswertemethoden sowie den angebrachten Korrekturen ab.

Im klassischen Ansatz werden beide Anteile der Gesamtunsicherheit (Stochastizität und Impräzision) mit Methoden der Stochastik behandelt. Dieses Vorgehen liefert wegen des quadratischen Varianzfortpflanzungsgesetzes nicht immer realistische Werte für die Unsicherheitsmaße der Ergebnisse, insbesondere wenn die Anzahl der Messungen groß genug gewählt wurde. Es scheint daher sinnvoll, eine konsequente getrennte Behandlung und Beschreibung der beiden qualitativ unterschiedlichen Effekte vorzunehmen. Intervalle eignen sich hierbei besonders für die Beschreibung der Impräzision.

Ein wesentliches Ergebnis der Arbeit ist die Entwicklung von Verfahren, mit denen Intervalle als Impräzisionsmaße für terrestrische geodätische Messungstypen abgeleitet werden. Durch Faustformeln werden die wesentlichen Abhängigkeiten der Intervallradien von den angewandten Mess- und Auswertemethoden, Korrekturenmodellen oder Sensoren in eine besonders praxisgerechte Form gebracht. Für Punktpositionen stellen Zonotope adäquat die Impräzisionsbereiche dar. Eine Kombination mit stochastischen Anteilen in Form von Konfidenzbereichen liefert erweiterte Unsicherheitsmaße, die die Gesamtunsicherheit der Punktposition geeignet wiedergeben. Sie bilden die Grundlage für eine aussagekräftige Interpretation von Messergebnissen.

Als dritter Schwerpunkt wird das Konzept der geodätischen Netzoptimierung um zwei Aspekte erweitert. Zum Einen wird gezeigt, dass sich Kenngrößen der Impräzision für das Zero und First Order Design ähnlich verhalten wie Kenngrößen der Stochastizität. Zum Anderen wird das „Intervalldesign“ eingeführt, um Mess- und Auswertemethoden, Korrekturenmodelle und Sensoren so anzuwenden bzw. auszuwählen, dass die Impräzision der Ergebnisse fest vorgegebene Sollwerte nicht übersteigt. Alle Untersuchungen und Ergebnisse der Arbeit werden an ausgewählten Beispielen erläutert.

## Abstract

In general, a geodetic model describes the nonlinear functional relationship between the observations and some parameters such as the coordinates of geodetic networks. Supposing exclusively normally distributed random measurement errors, a least-squares adjustment yields the best estimate of the unknown parameters. These assumptions and methods are standard in modern geodetic data analysis. But besides random variability (stochasticity), imprecision may occur. It is caused by unknown systematic deviations between the observations and the model which are due to simplifications or imperfect knowledge. Hence, the ratio between random effects and remaining systematics depends mainly on the applied measurement methods, the used correction models and analysis strategies.

In the classical approach, both effects are treated by stochastic means. This yields less realistic uncertainty measures of the results due to the quadratic error propagation which is intrinsic to this approach. Especially, if the number of observations is sufficiently increased, the standard deviations of the results can decrease beyond any limit. These characteristics contradict the experience of surveying practitioners. Hence, it seems reasonable to distinguish strictly between stochasticity and imprecision. Mathematically, imprecision can be treated by intervals.

As a main result of the thesis, different approaches are developed in order to derive intervals, which are adequate measures of imprecision of the geodetic observations. Approximative formulas describe compactly the dependences between the interval radius, sensor performance and network design. As a second topic, zonotopes and zonohedra are used to describe directly and in detail the imprecision of point positions. A combination of zonotopes and point confidence ellipses yields extended uncertainty measures. These are suitable to express adequately the whole uncertainty of the results. Furthermore, these measures are basic for realistic interpretations of the results of modern geodetic data analysis.

A third topic deals with network optimization. The classic concepts are enlarged by two new aspects concerning imprecision. First, it is shown that imprecision reacts similarly to stochasticity during zero and first order design

optimization. Secondly the “interval design“ is introduced. This approach focuses on combining measurements methods, correction models and sensors in such a way that the imprecision of the results does not exceed given upper limits. Numerical and graphical examples are given to emphasize the theoretical results and main ideas of the thesis.

# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b>	<b>5</b>
<b>1 Einleitung</b>	<b>9</b>
<b>2 Beschreibung von Impräzision durch Intervalle</b>	<b>11</b>
2.1 Motivation . . . . .	11
2.1.1 Messmodell und geodätisches Modell . . . . .	11
2.1.2 Modell und Unsicherheit . . . . .	12
2.1.3 Warum Intervallmathematik? . . . . .	15
2.2 Begriffe aus der Intervallmathematik . . . . .	17
2.2.1 Reelle Intervallrechnung . . . . .	17
2.2.2 Reelle Intervallalgebra . . . . .	19
2.2.3 Funktionen mit Intervallen als Variablen . . . . .	21
2.2.4 Intervalle und Fuzzy-Sets . . . . .	24
2.2.5 Intervalle und gleichverteilte Zufallsvariablen . . . . .	25
2.3 Bestimmung der Impräzision korrigierter Beobachtungen . . . . .	27
2.3.1 Vorbemerkungen . . . . .	27
2.3.2 Korrekturen und Messmethoden - Ein formalisierender Ansatz . . . . .	28
2.3.3 Impräzision der Einflussparameter . . . . .	32
2.3.4 Intervallauswertung des Aufbereitungsprozesses der Messung . . . . .	35
2.3.5 Intervalle aus der Sensitivitätsanalyse des Aufbereitungsprozesses der Messung . . . . .	36
2.3.6 Vergleich der Reduktion von Impräzision und Stochastizität durch Messmethoden . . . . .	38
2.3.7 Identifikation von Haupteinflussfaktoren . . . . .	41
<b>3 Beispiele für Intervallradien</b>	<b>44</b>
3.1 Intervallradien als Impräzisionsmaße: Beispiel Totalstation . . . . .	44
3.1.1 Zenitdistanzbeobachtung . . . . .	44
3.1.2 Richtungsbeobachtungen . . . . .	47
3.1.3 Streckenmessung . . . . .	48
3.2 Analyse der Haupteinflussfaktoren . . . . .	49
3.3 Diskussion der Koeffizienten der Faustformeln . . . . .	51
3.3.1 Vorbemerkungen . . . . .	51
3.3.2 Koeffizienten der Faustformel für die Streckenmessung . . . . .	52
3.3.3 Koeffizienten der Intervallradien der Richtungsbeobachtung . . . . .	53
3.3.4 Koeffizienten der Faustformel für die Zenitdistanz . . . . .	56
3.3.5 Minimale Koeffizienten . . . . .	56
3.4 Vergleich der Modellierungsmethoden der Intervallradien . . . . .	56

<b>4</b>	<b>Analyse geodätischer Netze unter Berücksichtigung systematisch wirkender Restunsicherheiten</b>	<b>59</b>
4.1	Parameterschätzung im Gauß-Markov-Modell . . . . .	59
4.1.1	Konzept . . . . .	59
4.1.2	Mathematisches Modell der dreidimensionalen geometrischen Ausgleichung . . . . .	61
4.1.3	Analyse geodätischer Netze nach Genauigkeit und Zuverlässigkeit . . . . .	63
4.2	Fortpflanzung der Impräzision der korrigierten Messungen auf die geschätzten Parameter . . . . .	66
4.2.1	Intervallauswertung des Parametervektors $\hat{\mathbf{x}}$ . . . . .	66
4.2.2	Wertebereich des Parametervektors $\hat{\mathbf{x}}$ . . . . .	67
4.2.3	Wertung . . . . .	69
4.3	Zonotope . . . . .	70
4.3.1	Grundlagen zu Zonotopen . . . . .	70
4.3.2	Wertebereich als Zonotop . . . . .	74
4.3.3	Kenngößen des Wertebereichs . . . . .	78
4.3.4	Geometrische Deutung . . . . .	80
4.3.5	Abschließendes Beispiel: Kenngößen für ein Überwachungsnetz . . . . .	80
4.3.6	Ersatzformen für den Wertebereich . . . . .	82
4.4	Verknüpfung von Impräzision und Stochastizität . . . . .	84
4.4.1	Konzept der Superposition . . . . .	85
4.4.2	Konzept des Erweiterungsprinzips in der Fuzzy-Theorie . . . . .	86
4.4.3	Vergleich einer rein stochastischen mit einer erweiterten Vorgehensweise . . . . .	87
4.5	Alternative Ansätze zur Fortpflanzung der Impräzision . . . . .	89
4.5.1	Konzept der Minimalen Maximalfehler . . . . .	89
4.5.2	Konzept der Set-Inversion . . . . .	91
<b>5</b>	<b>Optimierung geodätischer Messoperationen</b>	<b>93</b>
5.1	Grundlagen der Optimierung . . . . .	93
5.1.1	Skalare Optimierung . . . . .	93
5.1.2	Vektoroptimierung . . . . .	94
5.2	Klassische Geodätische Netzoptimierung - ein Überblick . . . . .	96
5.2.1	Motivation und Einordnung . . . . .	96
5.2.2	Zielfunktionen der klassischen Netzoptimierung . . . . .	97
5.2.3	Feste und freie Parameter in der Optimierung . . . . .	101
5.3	Optimierung der Impräzision: direkte Fragestellung . . . . .	103
5.3.1	Zielfunktionen für die Impräzision . . . . .	103
5.3.2	Reaktion globaler Zielfunktionen auf Variationen der Parameter . . . . .	104
5.3.3	Zusammenfassung und Wertung . . . . .	112
5.4	Optimierung der Impräzision: indirekte Fragestellung . . . . .	115
5.4.1	Idee und Zielsetzung . . . . .	116
5.4.2	Formulierung von Zielvorgaben . . . . .	117
5.4.3	Lösung des Intervalldesigns . . . . .	119

---

5.4.4	Beispiel und Wertung der Lösung des Intervalldesigns . . . . .	123
5.5	Gemeinsame Optimierung von Impräzision und klassischen Zielvorgaben . . . . .	125
5.5.1	Optimierung der erweiterten Unsicherheit . . . . .	125
5.5.2	Gemeinsame Optimierung von Impräzision, Stochastizität und Zuverlässigkeit . . . . .	126
<b>6</b>	<b>Zusammenfassung</b>	<b>131</b>
	<b>Literatur</b>	<b>134</b>





# 1. Einleitung

Eine wesentliche Aufgabe der Geodäsie ist neben der Modellierung und Beschreibung der Erdoberfläche und des Erdschwerefeldes die messtechnische Bestimmung der dazu nötigen Parameter. Diese sind beispielsweise die Koordinaten von Punkten ausgezeichneter Flächen wie der Erdoberfläche oder Ozeanoberfläche zur Beschreibung der geometrischen Figur der Erde oder physikalische Parameter zur Beschreibung des Erdschwerefeldes oder der Erdorientierung. In der vorliegenden Arbeit soll die Diskussion auf Punktkoordinaten geodätischer Überwachungsnetze fokussiert werden.

Die Abfolge der zur Parameterbestimmung nötigen Schritte reicht von der Planung der Messungen über die konkrete Durchführung mit anschließender Datenaufbereitung bis hin zur Berechnung der gesuchten Parameter (hier: Punktkoordinaten geodätischer Netze) und daraus abgeleiteter Größen. In der vorliegenden Arbeit wird davon ausgegangen, dass die Parameter nach der Methode der kleinsten Quadrate, einem Standardansatz in der Geodäsie, geschätzt werden. Neben den konkreten Werten für die geschätzten Parameter sind Maßzahlen von Interesse, mit deren Hilfe die Güte der Schätzwerte beurteilt werden kann. In der Geodäsie werden dafür ausschließlich stochastische Kenngrößen in Form von Unsicherheitsmaßen (Varianzen oder Konfidenzbereichen) verwendet.

Die vorliegende Arbeit soll einen wesentlichen Beitrag liefern, dieses rein stochastisch basierte Konzept entscheidend zu erweitern und zu ergänzen. Ziel ist es, eine adäquate Modellierung der Gesamtunsicherheit der Messung und deren Fortpflanzung auf die geschätzten Parameter zu ermöglichen. Es wird dabei davon ausgegangen, dass sich die Gesamtunsicherheit aus zwei Komponenten zusammensetzt. Neben einer zufälligen Variabilität, die durch stochastische Kenngrößen geeignet beschrieben wird, treten systematisch wirkende Effekte (Impräzision) auf. Letztere werden durch Diskrepanzen zwischen Modell und Daten hervorgerufen, die in Vernachlässigungen oder unvollständigem Wissen um die Modellierung begründet sein können. Systematisch wirkende Effekte weisen bei der Betrachtung einer konkreten Anwendung keinen zufälligen Charakter auf, so dass sie durch Zufallsgrößen nur wenig geeignet darstellbar sind. Es wird gezeigt, dass sich Intervalle hingegen sehr gut eignen, Impräzision zu charakterisieren und zu beschreiben. Eine lineare Fortpflanzung der Intervalle erlaubt, die Impräzision der Beobachtungen auf die geschätzten Parameter zu übertragen. Die Kombination beider Anteile (Stochastizität und Impräzision) unter Beibehaltung ihrer spezifischen Eigenschaften liefert ein realistisches Maß für die Gesamtunsicherheit der Schätzwerte. Dies ist für eine aussagekräftige Beurteilung der geschätzten Punktkoordinaten und daraus abgeleiteter Größen entscheidend.

Die vorliegende Arbeit gliedert sich in drei Schwerpunkte: Messung und Datenaufbereitung - Schätzung der Parameter - Optimierung und Entwurf geodätischer Netze und Versuchsanordnungen. Der letzte Punkt schließt dabei den Kreis zum ersten Themenkomplex „Messung und Datenaufbereitung“. In einem ersten Kapitel wird die Verwendung von Intervallen als Maße für Impräzision motiviert. Ausgangspunkt bildet ein Überblick und eine Wertung der Modellierung in der Geodäsie, insbesondere des dabei verwendeten stochastischen Ansatzes zur Beschreibung von Unsicherheit. Als Ergänzung und Erweiterung dieses Konzeptes werden Intervalle als Maße für Impräzision begründet. Die Grundlagen der Intervallmathematik werden vorgestellt, die eine formale Betrachtung und das Rechnen mit Impräzision erlauben. In einem letzten Abschnitt wird die geodätische Datenanalyse genau beleuchtet. Ansatzpunkt für die Überlegungen ist, dass jede Verarbeitung der originären Messungen bzw. Daten auch deren Unsicherheit beeinflusst. Es wird aufgezeigt, wie konkrete Werte für die Impräzision gewonnen werden, welche Größenordnung diese annehmen und wie sie durch den Aufbereitungsprozess der originären Messung auf die korrigierte Messung fortgepflanzt werden können. Die vorgestellten Konzepte werden dabei an konkreten Beispielen erläutert und diskutiert, um so den Bogen von der Theorie zur Praxis zu spannen.

Das dritte Kapitel, das eng mit dem zweiten verknüpft ist, veranschaulicht an ausgewählten Beispielen das Konzept der Intervalle als Impräzisionsmaße für geodätische Messungstypen, wie sie in geodätischen Überwachungsnetzen vorkommen. Ergänzend werden Faustformeln für die Impräzision terrestrischer Messungen wie Richtungen, Strecken und Zenitdistanzen angegeben. Diese beschreiben die wesentlichen Charakteristika und erlauben so einen schnellen Überblick über das Verhalten der Impräzision für die ausgewählten Messungstypen.

Eine zweite Einheit bildet die Fortpflanzung der Gesamtunsicherheit (Stochastizität und Impräzision) der korrigierten Beobachtungen auf die geschätzten Parameter. Hierzu wird die Intervallauswertung des Kleinsten-Quadrate-Schätzers angewendet, die von KUTTERER (1994) vorgestellt wurde. Die Diskussion wird im Hinblick auf die Schätzung von Punktkoordinaten in geodätischen Netzen geführt. Hierfür wird das geometrische Modell der dreidimensionalen Netzausgleichung skizziert und auf die Besonderheiten der Netzausgleichung als Spezialfall der Parameterschätzung eingegangen. Das Konzept der Intervallbox als Impräzisionsbereich für Koordinaten wird durch das Konzept der Zonotope verfeinert und verallgemeinert. Dadurch können detaillierte Aussagen

über die Form des Impräzisionsbereichs für Punktpositionen abgeleitet werden, was anhand verschiedener Beispiele veranschaulicht wird. Schließlich werden durch eine Kombination von Stochastizität und Impräzision erweiterte Konfidenzbereiche für Punktpositionen entwickelt, die deren Unsicherheit realistischer widerspiegeln.

Einen dritten Themenbereich bildet das Design und die Optimierung geodätischer Netze, dem das fünfte Kapitel gewidmet ist. Die Untersuchungen schließen den Kreis zum Themenkomplex „Messung und Datenaufbereitung“. Der Ingenieur und Wissenschaftler muss in der Regel bei der Lösung seiner Aufgabenstellungen verschiedene, gegebenenfalls sich widersprechende Zielvorgaben optimal erfüllen. Im Kontext der vorliegenden Arbeit soll die klassische Vorgehensweise der Geodäsie zur Lösung dieser Fragestellung um zwei Aspekte erweitert und ergänzt werden. Zum Einen kommt die Impräzision als neue Zielgröße zu den klassischen Größen wie Genauigkeit, Zuverlässigkeit oder Wirtschaftlichkeit hinzu, die für das gegebene Netz optimiert werden soll. Zum Anderen hängt die Gesamtunsicherheit der Ergebnisse entscheidend von der Messanordnung und Auswertestrategie ab. Die Diskussion - auch für die klassischen Zielgrößen - muss somit breiter angelegt werden, was am Beispiel von Überwachungsnetzen der Ingenieurgeodäsie gezeigt wird.

## 2. Beschreibung von Impräzision durch Intervalle

### 2.1 Motivation

Das Konzept der Beschreibung von Impräzision durch Intervalle ist als einer der „Bausteine für eine neue Fehlertheorie“ anzusehen, wie sie von KUTTERER (2002b) angeregt und vorgestellt wurde. Dieser Abschnitt soll den Leser für die speziellen Problem- und Ausgangsfragestellungen sensibilisieren, die die Notwendigkeit einer „neuen Fehlertheorie“ als Lösung bzw. Antwort nahelegen.

In einem ersten Schritt wird Grundsätzliches zur Modellierung angemerkt, die den theoretischen Hintergrund für die konkrete Durchführung und Analyse von Messungen sowie für die Ergebnisinterpretation bildet. Bereits bei der Modellierung, spätestens aber bei der Durchführung von Messungen, ist man mit „Fehlern“ konfrontiert. In einem zweiten Schritt wird die dabei angewandte stochastische Modellierung skizziert. Sie dominiert den Umgang mit „Fehlern“ in der Geodäsie. Es wird herausgearbeitet, welches die dabei explizit oder implizit angenommenen Arbeitshypothesen sind und welche Interpretationsmöglichkeiten bzw. welche Einschränkungen sie mit sich bringen. In einem letzten Schritt wird ein intervallmathematisches Konzept motiviert, das Defizite einer rein stochastischen Modellierung behebt und diese entscheidend ergänzt.

#### 2.1.1 Messmodell und geodätisches Modell

Ziel der Geodäsie ist letztlich, Aussagen über die Geometrie der Erdoberfläche als Ganzes oder in Teilen zu treffen und die Struktur des Schwerefeldes im Außenraum abzuleiten. Da die a priori unbekannt Parameter, die zur Beschreibung der Geometrie und des Schwerefeldes benötigt werden, in der Regel nicht direkt gemessen werden können, müssen diese indirekt aus Beobachtungen bestimmt werden. Dazu werden die Vorgänge und Zustände der realen Welt, für die man sich interessiert, in *geodätischen Modellen* approximiert und so parametrisiert, dass sie Messungen, für welche Sensoren zur Verfügung stehen, funktional beschreiben. In der geodätischen Praxis werden redundante Messungen durchgeführt, und die unbekannt Parameter des geodätischen Modells werden nach der Methode der kleinsten Quadrate geschätzt. Die als bekannt vorausgesetzten physikalischen Zusammenhänge, die den Messvorgang an sich darstellen, sind im *Messmodell* gegeben.

Der Modellbegriff wird hier bewusst eingeführt, um darauf hinzuweisen, dass implizit mit jeder Beschreibung in einem Modell die Realität nur approximativ dargestellt werden kann und verschiedene Modellvoraussetzungen oder Arbeitshypothesen akzeptiert werden. Die folgenden Überlegungen sollen einige Zielkonflikte andeuten, die sich bei der Modellierung ergeben können. Ein Modell sei charakterisiert durch die Anzahl der Parameter, die Komplexität ihrer Verknüpfung zur Beschreibung des Phänomens und die Voraussetzungen, die angeben, unter welchen Bedingungen das Modell gültig ist. Soll beispielsweise ein Modell mathematisch einfach handhabbar sein, so wird man zweckmäßig wenige Parameter einführen und die Komplexität der Operationen zwischen Parametern beschränken. Dies wird dazu führen, dass nicht mehr die Gesamtheit eines Phänomens modelliert werden kann, sondern nur noch Teilaspekte. Um dies zu verdeutlichen, sollten die nichterklärbaren Zustände explizit ausgeschlossen werden. Dies kann beispielsweise durch eine entsprechend schärfere Formulierung der Modellvoraussetzungen realisiert werden. HECK (1995, S.86) nennt z.B. die Grundlagen der Newtonschen Mechanik als Modellvoraussetzungen, die den Geltungsbereich der verwendeten Modelle beschränken, da relativistische Effekte oder Refraktion nicht erklärt sind. Weitere Voraussetzungen sind beispielsweise „Normalatmosphäre“, „Laborbedingungen“ oder „reibungsfreie Betrachtung“ in der Mechanik.

Ein offensichtlicher aber wesentlicher Aspekt der Wahl der Modelle ist, dass sie für die jeweilige Aufgabenstellung (Forschungszweck oder Anforderungen von Auftraggebern) die reale Situation *hinreichend* und *geeignet* beschreiben sollen. Diese Forderungen implizieren, wie oben bereits angedeutet, einen mehr oder minder großen Konflikt zwischen einer möglichst umfassenden, oft sehr komplexen Beschreibung auf der einen Seite und einer vereinfachenden, aber mathematisch gut handhabbaren Beschreibung wie dem Arbeiten in linearen Modellen oder mit gewöhnlichen Differentialgleichungen auf der anderen Seite. Auf die Frage, ob die erweiterten komplexen Modelle die realen Zusammenhänge *richtiger* oder auf der gleichen Approximationsebene der Realität *nur anders* darstellen, soll in dieser Arbeit nicht weiter eingegangen werden.

Abbildung 2.1 zeigt die Zusammenhänge zwischen der realen Welt mit den sich darin abspielenden Vorgängen und Zuständen sowie deren Abbildung in Modelle. Mit aufsteigender Ordnungszahl nimmt die Komplexität der Modelle und in der Regel die Parameteranzahl zu, wohingegen die Schärfe der Modellvoraussetzungen abnimmt, bis schließlich ein mehr oder weniger allgemeingültiges Modell erhalten wird. Ein vereinfachtes Modell ist somit ein Spezialfall eines umfassenden Modells.

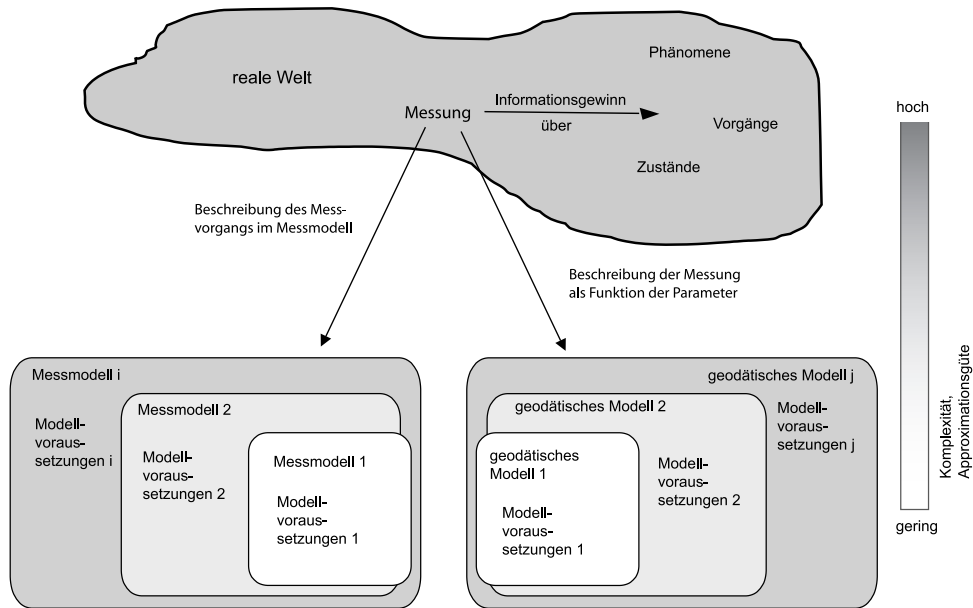


Abbildung 2.1: Zusammenhänge zwischen den Modellen

Es ist wichtig, die Modellvoraussetzungen explizit darzustellen, da nur auf dieser Grundlage eine realistische Beurteilung und ein Vergleich von Modellen möglich ist. Ebenso sind erst so die berechneten Ergebnisse und die gezogenen Schlussfolgerungen nachvollziehbar. In vielen technischen Anwendungen sind die Modellvoraussetzungen jedoch nur implizit gegeben, z.B. durch Arbeitshypothesen oder Postulate, die in einer Fachrichtung allgemein anerkannt sind. Dies lässt für Außenstehende verschiedene Spekulationen zu und erschwert es, die Berechnung von Ergebnissen nachvollziehen zu können. Es muss aber eingeräumt werden, dass eine explizite Darstellung der Modellvoraussetzungen verlangt, dass die modellierten Phänomene in ihrer ganzen Tiefe kognitiv erfasst wurden. Inwieweit dies heute bei sehr komplexen Problemstellungen durch den Einzelnen überhaupt möglich ist, sei dahingestellt.

Soll mit einem einfachen und somit gut überschaubaren geodätischen Modell gearbeitet werden, dann sind *Korrekturen* an den Messungen anzubringen. Diese überführen die Voraussetzungen des Messmodells auf die Modellvoraussetzungen des zur Lösung der geodätischen Aufgabenstellung gewählten mathematischen Ausgleichungsmodells. Im Detail müssen einerseits die Modellvoraussetzungen des Messmodells (*Beobachtungsrahmen*) im Vergleich zu den realen Bedingungen berücksichtigt werden. Hierunter ist z.B. die Betrachtung der atmosphärischen Bedingungen zum Zeitpunkt der Messung im Vergleich zu den Modellbedingungen wie z.B. Normalatmosphäre in einem Labor zu verstehen oder Korrekturen wegen relativistischer Effekte, um im Rahmen der Newtonschen Mechanik zu arbeiten. Andererseits führt der Wunsch nach mathematisch einfach handhabbaren geodätischen Modellen zu vereinfachenden Modellannahmen, die ebenfalls korrigiert werden müssen. Der Übergang zwischen Messung und geodätischem Modell ist somit fließend (FANJOR 1985) und bildet die Schnittstelle zwischen der Datenaufbereitung (dem Anbringen von Korrekturen im Kontext des Messmodells) und der Datenanalyse (der Parameterschätzung und Interpretation im Kontext des geodätischen Modells), vgl. KUTTERER (2002b, S.11) und Abbildung 2.2. In der vorliegenden Arbeit soll als Messung primär der Sensoroutput verstanden werden, auf den verschiedene Korrekturen angewandt werden. Als Ergebnis erhält man die *korrigierte Messung*, die die Eingangsgröße des Ausgleichungsmodells ist.

### 2.1.2 Modell und Unsicherheit

Die Diskrepanz zwischen Modell und Realität als auch die Messung an sich induzieren *Unsicherheit*, vgl. BANDEMER (1997) bzw. KUTTERER (2002b, S.16ff). Ziel ist es, die im geodätischen Modell geschätzten Parameter sinnvoll beurteilen zu können. Hierzu muss zunächst der Unsicherheitshaushalt der korrigierten Messung geeignet beschrieben werden. Geeignet bedeutet, dass qualitativ unterschiedliche Beiträge zum Unsicherheitshaushalt, die spezifische Charakteristika aufweisen, entsprechend unterschiedlich behandelt werden müssen. In einer weiteren Betrachtung muss diese Unsicherheit auf die geschätzten Parameter fortgepflanzt werden. Dieser Schritt wird in Kapitel 4 vollzogen.

Im Folgenden wird als „Bestandsaufnahme“ das klassische Vorgehen in der Geodäsie skizziert. Eine Übersicht hierzu wird in KUTTERER (2002b, S.16ff) gegeben. In der klassischen geodätischen Literatur findet man eine Unterteilung der Unsicherheit in drei Fehlertypen: grobe, systematische und zufällige Fehler. Im Kontext des geodätischen Modells wird das *stochastische Modell* eingeführt, das die Unsicherheit der korrigierten Beobachtungen, der Eingangsgrößen der Ausgleichung, beschreibt. Als eine allgemein akzeptierte Arbeitsgrundlage wird hierzu angenommen, dass zufällige Fehler normalverteilt sind. Auf dieser Basis können grobe Fehler durch Testverfahren wie Datasnooping (BAARDA 1968) oder robuste Schätzverfahren vor der eigentlichen Ausgleichung eliminiert werden, vgl. allgemein HUBER (1981) oder HAMPEL ET AL. (1986) und für geodätische Anwendungen beispielsweise WICKI (1998). Systematische Fehler sollen durch geeignete Messverfahren oder das Anbringen von Korrekturen reduziert bzw. eliminiert werden. KUTTERER (2002b, S.17) weist darauf hin, dass in moderner Literatur zur Ausgleichungsrechnung systematische Fehler nicht mehr explizit behandelt werden und dass es wegen der Rückläufigkeit der Bedeutung des Begriffs „Fehler“ passender scheint, den Begriff *Abweichung* zu verwenden. In der vorliegenden Arbeit werden abweichend dazu die Fachbegriffe wie Kippachs- oder Zielachsfehler weiter verwendet.

Paradoxerweise können systematische und grobe Fehler im klassischen Konzept praktisch nicht losgelöst von zufälligen Fehlern betrachtet werden, obwohl sie von ihrer Charakteristik her als unterschiedlich eingestuft werden. Zufällige Fehler bilden letztlich die Referenz bzw. den Maßstab, der entscheidet, ob grobe bzw. systematische Fehler vorliegen. Bei groben Fehlern ist der Zusammenhang direkt aus den oben genannten Testverfahren ersichtlich. Bei systematischen Effekten gilt: Setzt man eine gewisse Größenordnung des zufälligen Fehlers voraus, so liegt ein systematischer Fehler erst dann vor, wenn er von den zufälligen Fehlern separierbar ist, d.h. seine Größe signifikant nachgewiesen werden kann. In diesem Fall sind die Messmethoden oder die Korrekturen nicht *hinreichend*, um den Zusammenhang vollständig zu beschreiben. Eine Verbesserung der Messmethoden, verbesserte Korrekturen oder ein erweitertes geodätisches Modell sind mögliche Antworten. Die Größenordnung der zufälligen Fehler steuert somit, welche Effekte kompensiert werden müssen. Zusammenfassend lässt sich feststellen, dass die klassische Betrachtungsweise von Unsicherheit in der Geodäsie auf zufällige Fehler fokussiert ist.

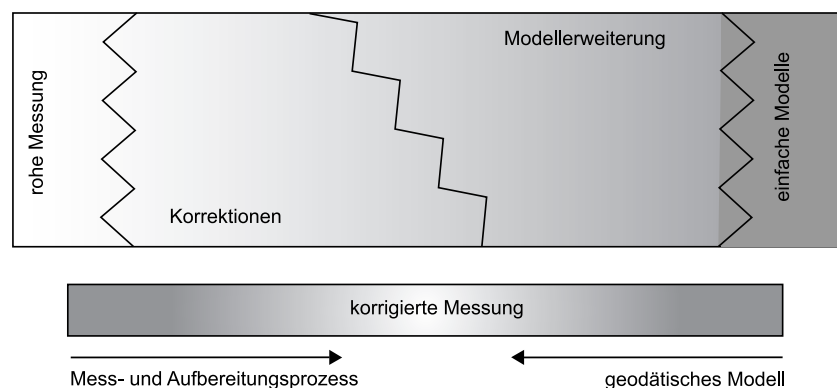


Abbildung 2.2: Wechselspiel zwischen Modell und Korrekturen

Wie die Effekte kompensiert werden, ist dem Anwender freigestellt und drückt sich im Wechselspiel zwischen geodätischem Modell und den Korrekturen aus, vgl. Abbildung 2.2. So kann ein Effekt entweder durch das Anbringen weiterer Korrekturen an die Beobachtungen beseitigt werden. Äquivalent dazu kann das geodätische Modell durch Einführung zusätzlicher Parameter erweitert werden, mit denen der Effekt geschätzt werden soll. Betrachten wir beispielsweise die Berücksichtigung des Einflusses der Troposphäre auf GPS-Messungen. Hier wird ein Anteil des Effektes durch Korrektur der Messungen eliminiert (beispielsweise durch das Modell nach SAASTAMOINEN (1973)). Da diese Korrektur den Gesamteffekt nur grob beschreibt, wird versucht, durch die Schätzung einer restlichen Laufzeitverzögerung (Zenitdelay) den Gesamteffekt besser zu erfassen. In der Regel gilt: Wird ein Effekt als Störgröße aufgefasst, so wird man ihn durch das Messmodell korrigieren. Andererseits kann diese Größe als Signal von Interesse sein, so dass sie im geodätischen Modell modelliert wird.

Neben den „Fehlerarten“ soll nun schlaglichtartig beleuchtet werden, an welchen Stellen des Aufbereitungs- bzw. Auswerteprozesses die klassische „Fehlerbetrachtung“ einsetzt und wie die Fehler fortgepflanzt werden. Hierdurch kann abgeschätzt werden, ob durch das Vorgehen das gesamte Unsicherheitsbudget abgedeckt ist oder ob Effekte nicht in die Rechnung mit einfließen. Im Kontext der elektronischen Distanzmessung beispielsweise stellen JOECKEL UND STOBBER (1999, S.99ff) verschiedene Fehler vor. Es wird das Auflösungsvermögen der

Phasenmessung, Phaseninhomogenitäten oder die Konstanz der Nullpunktskorrektur als von der Streckenlänge unabhängige Anteile genannt. Streckenlängenabhängige Anteile werden hauptsächlich durch Abweichung der Modulationsfrequenz und des Brechungsindex von Standardwerten hervorgerufen. Die Gesamtunsicherheit wird in der folgenden Form dargestellt, wobei sich die Koeffizienten als Summe der Einzelvarianzen der angegebenen Effekte berechnen.

$$\sigma_S^2 = a^2 + b^2 S^2 \quad (2.1)$$

RÜEGER (1996, S.184) unterscheidet explizit zufällige und systematische Fehler. Die systematischen Fehler werden dabei nach ihrer Abhängigkeit von der Streckenlänge in vier Klassen eingeteilt. Er nennt weitere instrumentelle Effekte, die bei JOECKEL UND STOBER (1999) nicht auftauchen, insbesondere Fehlereinflüsse, die nicht vom Instrument abhängen, wie atmosphärische Effekte oder Einflüsse der Zentrierung und Horizontierung sowie der verwendeten Reflektoren. Weiter weist RÜEGER (1996, S.220f) darauf hin, dass in den Herstellerangaben für die Koeffizienten  $a$  und  $b$  sowohl zufällige, „quasi-zufällige“ (kurzperiodische) als auch systematische Effekte enthalten sind. Daher sei das Modell (2.1) als a priori Modell für die Ausgleichung weniger geeignet. PARKINSON (1996, S.469-484) betrachtet den Fehlerhaushalt von GPS-Phasenmessungen und unterscheidet explizit zufällige (random) und systematische Anteile (bias), vergleiche hierzu beispielsweise PARKINSON (1996, Tab.2-4, S.482ff.). Trotz dieser Differenzierung werden beide Anteile als Varianzen aufgefasst und zur Gesamtvarianz addiert.

Diese wenigen Beispiele zeigen bereits zwei typische Merkmale der klassischen Unsicherheitsbetrachtung: Zum Einen werden je nach Anwendung unterschiedliche Effekte berücksichtigt oder vernachlässigt, wie der Vergleich der beiden Lehrbücher von RÜEGER (1996) und JOECKEL UND STOBER (1999) zeigt. Prinzipiell scheint es also offen und fallspezifisch zu sein, an welcher Stelle im Ablauf der Aufbereitung der Messung und für welche Parameter bzw. Korrekturen Unsicherheit betrachtet wird. Um in solchen Situationen Klarheit zu bekommen, ist es sinnvoll, den Forderungen des *Guide to the Expression of Uncertainty in Measurements (GUM)* (DEUTSCHES INSTITUT FÜR NORMUNG 1995, S.35ff) nachzukommen und alle Einflussfaktoren für die Unsicherheit zu protokollieren. Ähnlich argumentiert auch KUTTERER (2002b, S.60ff). Er schlägt vor, Fachleute der einzelnen Spezialdisziplinen sollten in einem Fragebogenkonzept alle Arten der Unsicherheit charakterisieren und zusammenstellen.

Zum Anderen werden, auch wenn qualitativ unterschiedliche Effekte wie zufällige und systematische Fehler explizit identifiziert werden, diese nicht unterschiedlich behandelt, vgl. PARKINSON (1996). Vielmehr werden beide Fehlertypen als Varianzen aufgefasst. Dieses Vorgehen entspricht den Vorschlägen des oben genannten Leitfadens (DEUTSCHES INSTITUT FÜR NORMUNG 1995). Als Konsequenz ergeben sich zu optimistische Maße für die Gesamtunsicherheit.

Die klassische Vorgehensweise ist in zweierlei Hinsicht kritisch zu bewerten. Einerseits sind, wie oben ausgeführt, Standardabweichungen bzw. Varianzen Steuerungsfaktoren für beispielsweise die Aufdeckbarkeit grober Fehler. Da eine rein stochastische Betrachtung verfolgt wird, können bestimmte Effekte vernachlässigt werden, weil sie in diesem Konzept keinen Einfluss auf die Standardabweichung des Ergebnisses haben. Dadurch wird die aktuelle Betrachtung übersichtlicher und mathematisch einfacher handhabbar. Ist jedoch die Varianz zu optimistisch angegeben, so werden relevante Beiträge zum Unsicherheitsbudget vernachlässigt.

Zu optimistisch abgegebene Unsicherheitsmaße führen andererseits in gewisser Weise zu zu optimistischen Aussagen. Beispielsweise kann bei statistischen Tests durch eine zu optimistische Varianz eine Aussage als signifikant bestätigt werden, wohingegen bei einer realistischen Angabe kein Nachweis erbracht werden kann. Trotz oder gerade wegen dieses Sachverhaltes scheint es sinnvoll, ein Konzept zu entwickeln, das den gesamten Unsicherheitshaushalt des Mess- und Auswerteprozesses mit einbezieht, denn ein solches Konzept ist als Fundament für aussagekräftige Analysen und Schlussfolgerungen unumgänglich.

Nach diesen Betrachtungen des klassischen Umgangs mit Unsicherheit sollen deshalb in einem weiteren Punkt einige Möglichkeiten aufgezeigt werden, das rein stochastische Konzept zu erweitern. Dieses soll so ergänzt werden, dass Beiträge zum Unsicherheitshaushalt mit unterschiedlichen Charakteristika auch durchgängig unterschiedlich beschrieben werden. Betrachten wir nochmals die Aufbereitung geodätischer Messungen. Es ist leicht einzusehen, dass jede Verarbeitung der originären Beobachtung auch ihre Unsicherheit beeinflusst. Einerseits sind für die einzelnen Korrekturen zusätzliche Informationen nötig, wie z.B. Modellkonstanten, die im Messmodell bereitgestellt werden, oder zusätzliche Messungen wie Temperatur oder Druck. Diese Informationen sind jedoch unsicher, so dass sich das Unsicherheitsbudget der korrigierten Messungen erhöht. Andererseits können Messmethoden (wie Messungen in zwei Lagen bei der Richtungsbeobachtung oder die Bildung von Linearkombination bei GPS-Messungen) einer starken Zunahme der Unsicherheit entgegenwirken, da sie verschiedene systematische Effekte auf Resteffekte reduzieren bzw. eliminieren.

Wie können alle diese Beiträge zum Unsicherheitsbudget adäquat beschrieben werden und welche Effekte sind vernachlässigbar? Um diese Fragen umfassend beantworten zu können, muss das gesamte Unsicherheitsbudget erarbeitet und analysiert werden. Hierzu ist der Mess- und Aufbereitungsprozess zu untersuchen, um einzelne Beiträge, ihre Charakteristika und ihr gegenseitiges Zusammenwirken zu identifizieren. Dies kann beispielsweise mit Hilfe von Fragebögen, vgl. KUTTERER (2002b, S. 60ff), realisiert werden, wobei für einzelne Teilaspekte Spezialisten der jeweiligen Disziplin zu befragen sind. Als konkrete Beispiele sind die Ansätze der IAG Spezial Studiengruppe 4.190 „Non-probabilistic assessment in geodetic data analysis“ zu nennen bzw. für die EDM RÜEGER (1996, S.184f) oder für das Nivellement RÜEGER UND BRUNNER (2000).

Es muss jedoch eingeräumt werden, dass einer vollständigen Erfassung und konsequenten Behandlung der Unsicherheit oftmals Beschränkungen bei der Realisierung der Messung entgegenstehen. Durch begrenzte technische, personelle oder finanzielle Möglichkeiten können nicht alle Zusatzinformationen erfasst oder Messungen ausgeführt werden, die als Eingangsparameter für die Berechnung der Korrektur nötig sind. In solchen Fällen muss auf Standardwerte und -modelle zurückgegriffen werden. In anderen Fällen liegt kein praxisrelevantes Modell für Korrekturen vor oder die verwendete Software ist nicht flexibel genug, um andere Korrekturen bzw. eine Abschätzung der Unsicherheit zuzulassen.

Neben den oben aufgeführten Vernachlässigungen sind noch andere Schranken gesetzt: Moderne Sensoren sind sehr komplex und meist mit geräteinterner Auswertesoftware gekoppelt, so dass das Gesamtsystem für den Anwender zu einer Black Box wird. Es ist somit nicht mehr zu durchschauen, welche Verarbeitungsschritte mit der ursprünglichen Messung schon durchgeführt wurden und wie diese die Unsicherheit des Sensoroutputs beeinflussen. Durch Herausgabe neuer Softwareversionen kann sich die Konfiguration des Sensorsystems in kurzen Zeitabständen verändern. Zusätzlich fehlen beispielsweise für moderne Tachymeter durchgreifende Untersuchungsmethoden, die verlässliche Aussagen zur Unsicherheit der angezeigten Messwerte geben. Hier sind die Hochschulinstitute aufgerufen, neue Verfahren zu entwickeln, um Erfahrungswerte für eine realistische Unsicherheitseinschätzung, wie sie zum Handwerkszeug eines Ingenieurs gehören, zur Verfügung zu stellen.

### 2.1.3 Warum Intervallmathematik ?

Ziel ist es, ein Konzept vorzuschlagen, das den gesamten Unsicherheitshaushalt adäquat darstellt. In dieser Arbeit wird vorausgesetzt, dass sich der Unsicherheitshaushalt mittels Stochastizität und Impräzision beschreiben lässt, vgl. KUTTERER (2002b). *Stochastizität* drückt die zufällige Variabilität einer Größe aus. *Impräzision* wird durch unbekannt systematische Abweichungen zwischen Beobachtungen und Modell hervorgerufen, die durch Vereinfachungen oder unvollständiges Wissen begründet sind. Das Verhältnis beider Anteile hängt maßgeblich von der jeweiligen Problemstellung und ihrer Modellierung sowie von den verwendeten Mess- und Auswertemethoden ab. Intervalle bilden gewissermaßen ein natürliches Konzept, um Impräzision zu beschreiben, da durch eine bekannte obere bzw. untere Intervallgrenze der Bereich angegeben wird, in dem ein Wert maximal variieren kann. Diese Variationsmöglichkeit drückt gerade die Impräzision, das fehlende Wissen über den konkreten Wert, aus.

Die Verwendung von Intervallen lässt sich einmal aus der Kritik an einer rein stochastischen Vorgehensweise im Konzept der Parameterschätzung nach der Methode der kleinsten Quadrate motivieren. Zum Anderen sind neue Fragestellungen von Interesse, wie die Bestimmung *garantierter Parameterbereiche*, auf die im klassischen Konzept direkt keine Antwort gegeben werden kann. Im Folgenden werden einige wichtige Ansatzpunkte kurz vorgestellt, vgl. hierzu auch KREINOVICH (1995).

**Stochastische Eigenschaften wenig bekannt:** Die stochastischen Eigenschaften von Messungen können a priori unbekannt sein. Dies gilt insbesondere für sogenannte zusätzliche Messungen wie Temperatur oder Druck, die für die Berechnung der Korrekturen der geodätischen Beobachtungen benötigt werden. Für sie ist der Stichprobenumfang meist sehr gering (vielleicht drei bis fünf Ablesungen), so dass die unbekannt Verteilung nicht durch Anpassungstests bestimmt werden kann, vgl. KREYSZIG (1979, Kap.15) oder KOCH (1997, S.295f). In der stochastischen Betrachtung bilden die zentralen Momente der zugrunde gelegten Verteilung (Mittelwert und Varianz) die wesentlichen Kenngrößen, mit deren Hilfe argumentiert wird. Die klassische Arbeitshypothese, normalverteilte Zufallsvariablen anzunehmen, scheint aufgrund des geringen Stichprobenumfangs ebenfalls wenig sinnvoll, da die Aussagen des zentralen Grenzwertsatzes (KREYSZIG 1979, S.200ff) nicht greifen. Zusätzlich zielen Mittelwerte als Kenngrößen auf die Betrachtung des durchschnittlichen Ereignisses, das sich erst bei Wiederholungen durchsetzt.

Sind die stochastischen Eigenschaften unbekannt, so kann die maximale Variabilität der Messwerte angegeben werden, die sich durch Intervalle beschreiben lassen, vgl. beispielsweise REMBE ET AL. (1999)



oder KIEFFER ET AL. (2000). Für viele Parameter ist die maximale Variabilität beispielsweise durch physikalische Gesetzmäßigkeiten oder technische Eigenschaften des Sensors beschränkt oder kann aus dem Erfahrungsschatz einfach abgeschätzt werden.

**Garantierte Parameter:** Bei vielen Anwendungen sind nicht durchschnittliche, sondern die maximal möglichen Werte von Interesse. So müssen bei sicherheitsrelevanten Konfigurationen vor allem die Extremfälle (*worst case*) untersucht und abgedeckt werden. Ebenso ist es sinnvoll, teure Versuchsanlagen so zu dimensionieren bzw. solche Sensoren und Messanordnungen zu verwenden, dass *alle* zu erwartenden Ereignisse auch wirklich untersucht werden können (BRAEMS 2002). Man spricht in diesem Zusammenhang von einer garantierten Parameterbestimmung und unterscheidet zwischen einer direkten Aufgabenstellung, der eigentlichen Parameterbestimmung und einer indirekten oder inversen Aufgabenstellung, bei der Sensoren und Messanordnung so ausgewählt werden sollen, dass die zu bestimmenden Parameter innerhalb vorgegebener Bereiche liegen. Technische Anwendungen mit Intervallen findet man beispielsweise bei COLDITZ (1997), REMBE ET AL. (1999) oder BRAEMS ET AL. (2003).

**Repräsentativität und Toleranz:** Neben dem eigentlichen Wert, der für die Rechnung verwendet wurde, interessiert oftmals auch seine Repräsentativität. Diese Information kann dem Wert in Form eines Intervalls mitgegeben werden. Intervalle können so als Bereiche aufgefasst werden, innerhalb denen typische Werte für Parameter liegen können. Betrachten wir als Beispiel den mittleren Refraktionskoeffizienten, der standardmäßig mit  $\kappa = 0.13$  angegeben wird. Dieser Wert ist jedoch sehr starken Schwankungen unterworfen, vgl. KAHMEN (1993, S.467). Durch das Hinzufügen der Variationsbreite für den Messungszeitraum in Form eines Intervalls  $[\underline{\kappa}, \overline{\kappa}]$  wird die Repräsentativität des Wertes mitgeliefert, die bei der Interpretation der Ergebnisse mit herangezogen werden sollte. Umgekehrt können die Intervallgrenzen auch als Toleranzen angesehen werden und somit feste Schranken beschreiben, innerhalb deren nur Werte akzeptiert werden.

**Rundungsfehler:** NUDING (1975) weist auf die semantische Problematik des Begriffs *reelle Zahlen* hin, der seiner Meinung nach zu viel Realität und Glaubwürdigkeit vortäusche. Zahlenwerte können im Rechner, auf Datenträgern oder in modernen geodätischen Messinstrumenten nur mit einer begrenzten Stellenanzahl gespeichert und verarbeitet werden. Rundungsfehler von  $\pm$  der Einheit der letzten Nachkommastelle sind somit unausweichlich. Eine Wurzel der Intervallmathematik liegt in der Aufgabenstellung, Rundungsfehler auf Rechenanlagen zu quantifizieren, vgl. KULISCH (1976), KULISCH UND MIRKANKER (1983) oder KULISCH (1989). In der vorliegenden Arbeiten soll das Thema Rundungsfehler und Numerik in geodätischen Rechnungen nicht im Vordergrund stehen. Ausführliche Arbeiten liegen zu diesem Thema beispielsweise von SCHMITT (1973) oder DINTER (2002) vor.

**Alternative Fehlerfortpflanzung:** Ein Kritikpunkt an der gängigen Fehlerfortpflanzung ist, dass unterschiedliche Einflüsse nicht konsequent mit unterschiedlichen Methoden behandelt und fortgepflanzt werden. Das Konzept der Intervalle liefert als Alternative ein lineares Fortpflanzungsgesetz für die Impräzision. Dadurch können zu optimistische Ergebnisse, wie sie mit dem quadratischen Fehlerfortpflanzungsgesetz erhalten werden, relativiert werden. Zufällige Anteile werden weiterhin quadratisch fortgepflanzt.

Prinzipiell besteht die Möglichkeit, die Gesamtunsicherheit oder systematische Fehler allein mit stochastischen Mitteln zu beschreiben. PELZER (1980) schlägt für die Beschreibung der Systematiken beispielsweise Korrelationsmodelle vor, die zu vollbesetzten Varianz-Kovarianz-Matrizen für die Beobachtungen führen. Es wird dabei auf das Elementarfehlermodell nach Bessel und Hagen zurückgegriffen. SCHWIEGER (1999) verwendet diesen Ansatz, um synthetische Korrelationsmatrizen für GPS-Messungen aufzustellen.

Eine weitere Alternative, unsichere Vorinformation über das konkrete stochastische Verhalten einzuführen, stellt die Bayes-Statistik zur Verfügung. In diesem Konzept kann unsichere Vorinformation durch nichtinformativ prior Wahrscheinlichkeitsdichten modelliert werden, vgl. KOCH (2000, S.57f). Mit dem Bayes-Theorem lassen sich daraus a posteriori Dichten für die geschätzten Parameter angeben. Als Pendant zu Konfidenzbereichen der klassischen Statistik werden mit dem Bayes-Ansatz Bereiche höchster posteriori Dichten (HPD-Regionen) zur Irrtumswahrscheinlichkeit  $1 - \gamma$  berechnet. Verschiedene Anwendungen wurden beispielsweise von KOCH (2000) für Varianzkomponentenschätzung, Berechnung von Konfidenzregionen oder die Mustererkennung angegeben.

Schließlich kann die Normalverteilung selbst der Problemstellung angepasst werden. RIESMEIER (1983, S.40ff) verwendet beispielsweise gestutzte Normalverteilungen, um Zentrierunsicherheiten mit zu berücksichtigen. Die Koppelung von Parametern mit gegebenenfalls verschiedenen Verteilungen muss allerdings aufwändig über Faltungen berechnet werden.

Auch wenn verschiedene Modellierungen zu gleichen Zahlenwerten für die Ergebnisse führen, so können doch verschiedene Konzepte zu Grunde liegen. Diese bestimmen letztlich die Unterschiede in der Interpretierbarkeit

und Interpretation der Ergebnisse. Diese Idee soll sich als roter Faden durch die gesamte Arbeit ziehen und durch anschauliche Beispiele unterstützt werden.

## 2.2 Begriffe aus der Intervallmathematik

In diesem Abschnitt werden Begriffe der reellen Intervallmathematik zusammengestellt, soweit sie für das Verständnis der Arbeit von Bedeutung sind. Die Grundidee der Intervallmathematik zielt auf eine Erweiterung der Zusammenhänge der Arithmetik und Analysis reeller auf intervallwertige Größen. In der vorliegenden Arbeit wird ausschließlich mit reellen Intervallen gerechnet, entsprechende Erweiterungen für komplexe Intervalle finden sich beispielsweise in ALEFELD UND HERZBERGER (1974, Kap.5). Die hier verwendete Schreibweise orientiert sich an ALEFELD UND HERZBERGER (1974, 1983) oder MAYER (1989). Auf Beweise soll vollständig verzichtet werden; sie können in den angegebenen Quellen jeweils nachgeschlagen werden.

### 2.2.1 Reelle Intervallrechnung

**Definition 1** Die abgeschlossene Menge

$$[a] := \{ t \in \mathbb{R} \mid \underline{a} \leq t \leq \bar{a}, \quad \underline{a}, \bar{a} \in \mathbb{R} \} \quad (2.2)$$

heißt *abgeschlossenes reelles Intervall*. Die Menge aller abgeschlossenen reellen Intervalle bezeichnet man mit  $\mathbb{IR}$ . Fallen untere und obere Grenze zusammen:  $\underline{a} = \bar{a}$ , so spricht man von einem *Punktintervall*.

Für reelle abgeschlossene Intervalle gelten die klassischen Mengenoperationen. Weiter kann jeder reellen Zahl  $a$  genau ein reelles Intervall der Form  $a = [a, a]$  zugeordnet werden. Da in der vorliegenden Arbeit ausschließlich mit reellen Intervallen gerechnet wird, wird im Folgenden auf den Zusatz *reell* verzichtet.

Neben der in Definition 1 eingeführten Intervalldarstellung durch seine Grenzen kann das Intervall äquivalent durch seinen Mittelpunkt  $a_m$  und Radius  $a_r$  angegeben werden

$$a_m := \frac{1}{2}(\bar{a} + \underline{a}), \quad a_r := \frac{1}{2}(\bar{a} - \underline{a}). \quad (2.3)$$

Fasst man die beiden Intervallgrenzen bzw. Mittelpunkt und Radius als geordnete Paare auf, so lässt sich (2.3) als Transformation darstellen:

$$(a_m \quad a_r)^T = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} (\underline{a} \quad \bar{a})^T, \quad (\underline{a} \quad \bar{a})^T = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} (a_m \quad a_r)^T. \quad (2.4)$$

**Definition 2** Sei  $[a]$  ein Intervall, so ist sein Durchmesser  $w([a])$  gegeben durch

$$w([a]) := 2 a_r = \bar{a} - \underline{a}. \quad (2.5)$$

Der Durchmesser eines Intervalls, der eng mit dem Intervallradius verknüpft ist, eignet sich als Kenngröße für Intervalle. Er gibt die maximale Variationsbreite für  $a$  an. Rechenregeln für den Durchmesser von Intervallen findet man beispielsweise bei ALEFELD UND HERZBERGER (1974, S.20ff).

**Definition 3** Bezeichne  $s(x)$  eine stetige einstellige Operation (*Standardfunktion*) in  $\mathbb{R}$ , dann ist durch

$$s([x]) := \left[ \min_{x \in [x]} s(x), \max_{x \in [x]} s(x) \right] \quad (2.6)$$

die zugehörige einstellige Operation (*Intervallstandardfunktion*) in  $\mathbb{IR}$  definiert, vgl. ALEFELD UND HERZBERGER (1974, S.3).

Standardfunktionen sind beispielsweise die Funktionen  $x^\alpha$  ( $\alpha \in \mathbb{R}$ ),  $e^x$ ,  $\ln x$  oder die trigonometrischen Funktionen. Ist  $s$  monoton wachsend in  $x \in [x]$ , so ergibt sich direkt  $s([x]) = [s(\underline{x}), s(\bar{x})]$ . Entsprechend gilt für monoton fallende Funktionen  $s([x]) = [s(\bar{x}), s(\underline{x})]$ .

**Beispiel 1** Sei  $[x] = [0, 1]$ , so gilt für die Intervallauswertung der monoton wachsenden Standardfunktion  $s(x) = e^x$ :  $e^{[x]} = e^{[0,1]} = [e^0, e^1] = [1, e]$ .

**Beispiel 2** Betrachtet man die Standardfunktion  $s(x) = x^2$  im Intervall  $[x] = [-1, 4]$ , so liefert die Auswertung der Intervallstandardfunktion:  $s([x]) = s([-1, 4]) = \left[ \min_{x \in [-1, 4]} x^2, \max_{x \in [-1, 4]} x^2 \right] = [0, 16]$ .

Voraussetzung fur die Auswertung der Standardfunktionen ist die Stetigkeit der Funktionen uber dem betrachteten Intervall.

**Beispiel 3** Betrachten wir hierzu die Intervalle  $[x] = [-1, 1]$ ,  $[y] = [1, 2]$ . So ist  $\frac{1}{[x]}$  nicht definiert, da die Funktion im gegebenen Intervall nicht definiert ist (Polstelle in  $x = 0$ ), wohingegen  $\frac{1}{[y]} = [\frac{1}{2}, 1]$  ergibt.

Mit diesen Definitionen fur Intervalle wurden die Grundlagen fur die Verknupfungen mehrerer Intervalle gelegt, die im Folgenden beschrieben werden, vgl. ALEFELD UND HERZBERGER (1974, Def.2, S.2) oder MAYER (1989, S.102).

**Definition 4** Seien  $[a], [b] \in \mathbb{IR}$  zwei Intervalle, so gilt fur ihre Verknupfung  $\circ \in \{+, -, \cdot, /\}$ :

$$[a] + [b] := [\underline{a} + \underline{b}, \bar{a} + \bar{b}], \quad (2.7)$$

$$[a] - [b] := [\underline{a} - \bar{b}, \bar{a} - \underline{b}], \quad (2.8)$$

$$[a] \cdot [b] := [\min\{\underline{a}\underline{b}, \underline{a}\bar{b}, \bar{a}\underline{b}, \bar{a}\bar{b}\}, \max\{\underline{a}\underline{b}, \underline{a}\bar{b}, \bar{a}\underline{b}, \bar{a}\bar{b}\}], \quad (2.9)$$

$$[a]/[b] := [a] \cdot \left[ \frac{1}{\bar{b}}, \frac{1}{\underline{b}} \right], \text{ falls } 0 \notin [b]. \quad (2.10)$$

Fur den Spezialfall, dass  $[a]$  in (2.9) eine Zahl ist, vereinfachen sich die Ausdrucke zu:

$$[c] = [a] \cdot [b] = a [b] = \begin{cases} [a\underline{b}, a\bar{b}], & a \geq 0 \\ [a\bar{b}, a\underline{b}], & a < 0 \end{cases}. \quad (2.11)$$

Mittelpunkt und Radius des Ergebnisintervalls lassen sich in diesem Fall unter Verwendung des Absolutbetrags  $|a|$  angeben:

$$c_m = a b_m, \quad c_r = |a| b_r. \quad (2.12)$$

MAYER (1989, S.103) gibt weitere Vereinfachungen fur die Formeln fur Multiplikation und Addition an, wenn die Lage der Intervalle bezuglich der Null bekannt ist.

Um die Unterschiede bzw. Parallelen zwischen dem Rechnen im Reellen und entsprechenden Operationen mit Intervallen zusammenzustellen, wird auf wichtige Eigenschaften der in Definition 4 eingefuhrten Verknupfungen eingegangen.

1. Das Ergebnis der Verknupfung zweier Intervalle ist wieder ein Intervall, d.h.  $[a] \circ [b] \in \mathbb{IR}$ .

2. Es gilt die *Einschließungseigenschaft*:

$$a \in [a], b \in [b] \Rightarrow a \circ b \in [a] \circ [b]. \quad (2.13)$$

Das Intervall  $\square \mathcal{M} := [a] \circ [b]$  ist die *Intervallhulle* der Menge  $\mathcal{M} = \{a \circ b \mid a \in [a], b \in [b]\}$ , d.h. die engstmogliche Einschließung von  $\mathcal{M}$  durch ein Intervall. Als Verallgemeinerung von (2.13) gilt die *Inklusionsmonotonie*:

$$[a] \subseteq [b], [c] \subseteq [d] \Rightarrow [a] \circ [c] \subseteq [b] \circ [d]. \quad (2.14)$$

3. Addition und Multiplikation von Intervallen sind assoziativ und kommutativ.

4. Das Neutralelement ist das Punktintervall  $0 = [0, 0]$  fur die Addition bzw.  $1 = [1, 1]$  fur die Multiplikation.

5. Da fur „echte“ Intervalle ( $\underline{a} \neq \bar{a}$ ) gilt, haben die Gleichungen

$$\begin{aligned} [a] + [x] &= [0, 0], \\ [a] \cdot [x] &= [1, 1] \end{aligned} \quad (2.15)$$

im Gegensatz zum Rechnen mit reellen Groen keine Losungen  $[x]$ . Folglich existiert kein inverses Element fur Addition bzw. Multiplikation.

6. Anstelle des Distributivgesetzes für reelle Operationen tritt im Allgemeinen die abgeschwächte Form der Subdistributivität:

$$[a] ([b] + [c]) \subseteq [a][b] + [a][c], \tag{2.16}$$

wobei die Gleichheit (Distributivität) beispielsweise gilt, wenn  $\underline{b} \geq 0, \underline{c} \geq 0$ .

**Definition 5** Seien  $[a], [b] \in \mathbb{IR}$ , so ist ihr Abstand (Hausdorff-Distanz) definiert durch:

$$d([a], [b]) := \max\{|\underline{a} - \underline{b}|, |\bar{a} - \bar{b}|\}. \tag{2.17}$$

Der Abstand zweier Intervalle  $[a], [b]$  eignet sich zu ihrem Vergleich. Die Rechenregeln sind beispielsweise in ALEFELD UND HERZBERGER (1974, S.14ff) zusammengestellt.

### 2.2.2 Reelle Intervallalgebra

Das Konzept der Intervalle wird auf Matrizen (Vektoren) erweitert, wobei die reellwertigen Matrixelemente bzw. Vektorkomponenten durch Intervalle ersetzt werden. Entsprechend den Operationen im Reellen sind Intervalloperationen zu definieren.

**Definition 6** Unter einem *Intervallvektor* soll der folgende Vektor

$$[\mathbf{x}] := ([x_1], \dots, [x_n])^T, [x_i] \in \mathbb{IR}, i = 1, \dots, n \tag{2.18}$$

und unter einer *Intervallmatrix* die folgende Matrix verstanden werden:

$$[\mathbf{A}] := ([a_{ij}]), [a_{ij}] \in \mathbb{IR}, i = 1, \dots, n, j = 1, \dots, m. \tag{2.19}$$

Die Menge der  $n \times 1$  Intervallvektoren wird mit  $\mathbb{IR}^n$  bezeichnet und die Menge der  $n \times m$  Intervallmatrizen mit  $\mathbb{IR}^{n \times m}$ . Fallen untere und obere Grenze zusammen:  $\underline{\mathbf{A}} = \overline{\mathbf{A}}$ , bzw.  $\underline{\mathbf{x}} = \overline{\mathbf{x}}$  so spricht man von einer *Punktmatrix* bzw. einem *Punktvektor*. Aus der Definition wird deutlich, dass die reellen Intervalle als  $1 \times 1$  Matrix aufgefasst werden können. Intervallvektoren bzw. Intervallmatrizen lassen sich ebenfalls durch Mittelpunkt und Radius darstellen, wobei mit  $\mathbf{x}_r$  der Vektor der Intervallradien und mit  $\mathbf{x}_m$  der Vektor der Intervallmitten bezeichnet wird.

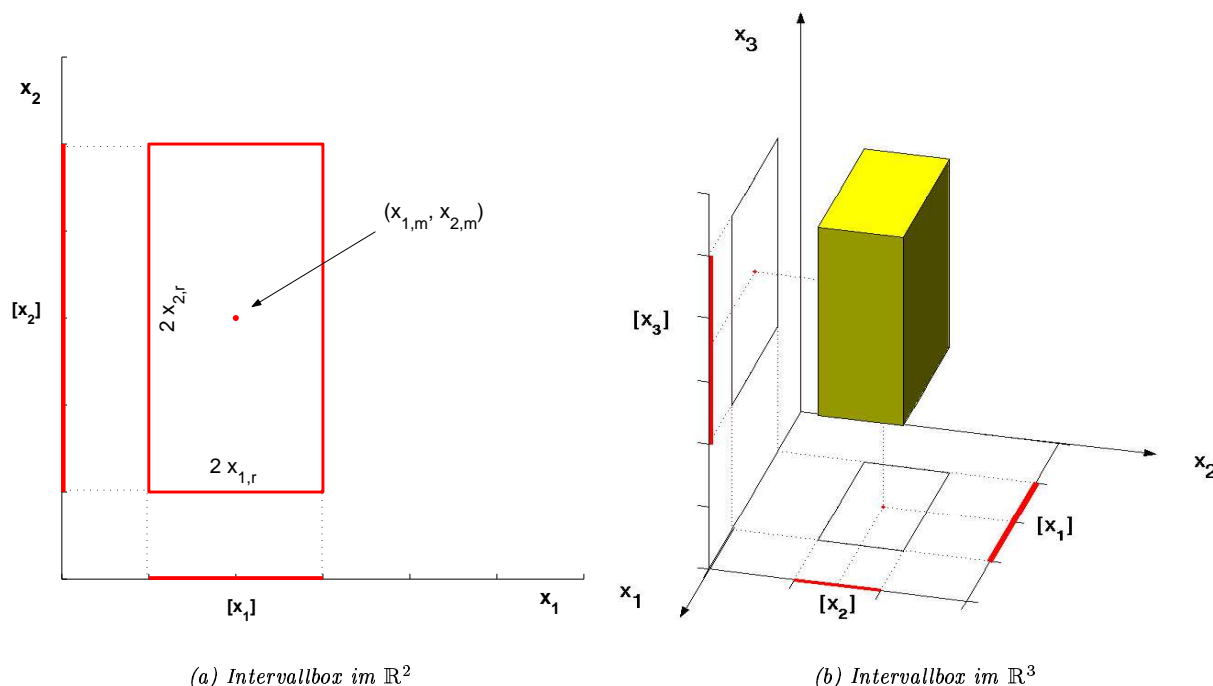


Abbildung 2.3: Darstellung der Intervallboxen

Ein Zugang zur geometrischen Veranschaulichung von Intervallvektoren liefert ihre Darstellung als kartesisches Produkt ihrer Komponentenintervalle (JAULIN ET AL. 2001, S.23ff):

$$[\mathbf{x}] = [x_1] \times [x_2] \times \dots \times [x_n]. \quad (2.20)$$

Geometrisch stellen Intervallvektoren Hyperquader (*Intervallboxen*) dar, wie sie in Abbildung 2.3 fur den zwei- und dreidimensionalen Fall veranschaulicht sind. Die gezeigten Intervallboxen lassen sich wie folgt konstruieren: Die Koordinaten des Schwerpunktes der Figur sind durch den Vektor der Intervallmitten angegeben. Im Zweidimensionalen ergeben sich Rechtecke, deren Seiten parallel zu den Koordinatenachsen des  $\mathbb{R}^2$  verlaufen. Ihre Langen sind identisch mit dem Durchmesser der Intervalle der einzelnen Komponenten:  $w([x_1]) = 2 x_{1,r}$ ,  $w([x_2]) = 2 x_{2,r}$ . Im Dreidimensionalen ergeben sich Quader, deren Kanten parallel zu den Koordinatenachsen des  $\mathbb{R}^3$  verlaufen und die Langen  $w([x_1]) = 2 x_{1,r}$ ,  $w([x_2]) = 2 x_{2,r}$ ,  $w([x_3]) = 2 x_{3,r}$  haben. Die Orthogonalprojektion der Intervallbox auf die einzelnen Koordinatenachsen ergibt gerade die Komponenten des Intervallvektors.

Die Addition und Subtraktion von Intervallvektoren ist ebenso wie die Multiplikation eines Intervallvektors mit einem Intervall komponentenweise erklart. Der Durchmesser eines Intervallvektors wird als Maximum der Durchmesser seiner Komponenten beschrieben. Weitere Rechenoperationen fur Intervallmatrizen sind beispielsweise in ALEFELD UND HERZBERGER (1974, Kap.10) oder NEUMAIER (1990) angegeben. Insbesondere soll darauf hingewiesen werden, dass der  $\mathbb{R}^n$  und  $\mathbb{R}^{n \times m}$  keine Vektorrume sind, da beispielsweise die inverse Operation zur Addition nicht erklart ist. Fur weitere Diskussion vergleiche KUTTERER (1994, S.17f).

Ein besonderer Blick soll abschlieend auf lineare Abbildungen gelenkt werden, wie sie in linearen Modellen bei der geodatischen Datenanalyse vorkommen:

$$\mathbf{x} = \mathbf{F} \mathbf{l}, \quad \mathbf{l} \in \mathbb{R}^n, \mathbf{x} \in \mathbb{R}^u, u < n. \quad (2.21)$$

Wird der Vektor  $\mathbf{l}$  durch den Intervallvektor  $[\mathbf{l}]$  ersetzt, so stellt  $[\mathbf{x}]$  die mittels  $\mathbf{F}$  linear ubertragenen Intervalle dar, wobei die Gleichheit komponentenweise zu sehen ist,

$$[\mathbf{x}] = \mathbf{F} \cdot [\mathbf{l}] \Leftrightarrow \begin{cases} \mathbf{x}_m &= \mathbf{F} \mathbf{l}_m \\ \mathbf{x}_r &= |\mathbf{F}| \mathbf{l}_r \end{cases} \quad (2.22)$$

$$\mathcal{Z} = \{ \mathbf{x} \in \mathbb{R}^u \mid \mathbf{x} = \mathbf{F} \mathbf{l}, \mathbf{l} \in [\mathbf{l}] \} \subseteq \square \mathcal{Z} = \mathbf{F} \cdot [\mathbf{l}] = [\mathbf{x}]. \quad (2.23)$$

Unter  $|\mathbf{F}|$  ist der elementweise Betrag der Matrix  $\mathbf{F}$  zu verstehen. Der Intervallvektor  $[\mathbf{x}]$  stellt die Intervallhulle der Menge  $\mathcal{Z}$  dar, d.h. ihre engste Einschlieung im gegebenen Koordinatensystem. Sei  $\mathbf{F}$  nun durch das Matrixprodukt  $\mathbf{F} = \mathbf{A} \mathbf{B}$  gegeben, so ist (2.22) wegen

$$(\mathbf{A} \mathbf{B}) \cdot [\mathbf{l}] \subseteq \mathbf{A} \cdot (\mathbf{B} \cdot [\mathbf{l}]) \quad (2.24)$$

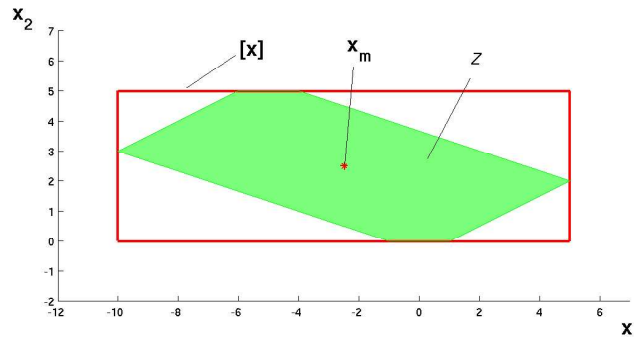
nur dann die kleinste Einschlieung, wenn die Klammerung wie auf der linken Seite durchgefuhrt wird (ALEFELD UND HERZBERGER 1974, Satz 4(9), S.151). Die Gleichheit in (2.24) gilt fur die Klasse der nichtnegativen Matrizen.

**Beispiel 4** Zur Veranschaulichung der Zusammenhange wahlen wir folgende Matrix und folgenden Intervallvektor:

$$\mathbf{F} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{pmatrix}, \quad [\mathbf{l}] = \begin{pmatrix} [-1, 1] \\ [0, 2] \\ [-3, 0] \end{pmatrix}. \quad (2.25)$$

Der Ergebnisintervallvektor ist dann  $[\mathbf{x}] = ([-10, 5], [0, 5])^T$ .

Der Ergebnisintervallvektor  $[\mathbf{x}]$  ist durch die achsparallele Intervallbox gegeben. Diese schliet die Menge  $\mathcal{Z}$ , den *Wertebereich der linearen Abbildung* oder *Komplex* (SCHMITT 1977), in der gegebenen Parameterbasis bestmoglich ein. Die Menge  $\mathcal{Z}$  ist konvex und zentralsymmetrisch bezuglich der Position  $\mathbf{x}_m$ . Auf weitere Eigenschaften wird in Kapitel 4.3 eingegangen. Der Wertebereich wird durch die achsparallele Intervallbox uberschatzt, da Korrelationen zwischen Intervallen bei der linearen Abbildung vernachlassigt werden. Wird der Intervallvektor  $[\mathbf{x}]$  weiterverarbeitet, beispielsweise in einem Iterationsverfahren  $[\mathbf{x}]^{(i+1)} = \mathbf{B} \cdot [\mathbf{x}]^{(i)}$ , so blaht sich das Ergebnisintervall  $[\mathbf{x}]^{(i+1)}$  immer weiter auf, vgl. (2.24). Bei der iterativen Berechnung dynamischer Systeme tritt diese uberschatzung beispielsweise besonders auf und wird als *wrapping effect* bezeichnet (LOHNER 1989).

Abbildung 2.4: Intervallbox  $[x]$  und Wertebereich  $Z$ 

### 2.2.3 Funktionen mit Intervallen als Variablen

**Definition 7** Sei  $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ ,  $[x] \subseteq D$  eine stetige Abbildung, so ist der Wertebereich von  $f$  bezüglich des Intervalls  $[x]$  gegeben zu:

$$\begin{aligned} \mathcal{W}_f([x]) &:= \{ f(x) \mid x \in [x] \} \\ &= \left[ \min_{x \in [x]} f(x), \max_{x \in [x]} f(x) \right]. \end{aligned} \quad (2.26)$$

**Definition 8** Unter einer *Intervallauswertung von  $f$*  soll das Ergebnisintervall  $f([x])$  verstanden werden, das beim Ersetzen des reellen Arguments  $x$  durch das Intervalls  $[x]$  und Anwendung der Rechenregeln der Intervallmathematik statt der entsprechenden reellen Operationen erhalten wird. Die Funktion ist hierzu in Standardfunktionen zu zerlegen. Diese werden nach (2.6) ausgewertet und die Ergebnisintervalle mit Hilfe der Rechenregeln (2.7)-(2.10) verknüpft.

**Beispiel 5** Betrachten wir die Funktion  $f(x) = \ln(2e^x - \sqrt{x})$  im Intervall  $[x] = [0, 1]$ . Die Funktion besteht aus drei Standardfunktionen, die ausgewertet werden:  $\sqrt{[x]} = \sqrt{[0, 1]} = [0, 1]$ ,  $e^{[x]} = [1, e]$ , vgl. Bsp. 1. Damit ergibt sich

$$\begin{aligned} f([x]) &= \ln(2 \cdot [1, e] - [0, 1]) \\ &= \ln([2, 2e] - [0, 1]) \\ &= \ln([1, 2e]) \\ &= [0, 1 + \ln(2)]. \end{aligned} \quad (2.27)$$

Eine grundlegende Problematik bei der Intervallauswertung von Funktionen liegt darin, dass Äquivalenzumformungen im reellwertigen Funktionsausdruck im Allgemeinen nicht für die entsprechende Intervallauswertung gelten. Dies liegt beispielsweise in der Subdistributivität für Intervalloperationen bzw. im Fehlen des inversen Elements der Addition und Multiplikation begründet. Das folgende Beispiel (JAULIN ET AL. 2001, S.30f, bzw. ALEFELD UND HERZBERGER 1974, S.36) soll dies drastisch verdeutlichen:

**Beispiel 6** Gegeben seien vier äquivalente Ausdrücke für die Funktion  $f(x)$ :

$$f_1(x) = x(x+1), \quad (2.28)$$

$$f_2(x) = x x + x, \quad (2.29)$$

$$f_3(x) = x^2 + x, \quad (2.30)$$

$$f_4(x) = \left(x + \frac{1}{2}\right)^2 - \frac{1}{4}. \quad (2.31)$$

Die Intervallauswertung in  $[x] = [-1, 1]$  liefert:

$$f_1([x]) = [x] \cdot ([x] + 1) = [-2, 2], \quad (2.32)$$

$$f_2([x]) = [x] \cdot [x] + [x] = [-2, 2], \quad (2.33)$$

$$f_3([x]) = [x^2] + [x] = [-1, 2], \quad (2.34)$$

$$f_4([x]) = \left([x] + \frac{1}{2}\right)^2 - \frac{1}{4} = \left[-\frac{1}{4}, 2\right] = \mathcal{W}_f([x]). \quad (2.35)$$

Je nach Wahl des Funktionsausdruckes wird der Wertebereich mehr oder weniger *überschätzt*, wobei nur die Intervallauswertung von  $f_4(x)$  dem Wertebereich entspricht. Allgemein ist die Beziehung zwischen Wertebereich und Intervallauswertung einer Funktion durch die *Einschließungseigenschaft* charakterisiert, vgl. ALEFELD UND HERZBERGER (1974, S.31):

$$\mathcal{W}_f([x]) \subseteq f([x]). \quad (2.36)$$

Im Folgenden soll diskutiert werden, unter welchen Voraussetzungen die Intervallauswertung dem Wertebereich entspricht bzw. wie dieser bestmöglich approximiert werden kann. Die Betrachtungen werden dabei auf stetig differenzierbare Funktionen  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $[x] \subseteq D$  erweitert. Im Vektor  $\mathbf{x}$  werden alle Parameter, Variablen und Konstanten der Funktion zusammengefasst, die als intervallwertig angenommen werden sollen.

Prinzipiell lässt sich die Bestimmung des Wertebereichs als globale Optimierungsaufgabe für den Ausdruck (2.26) begreifen, vgl. JAULIN ET AL. (2001, S.29). Je nach Art der zu untersuchenden Funktionen und Anzahl der Parameter ist die globale Optimierungsaufgabe nur mit sehr großem Aufwand lösbar. Es scheint also sinnvoll, die Einschließungseigenschaft (2.36) zu diskutieren.

Als erstes wird untersucht, für welche Gruppen von Funktionen Wertebereich und Intervallauswertung übereinstimmen. Zu dieser Gruppe gehören Funktionen, bei denen jeder der Parameter  $(x_1, \dots, x_n)$  höchstens einmal im Funktionsausdruck vorkommt, vgl. ALEFELD UND HERZBERGER (1974, S.32). Dies ist offensichtlich für die Funktion  $f_4$  im Beispiel 6 der Fall. Bei den anderen Darstellungen der Funktion treten die Parameter mehrfach auf. Sie werden bei der Intervallauswertung wie unabhängige Variablen behandelt, wodurch Abhängigkeiten wie beispielsweise  $[x] - [x] \neq [0, 0]$  vernachlässigt und zusätzliche Freiheitsgrade eingeführt werden. Dies führt zu einer echten Überschätzung. Weitere Funktionsgruppen, für die die Gleichheit gilt, sind in ALEFELD UND HERZBERGER (1974, S.32) aufgeführt.

Für beliebige Funktionen lassen sich keine allgemeinen Aussagen treffen, so dass jeder Einzelfall untersucht werden muss (MOORE 1969, Kap.6). Eine einfache Funktionendarstellung, in der die Approximationsgüte des Wertebereichs durch die Funktion abgeschätzt werden kann, ist die *zentrische Form*, vgl. ALEFELD UND HERZBERGER (1974, S.37f). Sie kann zum Beispiel aus dem Mittelwertsatz abgeleitet werden (JAULIN ET AL. 2001, S.33): Es existiert eine Zwischenstelle  $\xi \in [x]$  derart, dass für alle  $\mathbf{x} \in [x]$  gilt:

$$f(\mathbf{x}) = f(\mathbf{x}_m) + \left( \frac{\partial f}{\partial \mathbf{x}}(\xi) \right)^T (\mathbf{x} - \mathbf{x}_m). \quad (2.37)$$

Folglich gilt auch:

$$f(\mathbf{x}) \in f(\mathbf{x}_m) + \left( \frac{\partial f}{\partial \mathbf{x}}([x]) \right)^T (\mathbf{x} - \mathbf{x}_m) \quad (2.38)$$

und die Intervallauswertung von (2.38) liefert:

$$f([x]) \subseteq f_T([x]) := f(\mathbf{x}_m) + \left( \frac{\partial f}{\partial \mathbf{x}}([x]) \right)^T ([x] - \mathbf{x}_m). \quad (2.39)$$

Die rechte Seite von (2.39) entspricht einem Kegel mit Spitze in  $(\mathbf{x}_m, f(\mathbf{x}_m))$ . Die Variationsbreite des Gradienten beschreibt den Öffnungswinkel des Kegels. Die folgenden Abbildungen stellen den Zusammenhang graphisch dar. Der grau schraffierte Bereich veranschaulicht die Variationsregion für den Gradienten, durch den die Funktion und damit der Wertebereich approximiert werden. Die Größe des Bereichs hängt vom Durchmesser des Eingangsintervalls ab. Im Punkt der Intervallmitten stimmen die Approximation und der Funktionswert überein.

JAULIN ET AL. (2001) zeigen weiter, dass die Approximationsgüte vom Durchmesser der Eingangsintervalle abhängt, da diese den Einfluss der Variationsbreite des Gradienten steuern. Je kleiner der Durchmesser ist, desto besser approximiert die zentrische Form die Intervallauswertung der Funktion in der ursprünglichen Darstellung. ALEFELD UND HERZBERGER (1974) zeigen unter Voraussetzung der Lipschitzbedingung, dass bei Reduktion des Intervalldurchmessers der Parameter die Intervallauswertung gegen den Wertebereich konvergiert, im Falle einer zentrischen Form sogar quadratisch (ALEFELD UND HERZBERGER 1974, S.33ff).

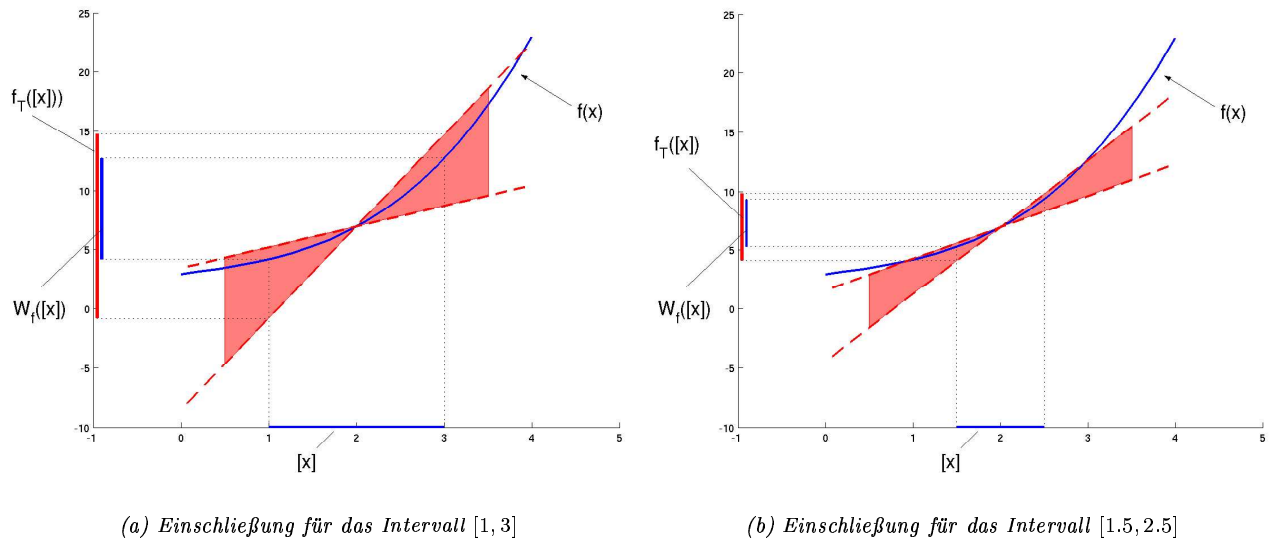


Abbildung 2.5: Darstellung der Intervalleinschließung des Wertebereichs  $\mathcal{W}_f([x])$  durch die zentrische Form  $f_T([x])$  für zwei unterschiedliche Durchmesser des Intervalls  $[x]$ . In Grau: Variationsbereich des Gradienten

Prinzipiell muss jedoch kritisch festgestellt werden, dass durch Anwendung des Mittelwertsatzes das Problem der Intervallauswertung der ursprünglichen Funktion auf die Auswertung des Gradienten bzw. der  $n$  partiellen Ableitungen verlagert wird, wobei sich im Allgemeinen der Aufwand zur Auswertung vergrößert. Dies kann nur abgefangen werden, wenn sich die Intervallauswertung des Gradienten wesentlich einfacher gestaltet als die der Ausgangsfunktion oder der Wertebereich des Gradienten sogar exakt angegeben werden kann. Zusätzlich sind Erfahrungswerte wichtig, um bei beliebigen Funktionsausdrücken diesen im Sinne einer besten Approximation des unbekanntes Wertebereichs umzuformen zu können, vgl. Beispiel 6. JAULIN ET AL. (2001, S.36) schlagen vor, für verschiedene Darstellungen der Funktion die Intervalleinschließungen auszuwerten und davon die Schnittmenge zu bilden. Die Effizienz dieses Vorgehens ist allerdings durch die Parameteranzahl beschränkt. Vergleiche für verschiedene Approximationen werden beispielsweise in JAULIN ET AL. (2001, S.36ff) angestellt.

Abschließend wird eine Strategie vorgeschlagen, die einerseits eindeutige Ergebnisintervalle liefert, die unabhängig von der Darstellung der Funktion und somit sehr gut für Anwender geeignet sind, die sich wenig mit Intervallmathematik auskennen. Andererseits soll eine ausreichende Approximation des Wertebereichs erhalten werden. Für alle Funktionen, bei denen die intervallauszuwertenden Parameter (Konstanten, Variablen) nur einfach vorkommen, kann der Wertebereich exakt direkt angegeben werden. In anderen Fällen soll der Wertebereich durch die folgende Funktion approximiert, im Allgemeinen aber nicht eingeschlossen werden:

$$\begin{aligned}
 f_A([\mathbf{x}]) &:= f(\mathbf{x}_m) + \left( \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}_m) \right)^T ([\mathbf{x}] - \mathbf{x}_m) \\
 &= f(\mathbf{x}_m) + \left( \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}_m) \right)^T [-\mathbf{x}_r, \mathbf{x}_r].
 \end{aligned}
 \tag{2.40}$$

Die Funktion  $f$  wird an der Stelle der zu betrachtenden Intervallmitten in eine Taylorreihe entwickelt und intervallausgewertet. Gleichung (2.40) stellt eine spezielle Funktion der Menge  $f_T([\mathbf{x}])$  aus (2.39) dar, da die partiellen Ableitungen lediglich an der Stelle der Intervallmitten bestimmt werden und keine Intervallauswertung des Gradienten vorgenommen wird. Da der Einfluss der Steigung mit dem Intervallradius gekoppelt ist, konvergiert  $f_A([\mathbf{x}])$  wie auch  $f_T([\mathbf{x}])$  gegen den Wertebereich bei verschwindenden Intervallradien.

Der Vergleich der beiden Abbildungen 2.6(a) und (b) veranschaulicht die Verbesserung der Approximation des Wertebereichs bei Halbierung der Intervalldurchmesser des Parameters  $x$ . Neben dieser Konvergenzeigenschaft lassen sich noch weitere Aussagen über die Approximation des Wertebereichs durch  $f_A([\mathbf{x}])$  treffen. Ist die Funktion im betrachteten Intervall entweder monoton wachsend oder monoton fallend und konvex von unten, d.h. die Hesse-Matrix  $\left( \frac{\partial^2 f}{\partial \mathbf{x}^2} \right)$  ist positiv semidefinit, so verläuft der Graph von  $f_A([\mathbf{x}])$  immer unterhalb des Graphen von  $f(\mathbf{x})$ . Folglich gilt:

$$f_A(\underline{\mathbf{x}}) \leq \min\{ \mathcal{W}_f([\mathbf{x}]) \} \leq f_A(\bar{\mathbf{x}}) \leq \max\{ \mathcal{W}_f([\mathbf{x}]) \}.
 \tag{2.41}$$



Entsprechendes gilt umgekehrt fur Funktionen, die konvex von oben sind. Der Extremfall einer schlechten Approximation tritt ein, falls in  $f(\mathbf{x}_m)$  ein lokales Extremum vorliegt, so dass  $f_A([\mathbf{x}]) = [f(\mathbf{x}_m), f(\mathbf{x}_m)]$  gilt. Allgemeine Aussagen konnen nur eingeschrankt getroffen werden, und die einzelnen Funktionen mussen im Detail untersucht werden. In Abschnitt 3.4 wird die Approximationsgute anhand einiger Beispiele fur Korrektionsfunktionen bei Streckenmessungen dargestellt werden.

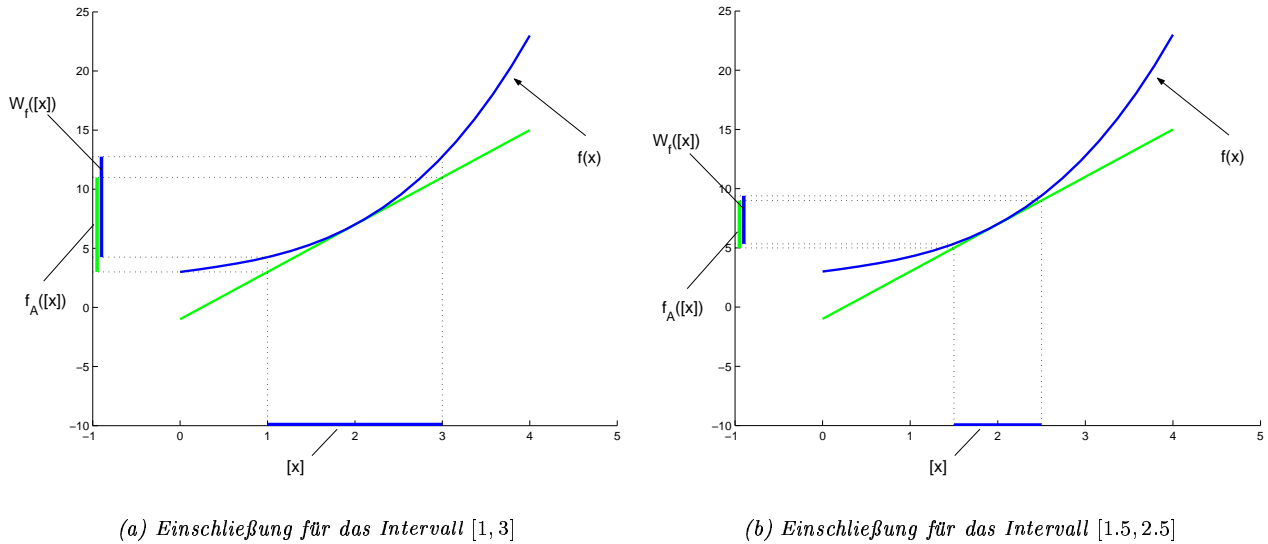


Abbildung 2.6: Darstellung der Approximation des Wertebereichs  $W_f([x])$  durch die Funktion  $f_A([x])$  fur zwei unterschiedliche Durchmesser des Intervalls  $[x]$ .

Abschließend soll kurz auf die Intervallauswertung vektorwertiger Funktionen eingegangen werden. Allgemein lasst sich festhalten: Gelten Eigenschaften fur alle Komponenten einer Funktion, so gelten sie auch fur die gesamte Funktion (JAULIN ET AL. 2001, S.29). Obwohl bei vektorwertigen Funktionen wie bei linearen Abbildungen (siehe Abschnitt 2.2.2) komponentenweise Exaktheit erreicht werden kann, uberschatzt im Allgemeinfeld der Vektor  $\mathbf{f}([x])$  den Wertebereich. Weitere Darstellungen zu vektorwertigen Intervallfunktionen finden sich bei ALEFELD UND HERZBERGER (1974) oder JAULIN ET AL. (2001).

## 2.2.4 Intervalle und Fuzzy-Sets

Dieser Abschnitt stellt den Zusammenhang zwischen Intervallen und unscharfen Mengen (*Fuzzy-Sets*) her. Es wird gezeigt, dass Intervalle spezielle Fuzzy-Sets sind. Die Darstellung orientiert sich an BANDEMER UND GOTTWALD (1993). Fur die Anwendung von Fuzzy-Sets in der Geodasie wird beispielsweise auf HEINE (1999) oder KUTTERER (1994, 2002b) verwiesen. Eine allgemeine ubersicht zu diesem Thema bietet das Handbuch RUPINI ET AL. (1998).

**Definition 9** Bezeichne  $\mathcal{U}$  das Universum oder die Grundmenge und  $\mathcal{M}$  eine beliebige Menge im Sinn der klassischen Mengentheorie. Dann wird die Zugehorigkeit zu einer Menge durch die *Indikatorfunktion*  $m_{\mathcal{M}} : \mathcal{U} \rightarrow \{0, 1\}$  formalisiert. Fur alle  $u \in \mathcal{U}$  gilt:

$$m_{\mathcal{M}}(u) = \begin{cases} 1, & \text{falls } u \in \mathcal{M} \\ 0, & \text{sonst.} \end{cases} \quad (2.42)$$

Die Indikatorfunktion kann verfeinert werden, so dass alle Zugehorigkeitsgrade zwischen 0 (Element gehort nicht zur Menge) und 1 (Element gehort zur Menge) zugelassen sind:  $m_{\tilde{A}} : \mathcal{U} \rightarrow [0, 1]$  (*Zugehorigkeitsfunktion*).

**Definition 10** Sei  $m_{\tilde{A}} : \mathcal{U} \rightarrow [0, 1]$  eine Zugehorigkeitsfunktion, so ist die *unscharfe Menge* definiert als

$$\tilde{A} := \{ (u, m_{\tilde{A}}(u)) \mid u \in \mathcal{U} \}. \quad (2.43)$$

Da jedes Intervall  $[a]$  eine (klassische) Menge darstellt, ist  $[a]$  auch ein Fuzzy-Set. Die Zugehörigkeitsfunktion ist dann die Indikatorfunktion, vgl. (2.42). Folglich ist die intervallmathematische Beschreibung in die Theorie der Fuzzy-Sets eingebettet.

**Definition 11** Unter einem strengen  $\alpha$ -Schnitt einer unscharfen Menge ist die folgende (klassische) Menge zu verstehen (BANDEMER UND GOTTWALD 1993, S.23f):

$$\tilde{A}_\alpha := \{ u \in \mathcal{U} \mid m_{\tilde{A}}(u) > \alpha, \quad \alpha \in [0, 1] \}. \quad (2.44)$$

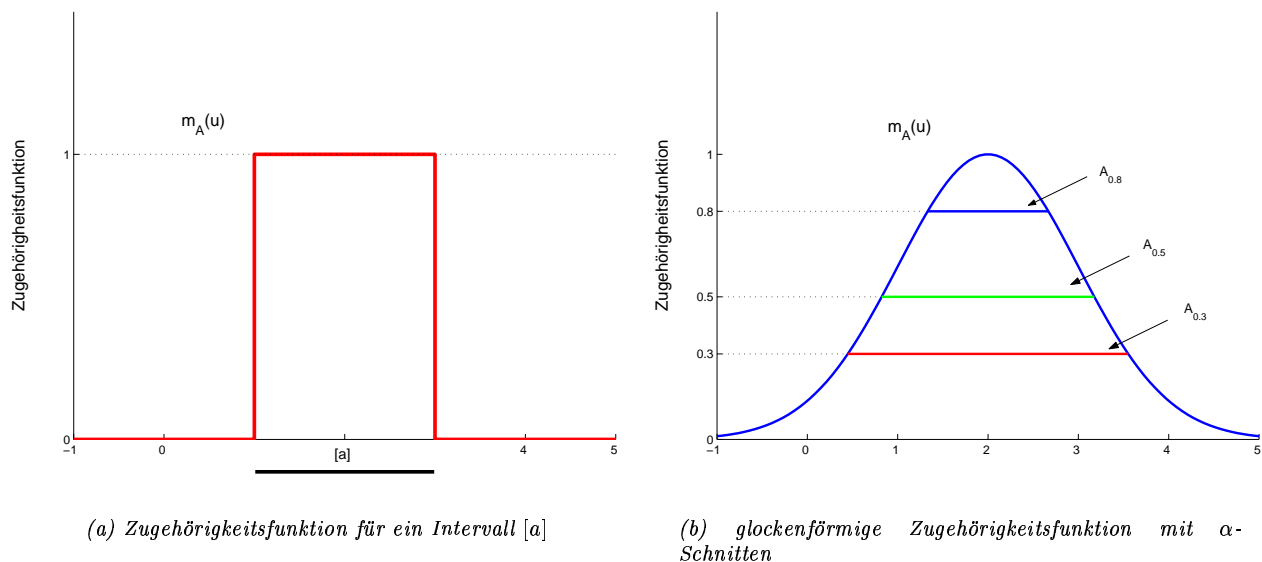


Abbildung 2.7: Zugehörigkeitsfunktionen und  $\alpha$ -Schnitte

Graphisch lässt sich der Zusammenhang wie folgt darstellen. Abbildung 2.7 (a) zeigt die Zugehörigkeitsfunktion für Intervalle (Indikatorfunktion). Sie hat den Wert Null außerhalb des Intervalls und Eins im Intervall  $[\underline{a}, \bar{a}]$ . Durch die Wahl anderer Zugehörigkeitsfunktionen als der Indikatorfunktion können die Aussagen der intervallmathematischen Betrachtung jederzeit verfeinert werden, deshalb beschränkt sich die vorliegende Arbeit explizit auf die intervallmathematische Beschreibung. Abbildung 2.7 (b) zeigt beispielsweise eine glockenförmige Zugehörigkeitsfunktion und  $\alpha$ -Schnitte für  $\alpha = 0.3, 0.5$  und  $0.8$ . Weitere Eigenschaften von Fuzzy-Sets können beispielsweise bei DUBOIS UND PRADE (1980), BANDEMER UND GOTTWALD (1993) oder VIERTL (1996) nachgelesen werden. In Abschnitt 4.4 wird nochmals auf Eigenschaften von Fuzzy-Sets zurückgegriffen, um mit dem *Erweiterungsprinzip nach Zadeh* Impräzision und klassische Fehlermaße elegant zu kombinieren (KUTTERER 2002b, S.67ff).

### 2.2.5 Intervalle und gleichverteilte Zufallsvariablen

Intervalle werden oft fälschlicherweise mit gleichverteilten Zufallsvariablen assoziiert. Der folgende Abschnitt soll kurz Analogien und deren Grenzen aufzeigen, auf die auch KUTTERER (1994, S.38f) eingegangen ist.

**Definition 12** Eine Verteilung für die Zufallsvariable  $x \in \mathbb{R}$  mit der Wahrscheinlichkeitsdichte

$$\phi_x(x) = \begin{cases} \frac{1}{b-a}, & \text{für } a < x < b \\ 0, & \text{sonst,} \end{cases} \quad (2.45)$$

heißt *Gleichverteilung*.

Kenngrößen von Zufallsvariablen sind die zentralen Momente einer Verteilung. Für gleichverteilte Zufallsvariablen berechnen sich der Mittelwert und die Varianz zu (KREYSZIG 1979, S.89):

$$\mu = \frac{a+b}{2}, \quad \sigma^2 = \frac{(b-a)^2}{12}. \quad (2.46)$$

Kenngrößen für Intervalle sind ihre obere und untere Grenze bzw. Intervallmitte und -Radius, wie sie in (2.2) vorgestellt wurden. Ein erster Vergleich auf Basis dieser Kenngrößen zeigt, dass die Intervallmitte dem Erwartungswert der Zufallsvariablen entspricht. Intervallradius und Varianz entsprechen sich nicht, vergleiche hierzu auch die Diskussion in KUTTERER (1994, S.39).

Grundlegend sind jedoch die Konzepte: Mit Begriffen der Stochastik gesprochen wird im Intervallansatz nur mit dem Trager der Wahrscheinlichkeitsdichte, dem Intervall  $[a, b]$ , gearbeitet. Die Art und Charakteristika der Dichtefunktion selbst spielen hierfür keine Rolle. Bei der stochastischen Betrachtungsweise hingegen enthalt jedoch gerade sie die Information über die Kenngrößen (zentrale Momente). Als wichtigste Konsequenz ergeben sich ganz unterschiedliche Aussagen über die Variabilitat der betrachteten GröÙe: einmal als Varianz und einmal als Durchmesser des Tragerintervalls  $[a, b]$ .

Diese Unterschiede treten deutlich hervor, wenn man beispielsweise die Grundrechenarten untersucht. Die Menge der Intervalle ist diesbezügl. abgeschlossen, d.h. das Ergebnis der Verknüpfung ist wieder ein Intervall. Addiert man hingegen zwei gleichverteilte Zufallsvariablen, so ist das Ergebnis keine gleichverteilte, sondern eine trapezoidal verteilte Zufallsvariable, wie das folgende Beispiel veranschaulicht, vgl. PAPOULIS (1991, S.136ff).

**Beispiel 7** Betrachten wir zwei unabhängige gleichverteilte Zufallsvariablen  $x$  und  $y$  mit ihren Dichten:

$$\phi_x(x) = \begin{cases} \frac{1}{b-a}, & \text{für } a < x < b \\ 0, & \text{sonst,} \end{cases} \quad \phi_y(y) = \begin{cases} \frac{1}{d-c}, & \text{für } c < y < d \\ 0, & \text{sonst.} \end{cases} \quad (2.47)$$

Es gelte weiter  $a < c < b < d$ . Die Dichte der Zufallsvariable  $z = x + y$  erhalt man über Konvolution (PAPOULIS 1991, S.136) aus:

$$\phi_z(z) = \int_{-\infty}^{\infty} \phi_y(z-x)\phi_x(x) dx. \quad (2.48)$$

Da das Integral (2.48) nur Werte ungleich Null annimmt, wenn der Integrand größer Null ist, können in (2.48) die Grenzen des Integrationsbereichs auf das Intervall  $[a, b]$  eingeschrankt werden. Wird weiter  $v := z - x$  substituiert, so kann (2.48) aquivalent geschrieben werden:

$$\phi_z(z) = \frac{1}{b-a} \int_{z-b}^{z-a} \phi_y(v) dv. \quad (2.49)$$

Die Ergebnisvariable  $z$  variiert im Intervall  $[a+c, b+d]$ . Zur Auswertung des Integrals sind drei Teilintervalle zu untersuchen. Als Ergebnis erhalt man bei der vorgegebenen Ordnung  $a < c < b < d$  und für  $b-a < d-c$ :

$$\phi_z(z) = \begin{cases} \frac{1}{b-a} \frac{1}{d-c} (z - (a+c)) & \text{für } z \in [a+c, b+c], \\ \frac{1}{d-c} & \text{für } z \in [b+c, a+d], \\ \frac{1}{b-a} \frac{1}{d-c} ((b+d) - z) & \text{für } z \in [a+d, b+d], \\ 0 & \text{sonst.} \end{cases} \quad (2.50)$$

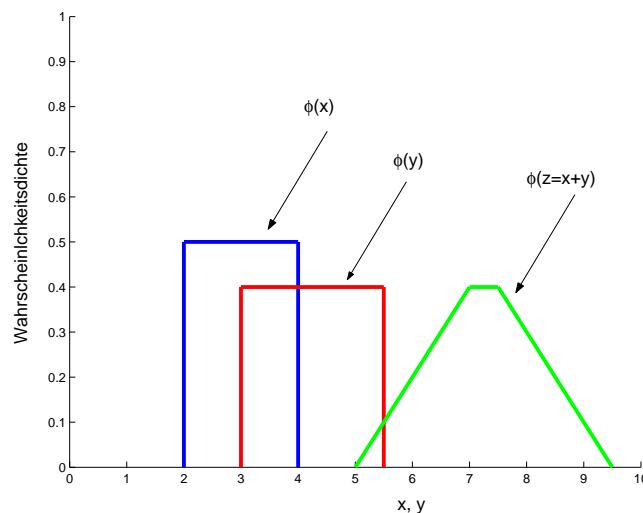


Abbildung 2.8: Darstellung der Wahrscheinlichkeitsdichten für gleichverteilte Zufallsvariablen

Abbildung 2.8 zeigt die Ausgangsdichten und die resultierende Wahrscheinlichkeitsdichte, die eine trapezoidale Verteilung und keine Gleichverteilung mehr ist. Zur Übersichtlichkeit wurden nur die von Null verschiedenen Funktionswerte dargestellt. Varianz und Mittelwert ergeben sich mit den entsprechenden Additionssätzen (KREYSZIG 1979, S.153ff):

$$\mu = \frac{a + b + c + d}{2}, \quad \sigma^2 = \frac{(b - a)^2 + (d - c)^2}{12}. \quad (2.51)$$

Abschließend lässt sich feststellen: Die Analogien, die auf den ersten Blick vorhanden scheinen, dürfen nicht darüber hinwegtäuschen, dass Intervalle und gleichverteilte Zufallsvariablen vom Konzept her grundverschieden sind. Ihre Anwendungen führen somit jeweils zu unterschiedlichen Interpretationen der Ergebnisse, einmal trägerbezogen und einmal bezogen auf die zentralen Momente der Verteilung.

## 2.3 Bestimmung der Impräzision korrigierter Beobachtungen

### 2.3.1 Vorbemerkungen

Ziel des Aufbereitungsprozesses der Messung ist die möglichst vollständige Reduktion der *Abweichungen* oder Diskrepanzen zwischen gemessenen Daten  $\ell$  und ihrer Modellierung im geodätischen Modell  $l$ . Diese lassen sich in einen deterministischen Anteil  $\delta$  und einen zufälligen Anteil  $\epsilon$  unterteilen:

$$\ell - l = \delta + \epsilon. \quad (2.52)$$

Die folgende Betrachtung konzentriert sich auf die deterministische Komponente  $\delta$ , wobei der stochastische Anteil  $\epsilon$  vorerst ausgeklammert wird. Die Abweichungen  $\delta$  setzen sich aus verschiedenen Effekten mit bekannten oder unbekanntem Charakteristika und Magnituden zusammen. So wird zur Beschreibung aller geometrischer Messungen (Richtungen, Strecken, Zenitdistanzen oder Raumstrecken) im geodätischen Modell die euklidische Geometrie zu Grunde gelegt. Abweichungen treten beispielsweise durch Refraktion im Ausbreitungsmedium oder durch die Gestalt des Schwerfeldes (Krümmung der Lotlinien) auf.

Neben solchen äußeren Effekten spielen auch Einflüsse der Instrumente eine Rolle. Bezugspunkte für die Modellierung der Messung sind die Sensorreferenzpunkte wie beispielsweise der Achsschnittpunkt aus Ziel-, Kipp- und Stehachse bei Totalstationen. Treten Instrumentenfehler auf, so kann dieser Referenzpunkt nicht realisiert werden. Bei GPS-Antennen ist das Antennenphasenzentrum der Referenzpunkt, wobei die Variation des Antennenphasenzentrums in Abhängigkeit von Elevation, Azimut und Messfrequenz berücksichtigt werden muss, vgl. beispielsweise ROTHACHER ET AL. (1996). Ebenso spielt die zeitliche Stabilität der Messfrequenz eine Rolle. Bei der elektronischen Distanzmessung (EDM) können Abweichungen durch Temperaturveränderungen (Erwärmungseffekte) oder Alterungsprozessen der Quarzoszillatoren auftreten, vgl. RÜEGER (1996, S.217ff) oder JOECKEL UND STÖBER (1999, S.82ff). Bei GPS-Messungen werden diese Effekte durch Uhrkorrektur modelliert. Die Abweichung  $\delta$  beschreibt somit den integralen Effekt, der sich durch Überlagerung aller Einzeleffekte ergibt.

Zur Reduktion oder Elimination lassen sich zwei Ansätze unterscheiden: das Anbringen von *Korrekturen* oder das Anwenden von *Messmethoden*. Unter einer Messmethode soll eine zur Reduktion geeignete Verknüpfung von Messungen verstanden werden. Zusammenfassend lässt sich der gesamte Aufbereitungsprozess als Abfolge verschiedener Korrekturen und Anwendung von Messmethoden auffassen. Hierauf wird im nächsten Abschnitt detailliert eingegangen. Als Resultat erhält man die *korrigierte Messung*, die als Eingangsgröße in die Ausgleichung eingeht. Die in ihr enthaltene Information ermöglicht die Schätzung der Parameter, enthaltene Restabweichungen verzerren die Schätzwerte.

Die Grundideen für alle weiteren Betrachtungen lassen sich in folgenden drei Punkten zusammenfassen:

- Die Abweichung zwischen Modell und Realität setzt sich aus verschiedenen Effekten wie den angesprochenen instrumentellen oder atmosphärischen Einflüssen zusammen.
- In der Regel bedient man sich bekannter Modelle und Ansätze aus der Physik, um formelmäßige Zusammenhänge zur Beschreibung und Erklärung der Effekte bereitzustellen. Zur Berechnung der Korrektur sind weitere Informationen notwendig wie konkrete Werte von Modellkonstanten oder zusätzliche Messungen wie beispielsweise Temperatur oder Druck. Prinzipiell muss jedoch festgehalten werden, dass nicht jeder Effekt in seiner Gesamtheit modelliert werden kann und daher *nichtmodellierbare Anteile* verbleiben.

- Die angewandte Korrektionsformel ist eine Vereinfachung grundlegender Gleichungen wie der Maxwell'schen Differentialgleichungen. Es lässt sich eine Hierarchie der Korrekturen angeben. Dabei wird postuliert, dass eine allgemeinere Form die Realität besser approximieren kann und der Anteil der nichtmodellierbaren Effekte abnimmt. Die Hierarchie ist nicht eindeutig und allgemein anerkannt. Es können auch durchaus mehrere Modelle für Korrekturen auf gleicher Stufe miteinander konkurrieren.

### 2.3.2 Korrekturen und Messmethoden - Ein formalisierender Ansatz

In diesem Abschnitt werden Korrekturen und Messmethoden durch mathematische Operationen formalisiert. Dieser Schritt mag abstrakt erscheinen. Er liefert aber das Handwerkszeug, das an späterer Stelle eine übersichtliche Betrachtung der Unsicherheitsfortpflanzung durch die Aufbereitungsprozesse bis zur korrigierten Messung gewährleistet. Konkrete Beispiele werden dieses Konzept in Abschnitt 3.1 veranschaulichen.

Voraussetzung für eine gezielte Reduktion von Effekten ist, wie in den vorherigen Abschnitten angesprochen, ein Zugang zu ihren wesentlichen Charakteristika. Dieser kann durch die Modelle der Physik geliefert werden, die ihrerseits primär auf eine Untersuchung und Beschreibung der Effekte zielen. Die Korrektur ist eine Funktion von verschiedenen Einflussparametern  $p_i$ :

$$k = f(\mathbf{p}). \quad (2.53)$$

Diese können beispielsweise rohe oder geräteintern korrigierte Messwerte selbst, zusätzliche (nichtgeodätische) Messungen wie Temperatur oder Druck und Informationen wie Modellannahmen und Modellkonstanten sein, die das Messmodell zur Verfügung stellt. Alle diese Werte sollen als Einflussparameter angesehen und in einem Vektor  $\mathbf{p}$  zusammengefasst werden. Es ist sinnvoll, die einzelnen Parameter zu klassifizieren.

- Eine erste Gruppe bilden die *Modellparameter*  $\mathbf{p}_M$ , die vom physikalischen Modell selbst abhängen, das die Korrektur beschreibt. Beispiele hierfür sind die Modellkonstanten.
- Davon sind die *Sensorparameter*  $\mathbf{p}_S$  zu unterscheiden, die die Eigenschaften des jeweils verwendeten Messinstruments kennzeichnen, wie technische Daten. Mittels dieser Parameter können Instrumente in verschiedene Leistungsklassen eingeteilt werden.
- Schließlich sollen *zusätzliche Messungen* als Parameter  $\mathbf{p}_Z$  aufgefasst werden, die beispielsweise zur Erfassung der aktuellen meteorologischen Bedingungen dienen, wie Temperatur, Druck oder Feuchte.

Die Parameter seien in der folgenden Reihenfolge im Vektor  $\mathbf{p}$  geordnet:

$$\mathbf{p} = (\mathbf{p}_M, \mathbf{p}_Z, \mathbf{p}_S)^T. \quad (2.54)$$

Betrachtet man die Aufbereitungsprozesse für Messungen unterschiedlichen Typs wie Richtungen, Strecken, Zenitdistanzen, GPS-Messungen, etc., so unterscheiden sich die Sensorparameter, die den Sensor an sich charakterisieren, von Messungstyp zu Messungstyp. Werden ähnliche Korrektionsmodelle bei verschiedenen Messungstypen verwendet, so können die gleichen Modellparameter in Korrekturen für verschiedene Messungstypen vorkommen. Die atmosphärischen Bedingungen beeinflussen alle Messungstypen, so dass die zusätzlichen Messungen für alle Messungstypen übergreifend die Umgebungsbedingungen beschreiben.

Umgekehrt verändert der Einsatz eines anderen Messinstruments nur die jeweiligen Sensorparameter. Eine Verbesserung der Bestimmung der atmosphärischen Bedingungen kommt dagegen allen dadurch beeinflussten Messungen zugute. Diese Argumentation wird bei der Analyse der Beiträge zur Gesamtunsicherheit wieder aufgegriffen und für eine gezielte Reduktion vertieft, vgl. Abschnitte 3.2 und 3.3 bzw. 5.4.

Neben den Parametern können die Korrekturen ebenfalls gruppiert werden. In Abschnitt 2.2.1 wurden Korrekturen zur Abstimmung von Modellvoraussetzungen eingeführt. Dementsprechend kann man Korrekturen in drei Gruppen einteilen:

- *Geräteinterne Korrekturen* beschreiben die Abweichungen zwischen dem Messmodell und dessen Realisierung in der aktuellen Geometrie bzw. Physik des Sensors. Diese Unterschiede werden beispielsweise durch Instrumentenfehler oder Instabilität von Messfrequenzen oder Uhren hervorgerufen. Oftmals werden geräteintern besondere Bezugswerte angenommen, die sich von den Standardwerten in der Physik unterscheiden (Normalatmosphäre). Diese Korrekturen werden als geräteinterne Korrekturen für alle Messungen im Vektor  $\mathbf{k}_S$  zusammengefasst.

- Die Abweichung zwischen den Modellvoraussetzungen des Messmodells und den realen Bedingungen werden durch die Korrekturen im Vektor  $\mathbf{k}_P$  betrachtet oder zumindest angenähert. Hierunter wird beispielsweise die Berücksichtigung der Refraktion als wesentlicher Faktor verstanden.
- Schließlich müssen die Modellvoraussetzungen des geodätischen Modells den realen Bedingungen angepasst werden. Darunter fällt beispielsweise die Wahl der Rechenbezugsfläche (Ebene, Kugel oder Ellipsoid) oder des Schwerefeldmodells. Letzteres charakterisiert die Form der Lotlinien, wobei zwischen parallelen oder zentralsymmetrischen Lotlinien, Lotlinien in Form von Ellipsoidnormalen eines Referenzellipsoids des Normalschwerefeldes oder doppelt gekrümmten Lotlinien unterschieden werden kann. Diese Information wird unter anderem für die Verknüpfung von Instrumentenreferenzpunkten mit Bodenpunkten oder definierten Stationspunkten verwendet. Die Korrekturen werden als Vektor  $\mathbf{k}_G$  zusammengefasst.

Neben dem direkten Anbringen der Korrekturen können die Charakteristika der Effekte genutzt werden, um Messanordnungen zu konstruieren, mit deren Hilfe sich systematisch wirkende Effekte bestimmen lassen und ihr Einfluss auf die Messung rechnerisch beseitigt werden kann. Grundidee aller Operationen ist die *Ausnutzung von Symmetrien* der systematisch wirkenden Effekte.

**Beispiel 8** Eine instrumentelle Einrichtung unter Ausnutzung von Symmetrien ist die Verwendung zweier diametral gelegener Teilkreisabgriffstellen bei der Richtungsbeobachtung. Hierbei werden Teilkreisexzentrizitäten geräteintern durch Mittelung der beiden Abgriffe eliminiert (COOPER 1987, S.153).

Treten Driften in Instrumenten auf, so können diese entweder durch Abgleich mit geräteinternen Vergleichsstrecken wie bei der EDM (innerer Lichtweg) oder durch die besondere Anlage der Messung reduziert werden, wenn eine geräteinterne Korrektur nicht direkt möglich ist.

**Beispiel 9** Betrachten wir die Gravimetermessung mit Federgravimetern. Diese Instrumente weisen eine deutliche nichtkonstante Drift der Federlänge auf. Es hat sich deshalb durchgesetzt, die Messanordnung so zu wählen, dass bei der Auswertung ein Driftpolynom geschätzt werden kann, vgl. TORGE (2001, S.181ff). Ein Beispiel hierzu ist das *Stepverfahren*, bei dem die Messungen in folgender Reihenfolge auf den verschiedenen Punkten durchgeführt werden: 1 – 2 – 1 – 2 – 3 – 2 – 3 – 4 . . . . Beim *Sternverfahren* lautet die Reihenfolge der Messungen 1 – 2 – 1 – 3 – 1 – 4 – 1 . . . . Werden die Messungen auf den einzelnen Punkten in ungefähr gleichen Zeitabständen ausgeführt, so ergibt sich nach der Differenzbildung bei linearer Drift ein konstanter Resteffekt. Zur Behandlung nichtlinearer Anteile können Koeffizienten höherer Ordnung für Driftpolynome geschätzt werden.

Verschiedene Effekte in der Umwelt, die sich beispielsweise durch Jahres- oder Tagesperioden kennzeichnen, weisen ebenfalls zeitliche Symmetrien auf.

**Beispiel 10** Bei der Anordnung der Richtungsmessung als Vollsatz werden zuerst Messungen aller Ziele in Fernrohrlage I und dann im Gegensinn in Lage II durchgeführt. Werden alle Beobachtungen in zeitlich gleichen Abständen gemessen und findet durch Sonneneinstrahlung bedingt eine lineare Torsion des Messpfeilers oder Stativs statt, auf dem das Instrument aufgebaut ist, so können aus der Mittelung der jeweils korrespondierenden Messungen in Lage I und II die Effekte auf einen konstanten Anteil reduziert werden. Diese fallen dann bei Reduktion auf ein gemeinsames Ziel heraus.

Wird die zeitliche Abfolge der Messungen erhöht, um beispielweise bewegte Objekte zu verfolgen oder zu leiten wie bei der Fahrzeugnavigation oder Baumaschinensteuerung, so können in der Regel keine Doppelmessungen oder Vergleichsmessungen durchgeführt werden. Die Reduktion von Effekten, die bisher von Messmethoden geleistet wurde, muss dann durch Korrekturen ermöglicht werden. Zudem treten weitere Instrumentenfehler auf, die bei den bisherigen statischen Messungen keinen Einfluss hatten (STAIGER 1998; HENNES 1999; HENNES UND INGENSAND 2000; HENNES UND KRICKEL 2000). HENNES (1999) nennt beispielsweise die nicht zeitgleich durchgeführte Messung der drei Polarelemente Richtung, Strecke, Zenitdistanz, die Problematik der Definition und Konstanz des Zeitbezuges einer Messung oder das weniger häufige Messen der geräteinternen Referenzstrecke, so dass sich Drifteffekte auswirken.

Im Folgenden werden exemplarisch einige wichtige Messmethoden, die in der Geodäsie zur Anwendung kommen, vorgestellt und durch Matrizen formalisiert.

**Differenzbildung:** Es ist anzunehmen, dass reale Vorgänge und Zustände wie atmosphärische Bedingungen zeitlich oder räumlich benachbart durchgeführte Messungen ähnlich beeinflussen (positive Korrelationen).

Dies gilt auch für Beobachtungen, die mit dem gleichen Instrumentarium gemessen werden. Die zu erwartenden Effekte werden sich durch gleiches Vorzeichen und ungefähr gleich großen Betrag auszeichnen. Ihr Einfluss wird reduziert, indem die Differenz beider Messungen gebildet wird. Die Größe des *Resteffektes* richtet sich danach, wie gut die Annahme der gleichen Magnitude auch tatsächlich zutrifft. Die Matrix der Messmethode wird formalisiert durch

$$\mathbf{M} = \begin{pmatrix} 1 & -1 \end{pmatrix}. \quad (2.55)$$

Mit den in (2.52) eingeführten Bezeichnungen lässt sich die korrigierte Messung bezüglich der deterministischen Abweichung darstellen als

$$l = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix} - \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = \ell_1 - \ell_2 + \delta_2 - \delta_1. \quad (2.56)$$

Für  $\delta_1 = \delta_2$  wird die Abweichung eliminiert. Gilt  $\delta_2 \neq \delta_1$ , so kann die Abweichung nicht vollständig eliminiert werden und es verbleiben Resteffekte. Einige Beispiele sollen die konkreten Anwendungen aufzeigen.

**Beispiel 11** Betrachten wir in einem ersten Beispiel GPS-Messungen näher. Pro Satellit-Empfänger Konfiguration liegen bis zu vier Messelemente vor: Code- und Phasenmessung auf den Frequenzen  $L_1$  und  $L_2$ . Durch Differenzbildung der Phasenmessungen auf beiden Frequenzen werden die geometrischen Anteile (Pseudostrecke Empfänger-Satellit) und damit die Orbitfehler eliminiert. Man spricht von der *geometriefreien Linearkombination*  $L_4$ , vgl. beispielsweise ROTHACHER UND SCHÄR (1995) oder HOFMANN-WELLENHOF ET AL. (2001, S.92ff).

Werden auf einer Station Signale von zwei Satelliten empfangen, so können durch deren Differenzbildung (*Einfachdifferenz*) instrumentelle Effekte der Empfänger wie Uhrfehler eliminiert bzw. reduziert werden. Die entsprechende umgekehrte Konfiguration erlaubt Uhrfehler des Satelliten zu eliminieren. Voraussetzung ist eine ausreichende Zeitsynchronisierung der Uhren. ROTHACHER UND SCHÄR (1995, S.108) geben  $1 \mu\text{s}$  für  $1 \text{ mm}$  radialen Fehleranteil an. Eine Koppelung der Einfachdifferenzen in *Doppel- und Tripeldifferenzen* erlaubt, die instrumentellen Anteile von Satellit und Empfänger zu eliminieren. Ein wesentlicher Nachteil der Differenzbildung insbesondere bei unkorrelierten Beobachtungen ist das Anwachsen des Messrauschens um ca. 41 % pro Differenzbildung. Hierauf wird noch in Abschnitt 2.3.6 detailliert eingegangen.

**Beispiel 12** Die Standard-Messmethode beim Nivellement (Rückblick - Vorblickablesung) stellt an sich schon eine Differenzbildung dar. Neben der Bestimmung des Höhenunterschieds werden aber auch instrumentelle Effekte wie Kollimationsfehler und Kompensatoreinspielfehler und die symmetrischen Anteile der Refraktion reduziert (RÜEGER UND BRUNNER 2000). Da diese instrumentellen Effekte zielweitenabhängig sind, wird die beste Reduktion bei Messung aus der Mitte erzielt.

**Mittlung:** Treten in beiden Messungen die gleichen Effekte additiv aber mit entgegengesetztem Vorzeichen auf, so wird ihr Einfluss reduziert, indem das Mittel aus beiden Messungen genommen wird. Die Matrix der Messmethode ist durch

$$\mathbf{M} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad (2.57)$$

gegeben und die korrigierte Messung lässt sich schreiben zu:

$$l = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = \frac{1}{2}(\ell_1 + \ell_2) - \frac{1}{2}(\delta_1 + \delta_2). \quad (2.58)$$

Für  $\delta_1 = -\delta_2$  kann die Abweichung eliminiert werden. Liegen die Effekte mit gleichem Vorzeichen vor, so können diese durch eine Mittlung nicht reduziert werden.

**Beispiel 13** Betrachten wir die Messung in zwei Lagen bei Richtungsbeobachtungen. Bei modernen Messgeräten wird für die instrumentellen Korrekturen folgendes einfaches Modell angenommen (KAHMEN 1993, S.136 (3.10)):

$$r_A = \rho + \frac{c_h}{\sin \zeta} + (Q(\rho) + i) \cot \zeta. \quad (2.59)$$

Dabei bezeichnet  $\rho$  den rohen Horizontalteilkreisabgriff,  $\zeta$  die Zenitdistanz und  $Q(\rho)$  die mit dem Zweiachsenkompensator bestimmte Stehachsschiefe in Kippachsrichtung. Kippachsfehler  $i$  und Zielachsfehler  $c_v$

treten bei Messungen in zwei Lagen aufgrund der Symmetrien bezüglich der Steh- und Kippachse mit unterschiedlichen Vorzeichen aber gleichgroßem Betrag auf, so dass sich bei der Differenzbildung aus beiden Lagen der Effekt herausmittelt. Auch wenn sich (2.59) als Standardmodell etabliert hat, so werden nur Effekte erster Ordnung berücksichtigt. Effekte höherer Ordnung (Taufehler) oder insbesondere das gleichzeitige Auftreten mehrerer Instrumentenfehler bleibt unberücksichtigt, vgl. MATTHIAS (1961). Für sehr steile Visuren von  $\zeta = 10 \text{ gon}$  betragen die Unterschiede in den Formeln ca  $0.1 \text{ mgon}$  (STRAHLBERG 1997).

**Quotientenbildung:** Die bisher vorgestellten additiven Verknüpfungen der Effekte mit der Messung dürften in vielen Fällen auf Approximationen oder Linearisierungen zurückzuführen sein. So kann beispielsweise eine multiplikative Verknüpfung durch eine additive ausgedrückt werden

$$l = \delta l = l - (\delta + 1)l := l + \check{\delta}. \quad (2.60)$$

Eine vollständige Elimination der Effekte ist jedoch nur bei der Quotientenbildung möglich. Die Matrix der Messmethode ist

$$\mathbf{M} = \begin{pmatrix} \frac{1}{l_2} & 0 \end{pmatrix} \quad (2.61)$$

und die korrigierte Beobachtung ergibt sich zu

$$l = \frac{\delta_1 l_1}{\delta_2 l_2}. \quad (2.62)$$

Für  $\delta_1 = \delta_2$  ist die korrigierte Messung frei vom Einfluss der Abweichung.

**Beispiel 14** Die Bildung von Streckenverhältnissen erlaubt die Reduktion von Refraktionseinflüssen erster Ordnung auf die Streckenmessung, vgl. JÄGER (1985). Durch die „erste Geschwindigkeitsreduktion“ wird der geräteinterne Brechungsindex  $n_0$  durch den mittleren repräsentativen Brechungsindex  $n_A$  ersetzt (RÜEGER 1996, S.74):

$$S = \frac{n_0}{n_A} S_0. \quad (2.63)$$

Das Verhältnis zweier Strecken, für die der mittlere Brechungsindex übereinstimmt, sind frei vom Einfluss der aktuellen Refraktion. HECK UND JÄGER (1986) zeigen die Vorteile dieses Ansatzes bei der Messung und Auswertung von Überwachungsnetzen auf.

**Linearkombinationen:** Die bisher betrachteten Methoden können als Spezialfälle der Linearkombination von Messungen aufgefasst werden. Linearkombinationen erlauben, Unterschiede der Größenordnung der Effekte mit zu berücksichtigen. Die Matrix der Messmethode ist

$$\mathbf{M} = \begin{pmatrix} m_1 & m_2 \end{pmatrix}. \quad (2.64)$$

**Beispiel 15** Betrachten wir die bei der Auswertung von GPS-Messungen häufig gewählte ionosphärenfreie Linearkombination näher (ROTHACHER UND SCHÄR 1995):  $L_3 := \frac{f_1^2}{f_1^2 - f_2^2} L_1 - \frac{f_2^2}{f_1^2 - f_2^2} L_2$ . Grundannahme für die Effizienz dieser Linearkombination ist, dass sich die Ionosphäre durch das einfache Modell beschreiben lässt:

$$\delta_{Ion} = -\frac{40.30}{f^2} TEC, \quad (2.65)$$

wobei  $TEC$  den gesamten Elektronengehalt angibt. Typische Werte liegen zwischen  $TEC = 10^{16} \dots 10^{18} \frac{e}{m^2}$ . Die Messfrequenz wird mit  $f$  bezeichnet. Für die beiden Trägersignale gilt:  $f_1 = 1575.42 \text{ MHz}$ ,  $f_2 = 1227.60 \text{ MHz}$ . KLOBUCHAR (1996, S.488ff) stellt die Vereinfachungsschritte zusammen, die nötig sind, um die Form (2.65) zu erhalten. Lässt sich das Verhalten der Ionosphäre in diesem einfachen Modell nicht realistisch beschreiben, bleiben mehr oder minder große Resteffekte.

**Kombination von Messmethoden:** In vielen Anwendungen kommen die „Grundformen“ der Messmethoden nicht alleine vor, sondern als Kombination. Dies kann durch die Verkettung der entsprechenden Matrizen ausgedrückt werden.



**Beispiel 16** Es soll nochmals die klassische Richtungsbeobachtung in Vollsätzen als einfaches und anschauliches Beispiel herangezogen werden. Hierzu werden mehrere Sätze in der oben beschriebenen Anordnung gemessen. Durch Mittelung der Lage I und II werden, wie im Beispiel 10 beschrieben, die Einflüsse der Instrumentenfehler eliminiert. Ebenso wird der Einfluss von Torsionen des Messpfeilers, auf dem das Instrument aufgebaut ist, auf einen konstanten Anteil reduziert. Durch die Reduktion auf das erste Ziel werden diese Anteile eliminiert. Dabei ist es wichtig, dass das erste Ziel möglichst im Horizont liegt und die entsprechende Richtungsmessung somit frei von Einflüssen des Stehachsfehlers ist. Bei einer fehlerbehafteten Referenzrichtung überträgt sich dieser Betrag bei der Differenzbildung auf alle anderen Richtungen. Ein Verdrehen des Horizontalteilkreises zwischen den einzelnen Sätzen verteilt - sofern das Verdrehen bei modernen Totalstationen überhaupt noch möglich ist - die Ablesungen über den gesamten Teilkreis. Dies entspricht einer Variation des Messungsrahmens für die Teilkreisablesung, wodurch Teilkreisfehler zufällig werden. Bei einer Mittelung aller Sätze ist somit zu erwarten, dass der Teilkreisfehler sich zu Null herausmittelt.

Zusammenfassend kann der Mess- und Aufbereitungsprozess als Verknüpfung von Korrekturen und Messmethoden aufgefasst werden. Dies soll formal wie folgt angegeben werden:

$$\begin{aligned}
 \mathbf{l} &= \mathbf{M}(\mathbf{M}_S \ell \circ \mathbf{k}_S \circ \mathbf{k}_P) \circ \mathbf{k}_G \\
 &= \underbrace{\mathbf{M}(\mathbf{l}_A \circ \mathbf{k}_P)}_{\text{Modell der Physik der Messung}} \circ \underbrace{\mathbf{k}_G}_{\text{Anpassung an das geodätische Modell}}.
 \end{aligned} \tag{2.66}$$

Der Vektor der Messwerte, die am Messgerät angezeigt werden, wird mit  $\mathbf{l}_A$  bezeichnet. Bei modernen Vermessungsinstrumenten ist dieser Wert bereits das Ergebnis einer geräteinternen Auswertung, die durch den Ausdruck  $\mathbf{M}_S \ell \circ \mathbf{k}_S$  angedeutet wird. Dies soll kurz skizziert werden. Beim Auslösen einer Messung wird in der Regel geräteintern eine Messreihe ermittelt. Bei der Streckenmessung sind dies 1000-2000 Einzelmessungen (JOECKEL UND STOBER 1999, S.85), beim Teilkreisabgriff mit dynamischer Abgriffmethode 512 Einzelabgriffe pro Umlauf an diametralen Teilkreisstellen (JOECKEL UND STOBER 1999, S.213ff). Diese Einzelmessungen werden mittels Filterungen oder durch einfache Median- oder Mittelbildung ausgewertet. Als Messwert wird nur das Ergebnis dieses Schritts angezeigt, das um zusätzliche, geräteinterne Korrekturen  $\mathbf{k}_S$  verbessert wurde. Die genauen Modelle und Auswerteschritte in der Gerätesoftware sind allerdings nur bedingt zugänglich, vgl. beispielsweise (INGENSAND 2001). Folglich kann nur die Unsicherheit des Anzeigewertes  $\mathbf{l}_A$  abgeschätzt werden. Diese Idee wird im Problembereich der Kalibrierung geodätischer Instrumente durch das Konzept der Systemkalibrierung umgesetzt (BRUNNER UND WOSCHITZ 2001), worauf im nächsten Abschnitt noch eingegangen werden soll.

### 2.3.3 Impräzision der Einflussparameter

In den beiden vorangegangenen Abschnitten wurden die Grundideen und eine Formalisierung der Aufbereitungsprozesses für die Messung vorgestellt. In diesem Abschnitt wird gezeigt, wie Intervalle als Maße für Impräzision konkret bestimmt werden können.

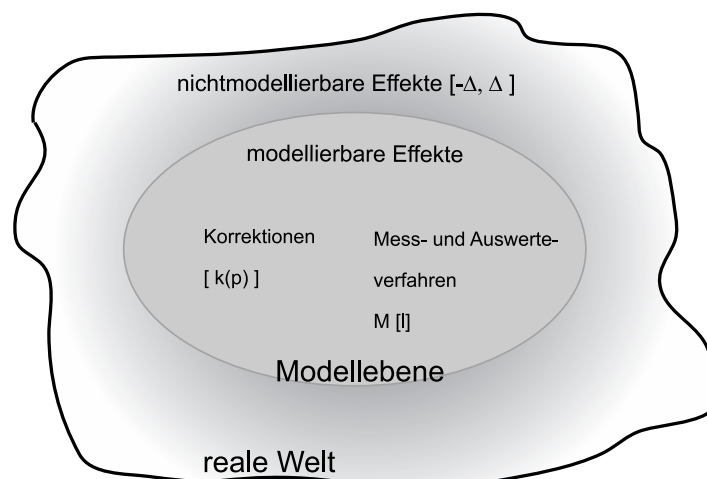


Abbildung 2.9: Ansatzpunkte für Impräzision

Da die korrigierten Messungen funktional von den Einflussparametern  $\mathbf{p}$  abhängen, lässt sich die Betrachtung der Impräzision auf diese Ausgangsgrößen zurückführen. Es sollen für die Einflussparameter obere und untere Schranken abgeleitet werden, innerhalb derer die konkreten Werte, die in die Korrekturen eingesetzt werden, *maximal* variieren können. Diese maximale Variationsmöglichkeit der Werte drückt das unzureichende Wissen über den konkreten Wert, also Impräzision aus.

Abbildung 2.9 veranschaulicht zwei Gesichtspunkte, die für die Beschreibung von Impräzision zu berücksichtigen sind. Beschränkt man sich auf die Ebene der Modelle, so führt man eine *innere Betrachtung* durch. Bezieht man jedoch die Realität mit ein, so zielt man auf eine *äußere Betrachtung*. Die Trennung in modellierbare und nichtmodellierbare Effekte verdeutlicht die Grenze der Beschreibbarkeit, die je nach Aufgabenstellung fließend ist.

### Ansatzpunkte auf der Modellebene

Bleiben wir zunächst auf der Modellebene. Ein Ansatzpunkt für die Impräzision ergibt sich bei näherer Betrachtung der Informationen (Modellkonstanten, zusätzliche Messungen), die zur Anwendung eines Korrektionsmodells nötig sind.

Ein wesentlicher Anteil eines Modells sind die *Modellkonstanten*. Diese sollen konventionsgemäß nur mit *signifikanten Ziffern* dargestellt werden. Als Aussage leitet sich daraus ab, dass der Wert der Modellkonstanten auf  $\pm 0.5 \times$  Einheit der letzten angegebenen Stelle signifikant ist, vgl. TIPLER (1994, S.6ff). Wird also davon ausgegangen, dass die Konstanten in den gängigen Korrektionsmodellen „physikalisch“ korrekt angegeben sind, so kann die Impräzision zur Hälfte der Einheit der letzten signifikanten Stelle abgeschätzt werden. Für manche Konstanten werden auch explizit andere Fehlerbänder angegeben, vgl. GROTEN (2000).

Für *zusätzliche Messungen* wie Temperatur und Druck gilt in der Regel, dass aus Kostengründen nur wenige Messungen ausgeführt werden oder sogar auf externe Angaben wie Daten von Wetterstationen des Wetterdienstes in der Nähe des Messgebiets zurückgegriffen wird, vgl. RÜEGER (1996, S.60). Werden im Rahmen der geodätischen Beobachtungen zusätzliche Messungen ausgeführt, so können aufgrund der geringen Stichprobe (vielleicht 3 Ablesungen) statistische Kenngrößen nicht aussagekräftig verwendet werden. Mittlere Effekte, die sich bei großen Wiederholzahlen ergeben würden, bilden sich nicht aus. Ebenso kann nicht davon ausgegangen werden, dass die Messungen normalverteilt sind. Auf jeden Fall kann eine entsprechende Hypothese wegen des geringen Stichprobenumfangs nicht getestet werden und entsprechende Grenzwertsätze der Stochastik für beliebige verteilte Zufallsvariablen (zentraler Grenzwertsatz, KREYSZIG (1979, Satz 76.1, S.200ff)) gelten nicht. In solchen Fällen bietet es sich an, die Impräzision sinnvoll abzuschätzen. Neben der *maximalen Abweichung zwischen Messwerten* kann auch die *Auflösung* des verwendeten Sensors angesetzt werden, die in den Handbüchern der Messinstrumente angegeben wird.

Ein weiterer Ansatzpunkt entsteht bei der Bewertung des Sensoroutputs geodätischer Instrumente. Diese setzen sich aus einer Vielzahl von Einzelsensoren zusammen. Man erhält in der Regel nur integrale Angaben für die Unsicherheit des Outputs, wobei die exakte Bestimmung dieser Werte seitens der Instrumentenhersteller nicht immer offengelegt ist, vergleiche beispielsweise RÜEGER (1996, S.220ff) oder BRUNNER UND WOSCHITZ (2001). In der geodätischen Messtechnik wird unter *Kalibrieren* die Bestimmung eines (funktionalen) Zusammenhangs zwischen angezeigten und tatsächlichen Messwerten verstanden (HEISTER UND STAIGER 2001). Für signifikante Abweichungen werden dann Korrekturfunktionen bestimmt. Sind die einzelnen Bauteile und Auswerteschritte nicht zugänglich, so kann nur das gesamte System kalibriert werden (*Systemkalibrierung*), vgl. beispielsweise HENNES UND INGENSAND (2000) oder BRUNNER UND WOSCHITZ (2001). Hierbei ist eine detaillierte Analyse einzelner Komponenten nicht möglich. Die *Komponentenkalibrierung* hingegen war lange Zeit state-of-the-art der geodätischen Geräteuntersuchung deren Ergebnisse sich in vielen Lehrbüchern niedergeschlagen haben, vgl. beispielsweise COOPER (1987), RÜEGER (1996), JOECKEL UND STÖBER (1999) oder DEUMLICH UND STAIGER (2002). Aktuell scheint noch kein Konsens zu bestehen, ob eine Systemkalibrierung einer Komponentenkalibrierung vorzuziehen sei. Einzelne Standpunkte und Argumentationen lassen sich beispielsweise im Tagungsband (HEISTER UND STAIGER 2001) nachlesen. Intervalle bieten in diesem Kontext eine geeignete Möglichkeit, Schranken für die Variabilität des Sensoroutputs bei geodätischen Messungen anzugeben.

Besteht jedoch Zugang zu allen Komponenten wie beispielsweise bei Eigenentwicklungen, so kann das Messsystem in diese zerlegt und die Impräzision des Sensoroutputs aus Tests oder Instrumentenspezifikationen abgeleitet werden. Solche Untersuchungen wurden beispielsweise von REMBE ET AL. (1999) für CCD-Kameras im Nanotechnikbereich oder von KIEFFER ET AL. (2000) für Ultraschallsensoren zur Robotersteuerung vorgenommen.

An dieser Stelle sei eingefügt, dass starre Intervallgrenzen mit einer gleichwertigen Zugehörigkeit aller Werte des Intervalls gegebenenfalls dem Wissen um die Physik der untersuchten Größe nicht gerecht werden. Hier können die Intervalle als Impräzisionsmaße durch Fuzzy-Sets verfeinert werden. VIERTL (1996, S.23ff) schlägt vor, Intensitäts- oder Grauwertverteilungen direkt als Zugehörigkeitsfunktionen zu interpretieren. Ein solches Vorgehen bietet sich beispielsweise zur Beschreibung des effektiven Wellenlängenbereichs einer Diode bei der elektronischen Distanzmessung (EDM) an. RÜEGER (1996, S.37, S.54) gibt den Variationsbereich mit  $\pm 10 \text{ nm}$  bezüglich der Nennwellenlänge der Diode an. Durch eine Intensitätsbetrachtung der Diode kann eine Zugehörigkeitsfunktion bestimmt werden. Somit wird im Grundsatz am Konzept der Intervalle als Impräzisionsmaße festgehalten, dieses aber als Fuzzy-Sets verfeinert, um der Physik besser Rechnung zu tragen.

Ein weiterer Ansatzpunkt für die Bestimmung der Impräzision der Einflussparameter ist das Ziel, *maximale Variationen* abzuschätzen. In anderen Fachdisziplinen (Fahrzeugbau, Elektrooptik, Verfahrenstechnik) werden diese Ansätze bereits verwendet, vgl. COLDITZ (1997) oder REMBE ET AL. (1999). Weitere Darstellungen sind unter den Stichworten *robust control* und *guaranteed estimation* zu finden. Durch die Verwendung von Intervallen kann garantiert werden, dass sich das Ergebnis, im Fall der Geodäsie die korrigierte Messung, innerhalb vorgegebener Schranken bewegt (*direkte Fragestellung*). Die *inverse Fragestellung*, aus der Vorgabe von Ergebnisintervallen maximale Intervalle der Einflussparameter abzuleiten, zielt auf ein Design der Beobachtungsmethoden und der verwendeten Sensoren, vgl. BRAEMS ET AL. (2003). Kleine Variationsbereiche für den Sensoroutput erfordern eine hohe Performance der Sensoren. Da damit hohe Kosten für den Sensor und für den gesamten Versuch bzw. die Messung verbunden sind, sind diese Analysen von Bedeutung. Solche Fragestellungen könnten gut in den Kontext der Geräteuntersuchungen integriert werden. BRAEMS (2002, Kap.6) stellt in ihrer Dissertation dazu Möglichkeiten vor.

Ein letzter Ansatzpunkt, der schon über eine reine Betrachtung auf der Modellebene hinausgeht, ist die Frage nach der *Repräsentativität* der in den Korrekturen verwendeten Größen. Hier können Intervalle den möglichen Wertebereich der vorkommenden Größen einschließen.

### Ansatzpunkte im Wechselspiel Modell - Realität

Im Spannungsfeld Modell - Realität soll nochmals auf den Gedanken der Repräsentativität zurückgegriffen werden. Betrachtet man die Herleitung von gängigen Korrektionsformeln wie die des Brechungsindex nach Barrell und Sears (RÜEGER 1999) oder neuerer Werte (CIDDOR 1996; CIDDOR UND HILL 1999), so handelt es sich dabei um Verknüpfungen verschiedener Experimente und Datensätze. Auf Probleme und gegebenenfalls Fehler, die bei solchen Verknüpfungen unterlaufen können, verweisen DEICHL (1984) oder RÜEGER (1999). Neben der Varianz des Mittels, das ein mittleres oder durchschnittliches Verhalten angibt, können minimale und maximale Ergebnisse der Experimente herangezogen werden, um die mögliche Variabilität bzw. die Impräzision abzuschätzen.

Betrachten wir die Bestimmung des mittleren Refraktionskoeffizienten  $\kappa$ . Eine Mittelung ergab den klassischen Wert mit Standardabweichung von  $\kappa = 0.13 \pm 0.04$ . Dieser Wert ist über den Tagesverlauf nicht konstant. Eine Alternative wäre also, wenn auf der Ebene dieses einfachen Modells gerechnet werden soll, eine größere Variationsbreite anzunehmen. Diese kann beispielsweise aus dem Tageslauf zu  $[\kappa] = [-3, 4]$  abgeschätzt werden, vgl. KAHMEN (1993, S.467) oder auch RÜEGER (1996, S.59). Dadurch würde der Impräzision, dem mangelnden Wissen um das tatsächliche Verhalten bzw. der fehlenden Repräsentativität dieses einen Wertes, Rechnung getragen.

Ein weiterer Ansatzpunkt liegt in der Betrachtung von *Rundungsfehlern*. In verschiedenen klassischen Abhandlungen und im Leitfaden (DEUTSCHES INSTITUT FÜR NORMUNG 1995) wird das Vorkommen jedes Wertes in einem symmetrischen Intervall um die letzte angegebene Stelle mit einem Radius, der die Hälfte der Einheit der letzten Stelle beträgt, als gleich wahrscheinlich angesehen. Rundungsfehler werden als Varianzen mit dem Wert  $\frac{1}{\sqrt{12}}$  der Einheit der letzten Stelle aufgefasst (DEUTSCHES INSTITUT FÜR NORMUNG 1995, S.56). Allerdings erlaubt erst die Verwendung von Intervallen (Wert  $\pm 0.5$  Einheit der letzten Stelle) die vollständige Erfassung der Rundungsfehler, da sich maximale Abweichungen an den Intervallrändern ergeben. Durch Rundung nach außen können diese fortgepflanzt und verifizierte Bereiche für die Ergebnisse erhalten werden. Hier ist eine der Wurzeln der Intervallmathematik zu sehen, vgl. KULISCH (1989). Geräteintern verarbeitete Korrekturen weisen Rundungsfehler auf. Ebenso wird der Anzeigewert, bzw. der im Messprotokoll übergebene Wert auf feste Nachkommastellen gerundet (MEIER-HIRMER UND MEIER-HIRMER 1997). Aus Speicherplatzgründen werden verschiedene zusätzliche Daten, wie Geoidmodelle oder Wetterdaten nur mit beschränkter Stellenzahl abgelegt.

Die Effekte, die sich im Spannungsfeld Modell - Realität ergeben, sollen als *Rest- und Rundungsfehler* aufgefasst werden. Für jede Messung ist ein neuer Einflussparameter  $\Delta_i$  einzuführen. Zusammengefasst ergibt sich der

Vektor  $\mathbf{p}_L = [-\Delta, \Delta]$ . Die Einflussparameter  $\mathbf{p}$  werden somit um einen zusätzlichen vierten Typ  $\mathbf{p}_L$  ergänzt und lassen sich darstellen als

$$\begin{aligned} [\mathbf{p}] &= \underbrace{\mathbf{p}_m}_{\text{Wert}} + \underbrace{[-\mathbf{p}_r, \mathbf{p}_r]}_{\text{Impräzision}}, \end{aligned} \quad (2.67)$$

wobei die Intervallmitte  $\mathbf{p}_m = (\mathbf{p}_M, \mathbf{p}_Z, \mathbf{p}_S, \mathbf{0})^T$  den eigentlichen Wert des Einflussparameters angibt und die Intervallradien  $\mathbf{p}_r = (\mathbf{d}\mathbf{p}_M, \mathbf{d}\mathbf{p}_Z, \mathbf{d}\mathbf{p}_S, \Delta)^T$  die Impräzision ausdrücken.

Abschließend soll noch der Frage nachgegangen werden, wie weit funktionale Abhängigkeiten und Zusammenhänge zurückverfolgt und in einzelne Einflussparameter zerlegt werden sollen. Grundsätzlich gilt: je weiter rückverfolgt wird, desto mehr Korrelationen und Abhängigkeiten können explizit berücksichtigt werden. Dies ist wesentlich für die Intervallauswertung, die in den nächsten Abschnitten angesprochen wird. Werden keine Abhängigkeiten zwischen Werten berücksichtigt und wird keine Zerlegung in gemeinsame Einflussparameter vorgenommen, so werden die Werte bei der Intervallauswertung als unabhängig angesehen. Die nichtmodellierten Abhängigkeiten führen zu Überschätzungen bzw. dem „wrapping effect“, vergleiche auch Abschnitt 2.2.3. Dadurch werden die Ergebnisintervalle unkontrolliert aufgebläht, so dass die Ergebnisse im Extremfall nicht mehr interpretierbar und als Impräzisionsmaße ungeeignet sind. In den folgenden Abschnitten wird gezeigt, wie die Intervalle der Einflussparameter auf die korrigierten Messungen fortgepflanzt werden. Dazu stehen zwei Ansätze zur Verfügung: zum Einen die *Intervallauswertung des Aufbereitungsprozesses der Messung* und zum Anderen die Ableitung der Intervalle aus der *Sensitivitätsanalyse des Aufbereitungsprozesses der Messung*.

### 2.3.4 Intervallauswertung des Aufbereitungsprozesses der Messung

Stehen die zu betrachtenden Intervalle für die Einflussparameter fest, so können die Korrekturen als Funktion dieser Parameter intervallausgewertet werden, wie dies in Abschnitt 2.2.2 vorgestellt wurde. Man erhält dann einen Ergebnisintervallvektor  $[\mathbf{l}]$  für den Vektor der korrigierten Messungen durch Intervallauswertung von (2.66)

$$[\mathbf{l}] = \mathbf{M} \cdot ([\mathbf{l}_A] \circ [\mathbf{k}_P]) \circ [\mathbf{k}_G] + [\Delta]. \quad (2.68)$$

Der erste Term gibt die Impräzisionsbetrachtung auf der Ebene des Modells an, der zweite Term  $\mathbf{p}_L$  schätzt die Unsicherheit nichtmodellierbarer Effekte ab. Für jede Messung ergibt sich ein Variationsbereich in Form des Intervalls  $[l_i, \bar{l}_i]$ . Innerhalb dieses Bereichs kann der korrigierte Messwert höchstens variieren, wenn die Einflussparameter mit den vorgegebenen Intervallen maximal abgeschätzt wurden. Wegen des mehrfachen Auftretens von Einflussparametern in den Korrekturfunktionen liefert das Intervall für jede korrigierte Messung in der Regel eine Überschätzung des eigentlichen Wertebereichs, vgl. Abschnitt 2.2.3. Die Vorgehensweise der Intervallauswertung des Aufbereitungsprozesses der Messungen weist folgende Vorteile auf:

- Der Wertebereich wird durch das Intervall  $[\mathbf{l}]$  eingeschlossen, d.h. alle Aussagen bezüglich des Wertebereichs bewegen sich auf der „sicheren Seite“. Die maximal möglichen Fälle oder der *worst case* im Kontext dieser Modellierung sind erfasst.
- Spezielle Softwareerweiterungen erlauben eine direkte und schnelle Umsetzung von (2.68). Die XSC-Erweiterung gängiger Programmiersprachen wie Pascal oder C stellen Intervalle als spezielle Operatoren und Intervalloperationen zur Verfügung (KLATTE ET AL. 1991; KLATTE 1993). Für MATLAB steht die kostenfrei zu beziehende INTLAB-Toolbox (RUMP 1999; RUMP 2002) bereit. Diese Spracherweiterungen erlauben einerseits monotone Rundungen zum Abfangen von Rundungsfehlern und andererseits das Rechnen mit Intervallen, wie in Abschnitt 2.2.2 vorgestellt. Eine schnelle Programmierbarkeit von Intervalloperationen kann durch Overloading der entsprechenden reellen Operationen erreicht werden. Damit stellt die Programmierung der Korrektur auch gleich die Umsetzung von (2.68) her.

Andererseits zeigt der Ansatz neben der Überschätzung des Wertebereichs noch einige Nachteile:

- Einer zeitsparenden Programmierung wie oben angedeutet kann entgegenstehen, dass sich ggf. die Operationsreihenfolge, wie sie im existierenden Programmcode abgelegt ist, nicht direkt zu einer Intervallauswertung eignet. Überschätzungen treten auf, wenn Variablen mehrfach im Funktionsausdruck vorkommen oder dieser eine für die Intervallauswertung ungünstige Form hat. Diesen Effekten kann durch Umformungen entgegen gewirkt werden, die genügend Erfahrungen im Umgang mit Intervallrechnung voraussetzen.

- Durch unterschiedliche Implementierungen der Intervallauswertung in verschiedenen Programmen, die mathematisch aquivalente Aufbereitsprozesse der Messungen abbilden, konnen im Extremfall unterschiedliche Ergebnisintervalle erzielt werden. Prinzipiell sollte ein solches Resultat vermieden werden, da dies in letzter Konsequenz aussagt, dass die Unsicherheit des Ergebnisses je nach Umsetzung im Programmcode variabel ist. Hier sind weitergehende Untersuchungen notwendig, um die Unterschiede fur verschiedene Implementierungen relevanter Formeln zu erfassen, was uber den Rahmen dieser Arbeit weit hinausgeht. Liegen genugend Erfahrungswerte vor, so wird man mit der Intervallauswertung besser umgehen konnen.
- Die kompakte Losung als Ergebnisintervall lasst keine Ruckschlusse auf Fragen nach Einflussparametern mit maximalen oder minimalen Beitragen direkt zu. Diese Fragestellungen sind jedoch in einer Design- bzw. Optimierungsphase von Messoperation oder Sensoren sehr nutzlich.

### 2.3.5 Intervalle aus der Sensitivitatsanalyse des Aufbereitungsprozesses der Messung

Ein alternativer Ansatz zur Ableitung von Intervallen basiert auf der Sensitivitatsbetrachtung des gesamten Messungs-Aufbereits-Prozesses. Die Sensitivitat einer Korrektur in Bezug auf die Einflussparameter wird mit Hilfe des totalen Differentials ermittelt. Hierzu wird die Funktion  $f$  aus (2.53) an der Stelle der Werte der Einflussparameter linearisiert, die der Intervallmitte der Intervalle fur die Einflussparameter  $p_{m,i}$  entspricht:

$$dk = \sum_{j=1}^{\nu} \frac{\partial f}{\partial p_j} dp_j = \mathbf{f}^T \mathbf{dp}, \quad (2.69)$$

wobei  $\mathbf{f}$  den  $\nu \times 1$  Vektor der partiellen Ableitungen darstellt und  $\mathbf{dp}$  den  $\nu \times 1$  Vektor der Einflussparameter. Fur den Wert der Korrektur wird keine Linearisierung angesetzt. Er ergibt sich aus  $k = f(\mathbf{p}_m)$ . Geht man entsprechend fur alle Korrekturen und Messmethoden vor und fasst die Ergebnisse in Vektoren zusammen, so erhalt man fur den Vektor der korrigierten Messungen den Zusammenhang

$$\mathbf{dl} = \mathbf{F} \mathbf{dp} + \mathbf{I} \Delta. \quad (2.70)$$

Der erste Term entspricht der Imprazisionsbetrachtung auf Modellebene, der zweite Term schatzt die nichtmodellierbaren Effekte wie Rest- und Rundungsfehler ab, wobei  $\mathbf{I}$  die  $n \times n$  Einheitsmatrix darstellt. Im Folgenden soll mit  $\bar{\mathbf{F}} := (\mathbf{F}, \mathbf{I})$  die um  $\mathbf{I}$  erweiterte Matrix  $\mathbf{F}$  verstanden werden und mit  $\mathbf{p}$  den Vektor aller vier Typen an Einflussparametern. Da im Folgenden von der Matrix  $\mathbf{F}$  kein Gebrauch mehr gemacht wird und somit keine Verwechslungsfahr besteht, soll mit  $\mathbf{F}$  die erweiterte Matrix bezeichnet werden.

Im Gegensatz zur Intervallauswertung, die in einem Guss Ergebnisintervalle liefert, werden hier die Intervalle aus den getrennt berechneten Intervallradien und Intervallmitten zusammengesetzt. Die Grundidee ist, dass das entstehende Intervall die maximal moglichen Abweichungen beschreiben soll. Dazu durfen sich keine Anteile bei der Matrix-Vektor-Multiplikation in (2.70) gegenseitig reduzieren. Dies kann erreicht werden, indem der Absolutbetrag der Elemente der Matrix  $|\mathbf{F}|$  verwendet wird. Interpretiert man die Differentiale in (2.70) als Intervallradien,  $\mathbf{p}_r := \mathbf{dp}$ , so ergibt sich der Zusammenhang fur die Intervallradien  $\mathbf{l}_r$  der korrigierten Beobachtungen (2.70) zu

$$\mathbf{l}_r = |\mathbf{F}| \mathbf{p}_r. \quad (2.71)$$

Die Intervallradien sind die Trager der Imprazision, die Intervallmitten sind die Trager der Stochastizitat der Beobachtungen. Unter Annahme einer Normalverteilung gilt:

$$\mathbf{l}_m \sim N(E(\mathbf{l}), D(\mathbf{l})), \quad (2.72)$$

wobei  $E(\cdot)$  den Erwartungswertoperator und  $D(\cdot)$  den Dispersionsoperator bezeichnet. Die Intervallmitten sollen sinnvollerweise den Werten der korrigierten Messung selbst entsprechen:

$$\mathbf{l}_m = \mathbf{M}(\mathbf{l}_A \circ \mathbf{k}_P) \circ \mathbf{k}_G. \quad (2.73)$$

Die Intervallmitten und -radien werden zu einem Intervall zusammengefasst, das sich mit (2.67) darstellen lasst:

$$[\mathbf{l}] = \mathbf{l}(\mathbf{p}_m) + \frac{\partial \mathbf{l}}{\partial \mathbf{p}} \cdot [-\mathbf{p}_r, \mathbf{p}_r] \quad (2.74)$$

$$= \mathbf{l}_m + \mathbf{F} \cdot [-\mathbf{p}_r, \mathbf{p}_r]. \quad (2.75)$$

Gleichung (2.74) entspricht in ihrer Form der allgemeinen Funktion  $f_A([\mathbf{x}])$ , die in Abschnitt 2.2.3 vorgestellt wurde. Die deutliche Trennung in den korrigierten Messwert  $\mathbf{I}_m$  und seine Impräzision  $\mathbf{F} \cdot [-\mathbf{p}_r, \mathbf{p}_r]$  wird durch die Darstellung (2.75) unterstrichen. Die Vorteile dieser Vorgehensweise liegen darin:

- Der Ansatz im Kontext der Sensitivitätsanalyse liefert eine eindeutige Lösung, da die Intervalle der Einflussparameter zuerst in Intervallmitte und -radius zerlegt werden. Eine getrennte Fortpflanzung kann somit unter Verzicht von Intervallmethoden ablaufen. Schließlich werden die Ergebnisse für die korrigierte Messung wieder zu Intervallen zusammengefügt, vgl. (2.74).
- Die lineare Form (2.71) ermöglicht einen direkten Zugang zu den Einflussparametern. Ihre Beiträge zum Gesamteinfluss setzen sich zusammen aus den Intervallradien der Einflussparameter und dem linearen Fortpflanzungsmechanismus, der in den partiellen Ableitungen abgebildet ist. Es können nun Einflussparameter ausgemacht werden, die den Wert des Intervallradius der korrigierten Beobachtungen beispielsweise dominieren oder nur einen geringen Einfluss auf ihn haben. Eine Variation der Auswertemethoden oder eine Verwendung verbesserter Korrektionsmodelle können die Magnitude der einzelnen Beiträge der Einflussparameter steuern. Auf die sich daraus ergebenden Interpretations- und Designmöglichkeiten soll in Abschnitt 2.3.7 bzw. 3.2 noch detailliert eingegangen werden.
- Es sind keine Veränderungen im Programmcode für die bestehende Implementierung der Korrektion nötig. Lediglich müssen Ergänzungen zur Berechnung der Intervallradien durchgeführt werden, beispielsweise in Form des zweiten Terms in Gleichung (2.75). Die Umsetzung dieses Ansatzes benötigt keine speziellen Kenntnisse der Intervallmathematik.

Nachteil dieses Ansatzes ist, dass der Wertebereich der korrigierten Messungen nicht eingeschlossen, sondern nur approximiert werden kann, worauf schon in Abschnitt 2.2.3 eingegangen wurde. Genauere Aussagen hierzu können für beliebige Funktionen von mehreren Veränderlichen nicht getroffen werden. Ebenso konnte keine a priori Abschätzung der Approximationsgüte des Wertebereichs gefunden werden. Es wird anhand von Korrekturen für die Streckenmessung im Abschnitt 3.4 gezeigt, dass der Wertebereich und die Intervalle, die sich aus einer Sensitivitätsanalyse berechnen lassen, gut übereinstimmen.

In der vorliegenden Arbeit wird das Kriterium der eindeutigen Lösung der vollständigen Einschließung beliebiger Funktionen vorgezogen. Alle nachfolgenden Untersuchungen beziehen sich daher immer auf die Intervalle, die mittels Sensitivitätsbetrachtung bestimmt wurden.

Die DIN 1319 sieht die Angabe der Unsicherheit nach (2.71) als eine alternative Darstellung der Unsicherheit an, insbesondere für Anwendungen, bei denen sichergestellt werden muss, dass die Messabweichung einen gewissen Höchstbetrag nicht überschreiten darf (DEUTSCHES INSTITUT FÜR NORMUNG 1998, S.22f). Die Ausführungen der DIN 1319 enden mit der Betrachtung der Kenngrößen für die maximal mögliche Messabweichung und dem Hinweis, dass diese nicht weiterverarbeitet werden sollen. Durch die Formalisierung der Impräzision als Intervalle wie in der vorliegenden Arbeit ist es jedoch möglich, Stochastizität und Impräzision zu kombinieren, was im folgenden Abschnitt kurz skizziert wird.

Abschließend soll auf Analogien der zwischen der Bestimmung von Intervallen aus der Sensitivitätsbetrachtung und der *affinen Arithmetik* eingegangen werden. Ausgangspunkt für diese Überlegungen spielt der „wrapping effect“. Von vielen Autoren wurde versucht, die Ursachen, die in der Vernachlässigung von funktionalen Abhängigkeiten zwischen Variablen liegen, zu mildern. Eine Zusammenstellung verschiedener Methoden findet man beispielsweise bei LOHNER (1989) oder KÜHN (1998). NEUMAIER (2002) stellt einen aktuellen Vergleich verschiedener Ansätze vor. Ein Verfahren, um den „wrapping effect“ auszuschalten, ist die affine Arithmetik. Erste Ideen hierzu findet man bereits bei HANSEN (1975), der von *generalized intervals* spricht. In den 90er Jahren wurden diese Ideen wieder aufgegriffen und für Anwendungen in der Computergraphik weiterentwickelt (COMBA UND STOLFI 1993). Die konkrete Fragestellung, die es zu lösen galt, war, eine verbesserte garantierte Einschließung von Nullstellen zu erhalten, die beispielsweise beim Schnitt von Flächen berechnet werden müssen.

Einen guten Überblick und Vergleich zur Intervallmathematik liefern DE FIGUIEREDO UND STOLFI (1997), worauf sich die folgenden Ausführungen beziehen. Unter einer *affinen Form* wird der folgende Ausdruck verstanden (DE FIGUIEREDO UND STOLFI 1997, S.43):

$$x = x_0 + x_1 \epsilon_1 + x_2 \epsilon_2 + \dots + x_n \epsilon_n. \quad (2.76)$$

Die  $\epsilon_i$  sind die sogenannten Fehlersymbole. Ihr exakter Wert ist unbekannt, es gilt aber:  $\epsilon_i \in [-1, 1]$ . Die Variablen  $x_i$  entsprechen den partiellen Ableitungen des Ausdrucks  $x$  nach den einzelnen Einflussparametern,

und  $x_0$  stellt einem konstanten Term dar. Bezeichnet man mit  $\mathbf{P}_r := \text{diag}(\mathbf{p}_r)$  die Diagonalmatrix, deren Diagonalelemente die Intervallradien  $\mathbf{p}_r$  sind, so lässt sich (2.75) als affine Form schreiben:

$$\mathbf{l} = \mathbf{l}_m + (\mathbf{F}\mathbf{P}_r)\boldsymbol{\epsilon}. \quad (2.77)$$

Die allgemeine Funktion  $f_A([x])$  aus (2.4) stellt ebenfalls eine affine Form dar. Sie erlaubt es, den „wrapping effect“ dadurch weitgehend auszuschalten, dass die funktionalen Abhängigkeiten zwischen den Variablen berücksichtigt werden.

Für beliebige Funktionen, die nicht in affiner Form vorliegen und somit nicht als Linearkombination der Fehlersymbole darstellbar sind, schlagen COMBA UND STOLFI vor, eine Ersatzfunktion in affiner Form zu wählen und ein weiteres Fehlersymbol  $\epsilon_k$  für die Approximationsgüte einzuführen:

$$f^* = f_A(\epsilon_1, \dots, \epsilon_n) + z_k \epsilon_k, \text{ mit: } z_k := \max_{\epsilon_1, \dots, \epsilon_n} |f^* - f_A|. \quad (2.78)$$

Der Faktor  $z_k$  soll so bestimmt werden, dass  $f_A$  das Polynom erster Ordnung ist, das  $f^*$  im Sinne *minimaler Maximalfehler* bestens approximiert. „Minimale Maximalfehler“ werden auch von anderen Autoren als Kriterium verwendet, um einen speziellen Wert aus einem Intervall auszuwählen (HEINDL UND REINHART 1976a). Die Anwendungen in der Geodäsie werden in Abschnitt 4.5.1 weiter besprochen.

### 2.3.6 Vergleich der Reduktion von Impräzision und Stochastizität durch Messmethoden

in Abschnitt 2.3.2 wurde gezeigt, wie durch Messmethoden und -anordnungen verschiedene systematisch wirkende Effekte eliminiert oder zu Resteffekten reduziert werden können. Im Folgenden soll am Beispiel der Differenzbildung und Mittelung verglichen werden, wie sich die Gesamtunsicherheit (Stochastizität und Impräzision) durch Anwendung von Messmethoden auf die korrigierten Messungen fortpflanzt. Exemplarisch sollen zwei Beobachtungen  $\ell_1$  und  $\ell_2$  vorliegen, aus denen durch Anwendung der Messmethode eine korrigierte Beobachtung  $l$  entsteht. Die Varianz-Kovarianz-Matrix, die die Stochastizität ausdrückt, sei durch  $\mathbf{C}_{\ell\ell}$  gegeben.

$$\mathbf{F} = \begin{pmatrix} \mathbf{f}^1 & 1 & 0 \\ \mathbf{f}^2 & 0 & 1 \end{pmatrix}, \quad \mathbf{C}_{\ell\ell} = \begin{pmatrix} \sigma_{\ell_1}^2 & \sigma_{\ell_1, \ell_2} \\ \sigma_{\ell_2, \ell_1} & \sigma_{\ell_2}^2 \end{pmatrix} \quad (2.79)$$

Die Impräzision sei durch den Zusammenhang  $\ell_r = |\mathbf{F}| \mathbf{p}_r$  angedeutet, wobei sich  $\mathbf{F}$  nach (2.79) in Zeilen zerlegen lässt.

**Differenzbildung:** Bei der Differenzbildung ist die Matrix der Messmethode zu  $\mathbf{M} = \begin{pmatrix} 1 & -1 \end{pmatrix}$  gegeben. Für die Stochastizität folgt nach dem Varianzfortpflanzungsgesetz:

$$\begin{aligned} \sigma_l^2 &= \mathbf{M} \mathbf{C}_{\ell\ell} \mathbf{M}^T \\ &= \sigma_{\ell_1}^2 - 2 \sigma_{\ell_1, \ell_2} + \sigma_{\ell_2}^2. \end{aligned} \quad (2.80)$$

Wird  $\sigma_{\ell_1}^2 = \sigma_{\ell_2}^2 = \sigma^2$  angenommen, so kann je nach Korrelation zwischen den Beobachtungen die Varianz der korrigierten Beobachtung Werte zwischen 0 und  $4 \sigma^2$  annehmen. Interessiert man sich insbesondere für die Größe der Varianz im Vergleich zu den Varianzen der Eingangsbeobachtungen, so kann festgestellt werden, dass für Korrelationen  $\varrho \leq \frac{1}{2}$  die Varianz der Differenz größer ist als die Varianz einer Einzelbeobachtung. Daraus lässt sich schließen, dass im Allgemeinen eine Differenzbildung zur Erhöhung des Messrauschens führt.

Für die Fortpflanzung der Impräzision ist es sinnvoll, auf die Einflussparameter zurückzugehen. Würden hingegen die Intervalle  $[l_1]$  und  $[l_2]$  direkt betrachtet, so würde angenommen, dass die Impräzision bei der Messungen vollständig unabhängig sei und diese getrennt (maximal) fortgepflanzt werde. Es liegen aber hier funktionale Abhängigkeiten von gemeinsamen Einflussparametern vor. Ein Nichtbeachten dieser Abhängigkeiten führt wie schon ausgeführt auf den „wrapping effect“, vgl. Abschnitt 2.2.2, so dass jede nachfolgende Operation den Impräzisionsbereich weiter aufbläht. Als Folge werden Intervalle berechnet, deren Aussagekraft für eine realistische Impräzisionsbeschreibung ungeeignet geworden ist. Das Ergebnisintervall für die korrigierte Messungen berechnet sich unter Berücksichtigung der gemeinsamen Abhängigkeiten:

$$\begin{aligned} l_r &= |\mathbf{M} \mathbf{F}| \mathbf{p}_r \\ &= |\mathbf{f}^1 - \mathbf{f}^2| \mathbf{p}_r + \Delta_1 + \Delta_2. \end{aligned} \quad (2.81)$$

Liegt eine ähnliche Modellierung für beide Beobachtungen vor, so stimmen die Koeffizienten der Zeilenvektoren  $\mathbf{f}^1$  und  $\mathbf{f}^2$  gut überein, so dass die modellierbaren Effekte eliminiert oder zumindest stark reduziert werden. Die nichtmodellierbaren Effekte, die durch die Intervalle  $\Delta_1$  und  $\Delta_2$  ausgedrückt werden, addieren sich jedoch.

**Fazit:** Die Differenzbildung reduziert die zahlenmäßig großen Beiträge der modellierbaren Effekte der Impräzision, nicht modellierbare Effekte werden jedoch summiert. Eine Differenzbildung zwischen Messungen führt im Allgemeinen dazu, dass sich das Messrauschen erhöht.

**Mittlung:** Bei der Mittelung ist die Matrix der Messmethode durch  $\mathbf{M} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}$  gegeben. Mit dem Varianzfortpflanzungsgesetz erhält man für die Varianz der korrigierten Beobachtung:

$$\begin{aligned} \sigma_i^2 &= \mathbf{M} \mathbf{C}_{\ell\ell} \mathbf{M}^T \\ &= \frac{1}{4} (\sigma_{\ell_1}^2 + 2 \sigma_{\ell_1, \ell_2} + \sigma_{\ell_2}^2). \end{aligned} \quad (2.82)$$

Werden wieder gleich große Varianzen für die Beobachtungen betrachtet, so kann je nach Korrelation zwischen den Messungen, die Varianz der korrigierten Beobachtungen zwischen 0 und  $\sigma^2$  liegen. Somit wird durch Mittelung die Varianz verringert.

Werden  $n$  unkorrelierte Messungen mit gleich großen Varianzen gemittelt, so errechnet sich die Varianz der korrigierten Messung zu:

$$\sigma_i^2 = \frac{\sigma^2}{n}. \quad (2.83)$$

Bei Mittelung von Mehrfachmessung wird die Varianz des Mittels vermindert. Offen bleibt, bis zu welchem  $n$  die Aussagen und Folgerungen noch realistisch sind.

Für die Fortpflanzung von Impräzision bei der Mittelung sind zwei Fälle zu unterscheiden. Zuerst soll nochmals die Mittelung von Messungen betrachtet werden, bei denen sich die Effekte im Vorzeichen unterscheiden. Dabei werden wieder die funktionalen Abhängigkeiten berücksichtigt. So folgt für die Impräzision der korrigierten Messung:

$$\begin{aligned} l_r &= |\mathbf{M} \mathbf{F}| \mathbf{p}_r \\ &= \frac{1}{2} |\mathbf{f}^1 + \mathbf{f}^2| \mathbf{p}_r + \frac{1}{2} (\Delta_1 + \Delta_2). \end{aligned} \quad (2.84)$$

Gilt  $\mathbf{f}^1 \approx -\mathbf{f}^2$  und  $\Delta_1 \approx \Delta_2$ , so ist die Impräzision des Ergebnisses vergleichbar mit der Impräzision bei der Differenzbildung.

Als zweiter Fall soll die Mittelung von Mehrfachmessungen untersucht werden. Hierbei treten die Effekte mit gleicher Magnitude und gleichem Vorzeichen auf,  $\mathbf{f}^1 \approx \mathbf{f}^2$ . Daher heben sich in (2.84) die Beiträge nicht gegenseitig auf. Im Gegensatz zur Varianzfortpflanzung erhält man eine Impräzision des Ergebnisses, die genauso groß ist wie die Impräzision der Eingangsgrößen. Dieser Sachverhalt ändert sich nicht, wenn beliebige viele Eingangsmessungen gemittelt werden.

Das Ergebnis lässt sich wie folgt interpretieren: Jede Messanordnung impliziert unausweichlich Unsicherheit für die Bestimmbarkeit der gesuchten Messgrößen. Bleibt man im Konzept der gewählten Messanordnung, so kann dieser Anteil der Unsicherheit nicht reduziert werden, da er immanent mit der Messanordnung verbunden ist. Erst das Verlassen oder Variieren des Messungsrahmens beispielsweise durch Verwendung alternativer Messanordnungen wird die damit verbundene Unsicherheit reduzieren. Als **Fazit** bleibt festzuhalten: Mittelung von Mehrfachmessungen reduziert die Varianz der korrigierten Messung; ihre Impräzision ist identisch mit der Impräzision der Ausgangsmessungen.

Da beide Konzepte unterschiedliche Aspekte der Gesamtunsicherheit beschreiben und eine unterschiedliche Unsicherheitsfortpflanzung aufweisen, liegt es nahe, durch eine geeignete Kombination von Stochastizität und Impräzision die Gesamtunsicherheit von Messungen treffend zu beschreiben. Dabei sollen Ergebnisse erzielt werden, die sich mit den Erfahrungen im Umgang mit Daten decken, d.h. nur ein gewisser Anteil der Unsicherheit kann durch Mehrfachmessung reduziert werden, nämlich die Stochastizität. Diese Reduktion lässt sich für die Gesamtunsicherheit nicht beliebig fortsetzen; hier bildet die Impräzision eine untere Schranke.

Für die Stochastizität wird  $\sigma = 1$  und für die Impräzision  $l_r = 0.3$  für eine Einzelmessung angesetzt, vergleiche den linken Teil der Graphik 2.10. Beide Effekte lassen sich *unter Beibehaltung ihrer Charakteristika* mengentheoretisch kombinieren. Die *erweiterte Unsicherheit* ergibt sich aus dem statistischen  $(1 - \gamma)$ -Konfidenzbereich, der vom Intervall überlagert wird, worauf in Abschnitt 4.4 noch detailliert eingegangen wird

$$k_{1-\gamma}^* := k_{1-\frac{\gamma}{2}} \sigma + l_r, \quad (2.85)$$



wobei  $k_{1-\gamma}^*$  als Radius des *erweiterten Konfidenzbereichs* bezeichnet werden soll. Auf der rechten Seite von Abbildung 2.10 ist der Verlauf der Gesamtunsicherheit sowie der Anteile Stochastizität und Impräzision bei der Ausführung von Mehrfachmessungen dargestellt. Mit Erhöhung der Anzahl der Mehrfachmessungen  $n$  wird die Standardabweichung des Mittels reduziert, die Impräzision bleibt hingegen konstant. Selbst bei sehr großer Anzahl an Messungen kann das Unsicherheitsniveau  $l_r$  von der erweiterten Unsicherheit nicht unterschritten werden, es stellt vielmehr die Asymptote des Verlaufs dar.

Vergleichend hierzu soll die Vorgehensweise, die vom Leitfaden zum Umgang mit Unsicherheit (DEUTSCHES INSTITUT FÜR NORMUNG 1995) vorgeschlagen wird, beleuchtet werden. Ziel dieses Konzeptes ist es, eine universelle, einheitliche und einfache Methode zur Behandlung des Unsicherheitshaushalts zu liefern. *Universell* bedeutet in diesem Zusammenhang, dass die Methode für alle Praxisfälle angewendet werden kann und soll. Unter der *Einheitlichkeit* ist zu verstehen, dass keine Unterscheidung in systematisch wirkende und zufällige Unsicherheitskomponenten vorgenommen werden soll. Dies wird damit begründet, dass die Unterteilung in systematische und zufällige Anteile einerseits nicht immer direkt ersichtlich ist und andererseits kontextbezogen sein kann (DEUTSCHES INSTITUT FÜR NORMUNG 1995, S.48ff). Es wird weiter argumentiert, dass ein Kontextbezug dem Ziel der Universalität entgegen stehe.

Der Leitfaden geht von einer rein stochastischen Betrachtungsweise aus und postuliert in diesem Sinne die *Wesensgleichheit* aller Beiträge (DEUTSCHES INSTITUT FÜR NORMUNG 1995, S.47ff). Es wird lediglich unterschieden, auf welche Art die Varianzen, die als Kenngrößen dienen, ermittelt wurden: die Kategorie A beschreibt Varianzen, die aus Beobachtungen geschätzt wurden (S.9ff), Kategorie B solche, die mit anderen Methoden ermittelt wurden (S.12ff). *Einfach* bedeutet schließlich, dass im Allgemeinen vom *Normalfall* ausgegangen wird (S.67ff), der beispielsweise durch die Annahme „gutartiger“ Verteilungen wie der Normal- oder Gleichverteilung, gleich großer Einzelbeiträge zur kombinierten Unsicherheit, vollständiger Korrektur aller Systematiken oder der Gültigkeit des zentralen Grenzwertsatzes gekennzeichnet ist. Im „Normalfall“ vereinfacht sich die Betrachtungsweise sehr.

Kennzeichnend für den Leitfaden sind die oben genannten Grundgedanken. Verschiedene alternative Darstellungsmöglichkeiten werden zwar skizziert, aber zugunsten eines einheitlichen Gesamtkonzeptes nicht zu Ende verfolgt, sondern argumentativ in dieses eingepasst. Beispielsweise werden Konfidenzbereiche für kombinierte Zufallsvariablen unterschiedlicher Verteilungen unter Vernachlässigung dieser Information nur auf Basis der Varianzen berechnet, wobei auf die damit verbundene Problematik hingewiesen wird (DEUTSCHES INSTITUT FÜR NORMUNG 1995, S.23). Insbesondere muss kritisch hinterfragt werden, ob die Anwendungen in der Geodäsie immer den vereinfachenden Bedingungen des Leitfadens genügen, beispielsweise, ob alle Korrekturen angebracht werden oder ob dies gar nicht möglich ist, ob die Beiträge aller Einflussparameter gleich groß sind oder ob nicht doch dominante Effekte vorliegen können.

Wird die Gesamtunsicherheit wie im Leitfaden vorgeschlagen als stochastisch aufgefasst und beide Varianzen einfach addiert, so ergibt sich die in Abbildung 2.10 hell dargestellte *kombinierte Unsicherheit* von  $u_c = 1.04 = \sqrt{\sigma^2 + l_r^2}$  für die Einzelmessung. Bei einer Mittelung von Mehrfachmessungen wird die Gesamtunsicherheit des Mittels in Abhängigkeit von  $n$  beliebig klein. Die Frage nach der Anzahl der Messungen  $n$ , für die so bestimmte Größen noch realistisch und aussagekräftig sind, bleibt dann weiter offen.

Umgekehrt ist eine rein intervallmathematische Beschreibung auch nicht geeignet, beide Anteile (Stochastizität und Impräzision) adäquat zu beschreiben, da eine Reduktion zufälliger Abweichungen in einem solchen Konzept nicht möglich ist. Abschließend muss nochmals unterstrichen werden, dass sich zwei Anteile der Unsicherheit (Stochastizität und Impräzision) mit unterschiedlichen Eigenschaften identifizieren und mit unterschiedlichen Mitteln adäquat beschreiben lassen: Varianzen eignen sich zur Beschreibung zufälliger Variabilität und Intervallradien zur Beschreibung von systematisch wirkenden Effekten.

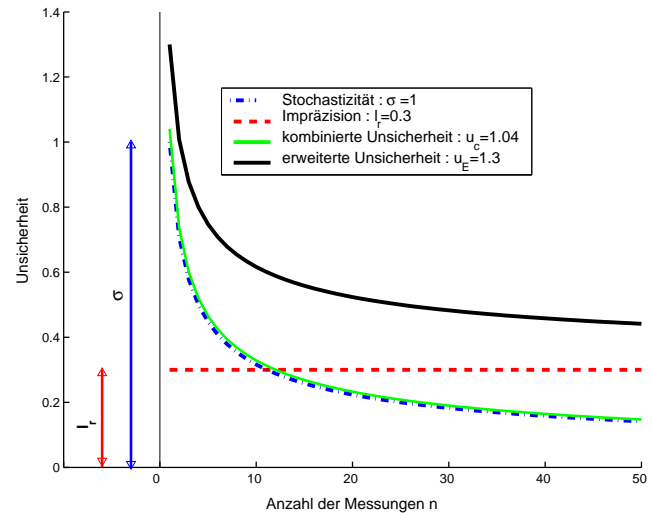


Abbildung 2.10: Schematisierung der Reduktion der Unsicherheit bei Mehrfachmessung. Die erweiterte Unsicherheit berechnet sich nach (2.85), die kombinierte Unsicherheit aus  $u_c^2 = \sigma^2 + l_r^2$ .

### 2.3.7 Identifikation von Haupteinflussfaktoren

Der wesentliche Vorteil der Ableitung der Intervalle aus der Sensitivitätsanalyse ist der direkte, lineare Zusammenhang zwischen den Eingangsintervallen der  $\nu$  Einflussparameter und ihrem Beitrag zum Intervallradius der korrigierten Beobachtungen. Mit Gleichung (2.71) lässt sich dieser darstellen:

$$\mathbf{l}_r = |\mathbf{F}| \mathbf{p}_r = |\mathbf{F} \mathbf{P}_r| \mathbf{e} = \mathbf{\Psi} \mathbf{e}, \quad (2.86)$$

wobei  $\mathbf{P}_r := \text{diag}(\mathbf{p}_r)$  die  $\nu \times \nu$  Diagonalmatrix ist, deren Diagonalelemente die per Definition positiven Radien der Einflussparameter sind und  $\mathbf{e} = (1, \dots, 1)^T$  den  $\nu \times 1$  Einsvektor darstellt. Kenngrößen, die sich eignen, den Beitrag der Einflussparameter zu quantifizieren, sind die Elemente  $\psi_{i,j}$  der Matrix  $\mathbf{\Psi}$ . Sie beschreiben den absoluten Beitrag des  $j$ -ten Einflussparameters zum Intervallradius der  $i$ -ten korrigierten Beobachtung. Interessiert man sich für die relativen Beiträge, so können diese wie folgt dargestellt werden:

$$\mathbf{\Psi}^* = \mathbf{L}_r^{-1} \mathbf{\Psi}, \quad \mathbf{L}_r := \text{diag}(\mathbf{l}_r). \quad (2.87)$$

Bei einer großen Anzahl an Einflussparametern erhöht eine Grauwertdarstellung der Matrizen  $\mathbf{\Psi}$  bzw.  $\mathbf{\Psi}^*$  die Lesbarkeit und Identifikation dominanter Beiträge.

**Beispiel 17** Betrachten wir die instrumentellen Korrekturen für die Richtungsbeobachtung, wie sie in einer einfachen Modellierung in KAHMEN (1993, Gl.(3.10) S.136 bzw. S.199) angegeben sind. Liegen detailliertere Informationen vor, so kann das einfache Modell (2.88) entsprechend erweitert und verbessert werden:

$$r_A = \rho + \frac{c_h}{\sin \zeta_A} + (Q(\rho) + i) \cot \zeta_A, \quad \text{mit} \quad \zeta_A = \zeta + i_{EX} + c_v + L(\rho), \quad (2.88)$$

wobei

$r_A$	die am Gerätedisplay angezeigte Richtung,
$\rho$	die rohe Teilkreisablesung,
$c_h, c_v$	den horizontalen bzw. vertikalen Kollimationsfehler,
$Q(\rho), L(\rho)$	die durch den Zweiachskompensator ermittelte Stehachsneigung in Kippachs- bzw. Zielachsrichtung,
$i$	den Kippachsfehler,
$\zeta_A$	die am Gerätedisplay angezeigte Zenitdistanz,
$\zeta$	die rohe Vertikalkreisablesung und
$i_{EX}$	den Indexfehler

darstellt. Der Vektor der sensorspezifischen Einflussparameter  $\mathbf{dps}$  setzt sich mit den Instrumentenfehlern folgendermaßen zusammen  $\mathbf{dps} = (dc_h, dc_v, di, dQ(\rho), dL(\rho), di_{EX})^T$ . Rundungsfehler  $\Delta_{r_A}$ ,  $\Delta_{\zeta_A}$  und Restfehler  $d\rho$ ,  $d\zeta$  werden als nicht modellierbare Effekte zusammengefasst:  $\mathbf{\Delta} = (d\rho_1, \Delta_{r_A,1}, \dots, d\rho_n, \Delta_{r_A,n}, d\zeta_1, \Delta_{\zeta_A,1}, \dots, d\zeta_n, \Delta_{\zeta_A,n})^T$ . Pro Messung (Richtung wie Zenitdistanz) wird für diese Effekte je ein Parameter eingeführt. Die Matrizen  $\mathbf{F}_S$  und  $\mathbf{F}_L$  haben für eine Richtungsbeobachtung in Lage I (erste Zeile) bzw. in Lage II (zweite Zeile) unter Berücksichtigung der oben gewählten Reihenfolge der Einflussparameter folgende Gestalt:

$$\mathbf{F}_S = \begin{pmatrix} \frac{1}{\sin \zeta_I} & -\frac{\cos \zeta_I c_h + Q(\rho_I) + i}{\sin^2 \cot \zeta_I} & \cot \zeta_I & -\frac{\cos \zeta_I c_h + Q(\rho_I) + i}{\sin^2 \zeta_I} & \cot \zeta_I & -\frac{\cos \zeta_I c_h + Q(\rho_I) + i}{\sin^2 \zeta_I} \\ \frac{1}{\sin \zeta_{II}} & -\frac{\cos \zeta_{II} c_h + Q(\rho_{II}) + i}{\sin^2 \zeta_{II}} & -\cot \zeta_{II} & \frac{\cos \zeta_{II} c_h + Q(\rho_{II}) + i}{\sin^2 \zeta_{II}} & \cot \zeta_{II} & -\frac{\cos \zeta_{II} c_h + Q(\rho_{II}) + i}{\sin^2 \zeta_{II}} \end{pmatrix},$$

$$\mathbf{F}_L = \begin{pmatrix} 1 & 1 & 0 & 0 & -\frac{\cos \zeta_I c_h + Q(\rho_I) + i}{\sin^2 \zeta_I} & -\frac{\cos \zeta_I c_h + Q(\rho_I) + i}{\sin^2 \zeta_I} & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & -\frac{\cos \zeta_{II} c_h + Q(\rho_{II}) + i}{\sin^2 \zeta_{II}} & -\frac{\cos \zeta_{II} c_h + Q(\rho_{II}) + i}{\sin^2 \zeta_{II}} \end{pmatrix}. \quad (2.89)$$

Aus (2.89) ist ersichtlich, dass die sensorspezifischen Effekte zenitdistanzabhängig sind. Mit zunehmend flacheren Visuren nimmt der Beitrag der Effekte ab, was dem allgemeinen Erfahrungsschatz entspricht. Für die folgenden Rechnungen wurde ein fehlerfreies Instrument angenommen, d.h. die Werte für  $\mathbf{ps}$  sind ungefähr Null. Dadurch

verschwindet auch rechnerisch der Einfluss der Beiträge der Zenitdistanz, vgl. zweite, vierte und sechste Spalte in  $\mathbf{F}_S$ . Die konkreten Werte für die Eingangsintervalle können wie in Abschnitt 2.3.4 beschrieben bestimmt werden. Als ein Beispiel für eine moderne Totalstation soll das Instrument TPS 1101 von Leica Geosystems verwendet werden (LEICA GEOSYSTEMS 1998; ZEISKE 1999). Entsprechende Untersuchungen können für alle Instrumente anderer Leistungsklassen und Hersteller durchgeführt werden.

Einflussgrößen	Leica TPS1101 Werte in [mgon]	Bemerkungen
$dc_h$	0.1	angenommene Unsicherheit des horizontalen Kollimationsfehler. Falls der Kollimationsfehler experimentell im Labor oder im Feld bestimmt wird, so ist er durch die Auflösung des Teilkreisabgriffes beschränkt.
$dc_v$	0.1	angenommene Unsicherheit der Indexabweichung durch vertikalen Kollimationsfehler. Auch hier ist die Auflösung gerätebedingt beschränkt.
$di$	0.05	angenommene Unsicherheit des Kippachsfehlers.
$dQ(\rho)$	0.2	Unsicherheit der durch den Zweiachskompensator ermittelten Stehachsneigung in Kippachrichtung. Diese Größe ist durch die Kompensatoreinspielgenauigkeit gegeben, vgl. LEICA GEOSYSTEMS (1998, S. 126) oder ZEISKE (1999)
$dL(\rho)$	0.2	Unsicherheit der durch den Zweiachskompensator ermittelten Stehachsneigung in Zielrichtung.
$di_{Ex}$	0.05	angenommene Unsicherheit der Indexabweichung durch Exzentrizität.
$d\rho$	0.05	Nichtmodellierter Restfehler der Horizontalteilkreisablesung.
$\Delta r_A$	0.05	Rundungsfehler zur Anzeige bzw. Speicherung in einem Stellenzahl beschränktem Wert. Die Unsicherheit wird zur Hälfte der letzten Anzeigestelle angesetzt.
$d\zeta$	0.05	Nichtmodellierter Restunsicherheit der Vertikalteilkreisablesung.
$\Delta \zeta_A$	0.05	Rundungsfehler zur Anzeige bzw. Speicherung in einem Stellenzahl beschränktem Wert.

Tabelle 2.1: Unsicherheit der Eingangsgrößen bei der Zenitdistanz- und Richtungsmessung

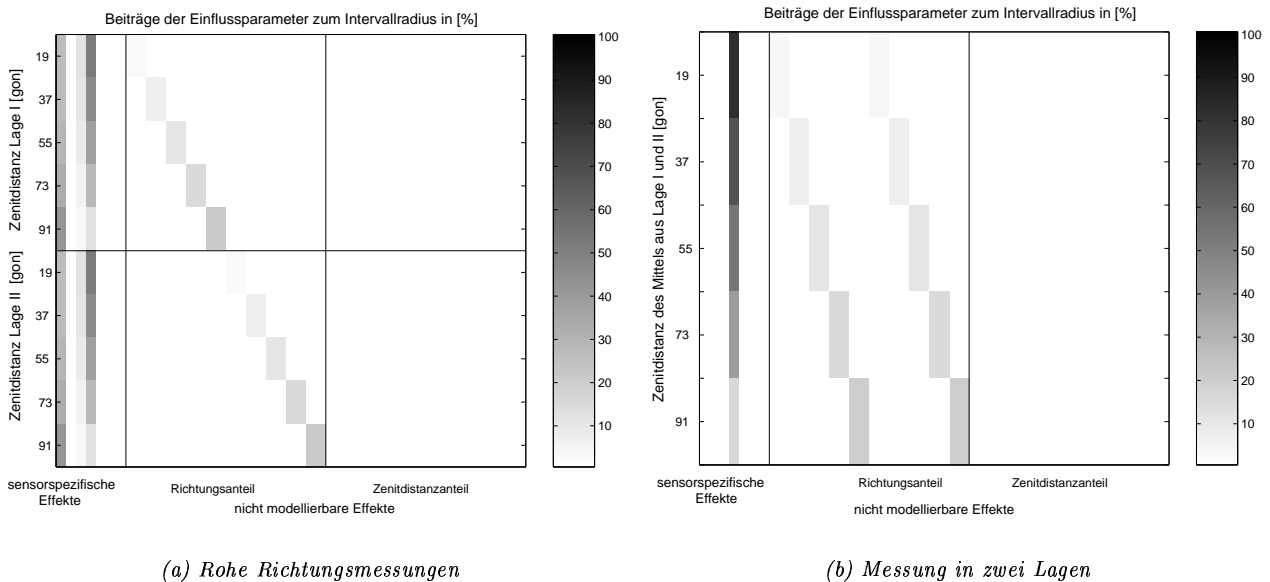


Abbildung 2.11: Relative Beiträge zum Intervallradius der Richtungsbeobachtung

Abbildung 2.11 zeigt einen Ausschnitt der Matrix  $\Psi$  für Richtungsbeobachtungen. In der Teilabbildung (a) sind die Beiträge für fünf Messungen in Lage I bei Variation der Zenitdistanz zwischen 19 gon und 91 gon und den entsprechenden fünf Messungen in Lage II als Grauwerte dargestellt. Die Spalten von  $\Psi$  beziehen sich auf die Einflussparameter in der oben angegebenen Reihenfolge: die ersten sechs Spalten stellen die Beiträge der sensorspezifischen Parameter dar, die restlichen Spalten stehen für Rundungs- und Restfehler. Dominierende Beiträge liefern die sensorspezifischen Parameter, die je nach Zenitdistanz zwischen 58% und 92% ausmachen. Die Rest- und Rundungsfehler der Zenitdistanz sowie die Indexabweichung spielen bei fehlerfreien Instrumenten keine Rolle.

In der Teilabbildung (b) sind die Einflüsse nach der Anwendung der Messmethode dargestellt. Durch die Messung in zwei Lagen und Mittelung kann der Einfluss der instrumentenspezifischen Parameter mit Ausnahme der Anteile des Stehachsfehlers eliminiert werden, da letzterer keine Symmetrien bezüglich Lage I und II aufweist. Durch die Reduktion bekommen Rundungs- und Restfehler mehr Gewicht. Abschließend sollen die Intervallradien angegeben werden, die bei Variation der Zenitdistanz erhalten werden.

Zenitdistanz $\zeta_I$ in[gon]	19.0	37.0	55.0	73.0	91.0
Intervallradien in Lage I	1.05	0.57	0.39	0.29	0.22
Intervallradien in Lage II	1.05	0.57	0.39	0.29	0.22
Intervallradien bei Mittelung Lage I und II	0.55	0.31	0.22	0.16	0.11

*Tabelle 2.2: Intervallradien der korrigierten Richtungsbeobachtung in [mgon] bei Variation der Zenitdistanz*

### 3. Beispiele für Intervallradien

Das folgende Kapitel soll zeigen, wie Intervallradien als Maße für Impräzision konkret bestimmt werden können. Erste Ansätze hierfür wurden von SCHÖN UND KUTTERER (2001a,b,c) für Streckenmessungen angegeben. In einem ersten Abschnitt wird dies exemplarisch für die terrestrischen Messungstypen Richtungen, Strecken und Zenitdistanzen durchgeführt, die mit einer modernen Totalstation gemessen werden können. Neben der vollständigen Darstellung der Abhängigkeiten der Intervallradien von den Einflussparametern werden Faustformeln für die einzelnen Messungstypen angegeben, die das wesentliche Verhalten der Intervallradien klar charakterisieren und herausarbeiten.

In weiteren Abschnitten werden einerseits die Haupteinflussfaktoren, die die Größe der Intervallradien ausmachen identifiziert. Andererseits werden die Koeffizienten der Faustformeln diskutiert. Diese Überlegungen bilden die Basis für eine Optimierung von Mess- und Auswertestrategien im Rahmen des Intervalldesigns, das in Abschnitt 5.4 eingeführt wird. Abschließend werden am Beispiel der Streckenmessung die verschiedenen Modellierungsmethoden der Intervallradien verglichen, die in den Abschnitten 2.3.4 und 2.3.5 vorgestellt wurden.

Es wird versucht, beim Vorgehen in den einzelnen Abschnitten deutlich zu unterscheiden in Zahlenwerte und Größenordnungen, die durch Literaturstellen „gesichert“ sind und in solche Angaben, die plausibel sind, aber hypothetischen Charakter aufweisen. Sie sollen eine Diskussionsgrundlage bilden und die Werte und Größenordnungen der Intervallradien der Messungen nachvollziehbar machen. Nur durchgreifende Geräteuntersuchungen bzw. Analysen der Messanordnungen können diese Werte und Größenordnungen bestimmen.

#### 3.1 Intervallradien als Impräzisionsmaße: Beispiel Totalstation

In diesem Abschnitt wird die Impräzision für die terrestrischen Messungstypen Richtungen, Strecken und Zenitdistanzen abgeleitet, die mit einer modernen Totalstation gemessen werden. Da alle terrestrischen Messungen beispielsweise durch die Atmosphäre ähnlich beeinflusst werden, ist es sinnvoll, bei der Auswertung keine Trennung in Lage und Höhe vorzunehmen, sondern alle Messungen gemeinsam zu betrachten. Im Folgenden werden die Einzelbeiträge aller Sensorkomponenten sowie der Refraktion zusammengefasst. Somit werden die Intervallradien der korrigierten Messungen erhalten, die der euklidischen Distanz zwischen den Instrumentenreferenzpunkten zugeordnet sind. Einen Überblick über die Messverfahren und Korrekturen findet man in den Lehrbüchern wie COOPER (1987), KAHMEN (1993), MÖSER (2000) oder DEUMLICH UND STAIGER (2002). Für die elektronische Streckenmessung bilden RÜEGER (1996) und JOECKEL UND STÖBER (1999) Referenzen. Eine Übersicht der Funktionsprinzipien geodätischer Sensoren findet man bei SCHLEMMER (1996). Viele Einzeldarstellungen zu speziellen Fragestellungen sind in den geodätischen Fachzeitschriften oder in Tagungsbänden enthalten, auf die nicht im Einzelnen eingegangen werden kann.

##### 3.1.1 Zenitdistanzbeobachtung

###### Korrektionsmodell

Unter Berücksichtigung der instrumentellen Einflüsse und der Refraktion ergibt sich für die korrigierte Zenitdistanzbeobachtung  $\zeta^*$  folgendes einfaches Modell:

$$\zeta^* = (\zeta + i_{Ex} + c_v + L(\rho)) - \frac{1}{2} \frac{\sin \zeta}{n_A} \frac{\partial n_A}{\partial z^*} \varrho - \frac{\sin \alpha_A}{S} \left( \frac{n_A \cos \alpha_A d}{\sqrt{n_R^2 - n_A^2 \sin^2 \alpha_A}} - e \right) \varrho. \quad (3.1)$$

Der erste Term beschreibt die instrumenteninternen Korrekturen  $\mathbf{k}_S$ , vgl. KAHMEN (1993, S.199). Hierbei bezeichnet

- $\zeta, \rho$  die rohe Vertikal bzw. Horizontalkreisablesung,
- $i_{Ex}$  den Indexfehler,
- $c_v$  den vertikalen Kollimationsfehler und
- $L(\rho)$  die durch den Zweiachskompensator ermittelte Stehachsneigung in Zielachsrichtung.

Der zweite Term beschreibt die Korrektur des Einflusses der Refraktion. Die Formeln von MORITZ (1961, 1962) bzw. BROCKS (1939) lassen sich für die hier angenommen konstanten Brechungsindexgradienten direkt integrieren, so dass sich die obige Form ergibt. Hierbei bedeutet:

- $n_A$  den Brechungsindex der Atmosphäre, wie er beispielsweise mittels aktueller meteorologischer Werte aus den Formeln von Barrell und Sears (RÜEGER 1996, S.51) abgeleitet werden kann,  
 $\frac{\partial n_A}{\partial z^*}$  den Brechungsindexgradient im Standpunkt in Richtung des lokalen Zenits und  
 $S$  die Streckenlänge.

Der dritte Term beschreibt den Einfluss einer Schräganzielung des Tripelprismas. Dadurch ergibt sich ein Versatz des Prismenhauptpunktes, der in ein Winkelmaß umgerechnet wird (RÜEGER 1996, S.162ff). Folgende Abkürzungen wurden gewählt:

- $\alpha_A$  beschreibt den Winkel der Schräganzielung,  
 $d$  den Abstand der Prismenspitze von der Grundseite (Eintrittsfläche),  
 $e$  den Abstand der Prismenstehachse von der Grundseite und  
 $n_R$  Brechungsindex des Prismas.

Mit Hilfe der Sensitivitätsanalyse werden die Intervallradien nach (2.69)-(2.71) abgeleitet. Auf die Angabe der benötigten partiellen Ableitungen wird hier aus Platzgründen verzichtet. Im Folgenden soll das Beispiel der Messung in einer Lage betrachtet werden.

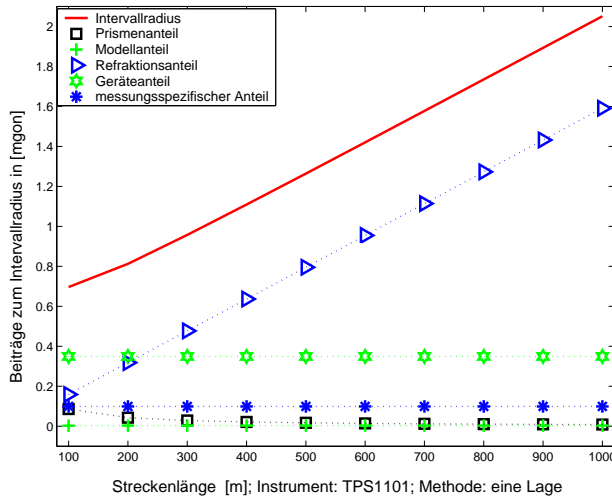
1. Prismenanteil	Brechungsindex $n_R$ , Prismenkonstanten $d, e$ Einfallswinkel $\alpha_A$
2. Modellanteil (Modellkonstanten)	Erdradius Näherungswerte aus Koordinaten für Strecke, Richtung, Zenitdistanz Näherungswerte der Koordinaten Bezugswerte der meteorologischen Modelle (Barrell Sears) $\alpha, p_0, e_0$ (JOECKEL UND STOBER 1999, S.73f) Modellkonstanten aus Magnus Tetens $c_1, c_2, c_3, c_4, c_5$ (JOECKEL UND STOBER 1999, S.75f) Modellkonstanten der Cauchy-Entwicklung $A, B, C$ (JOECKEL UND STOBER 1999, S.71f)
3. Refraktionsanteil	Feucht- und Trockentemperatur $\vartheta_w, \vartheta$ , Druck $p$ , Wasserdampfpartialdruck $p_w$ , Relative Feuchte $RF$ , Refraktionskoeffizient $\kappa$ , Brechungsindex der Atmosphäre $n_A$ , Temperaturgradienten $\frac{\partial T}{\partial x^*}, \frac{\partial T}{\partial y^*}, \frac{\partial T}{\partial z^*}$ , Brechungsindexgradienten $\frac{\partial n_A}{\partial x^*}, \frac{\partial n_A}{\partial y^*}, \frac{\partial n_A}{\partial z^*}$ , Konstanten der Brechungsindexmodelle
4. Geräteanteil	meteorologische Referenzwerte der Instrumente Trägerwellenlänge der EDM $\lambda$ , Frequenz $f_0$ , Additionskonstante $d_0$ Horizontaler und vertikaler Kollimationsfehler $c_h, c_v$ Kippachsfehler $i$ Quer- und Längsanteil des Stehachsfehlers $Q(\rho), L(\rho)$ Indexfehler $i_{E_x}$ Referenzwellenlänge der Richtungs- und Zenitdistanzbeobachtung $\lambda_{opt}$
5. Messungsspezifischer Anteil	Restfehler $dS, d\rho, d\zeta$ Rundungsfehler $\Delta_S, \Delta_{r_A}, \Delta_{\zeta_A}$

Tabelle 3.1: Übersicht über die Einflussparameter

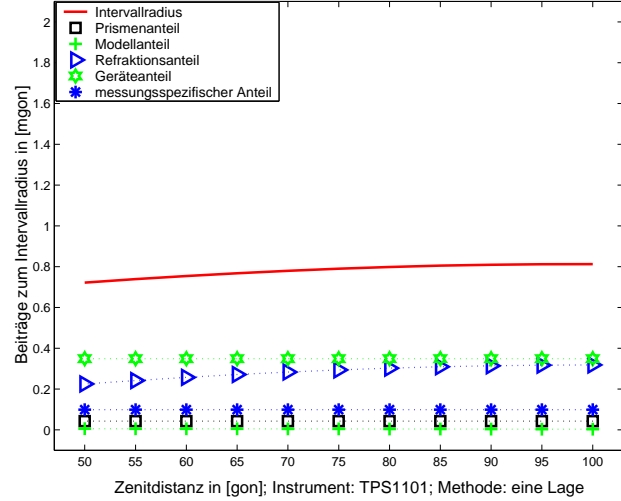
### Intervallradien für Zenitdistanzbeobachtungen in einer Lage

Die Intervallradien werden beispielhaft für das Instrument TPS1101 von Leica Geosystems abgeleitet. Die angenommenen Intervalle für die Einflussparameter und ihre Bezeichnungen wurden in Beispiel 17 eingeführt. Als Prisma wird das GPR1 von Leica Geosystems verwendet. Die Prismengeometrie sei besser als  $\pm 0.5 \text{ mm}$  bekannt (RÜEGER 1996, S.257), d.h.  $[d] = [39.1, 40.1] \text{ mm}$  und  $[e] = [25.7, 26.7] \text{ mm}$ . Der Brechungsindex des BK7 Glases des Reflektors  $n_R$  ist auf  $\pm 5 \text{ ppm}$  bekannt (RÜEGER 1996, S.155), d.h.  $[n_R] = 1.5 \pm 5 \text{ ppm}$ . Die Prismen können mit einem Diopter oder nach Augenmaß sicher besser als  $d\alpha_A = \pm 0.7 \text{ gon}$  ausgerichtet werden (RÜEGER 1996, S.163). Es soll eine Schräganzielung von  $10 \text{ gon}$  modelliert werden, so dass sich  $[\alpha_A] = [9.3, 10.7] \text{ gon}$  ergibt. Weiter wird das Refraktionsmodell nach MORITZ (1961, 1962) bzw. BROCKS (1939) mit einem konstanten Brechungsindexgradienten  $\frac{\partial n_A}{\partial x, y, z} = -0.1 \pm 0.05 \frac{\text{ppm}}{\text{m}}$  angesetzt. Die meteorologischen Rahmenbedingungen seien

am Standpunkt zu  $[\vartheta] = [19.8, 20.2]^\circ C$ ,  $[p] = [999.5, 1000.5] hPa$ ,  $[RF] = [77, 83] \%$  angenommen, vgl. für die Unsicherheiten RÜEGGER (1996, S.61ff). Tabelle 3.1 gibt einen Überblick über die Einflussparameter, die in fünf Gruppen zusammengefasst und in den Abbildungen dargestellt werden.



(a) Streckenabhängigkeit bei horizontalen Visuren



(b) Zenitdistanzabhängigkeit bei konstanter Streckenlänge von 200 m

Abbildung 3.1: Beiträge zum Intervallradius der Zenitdistanzbeobachtung bei Messung in einer Lage

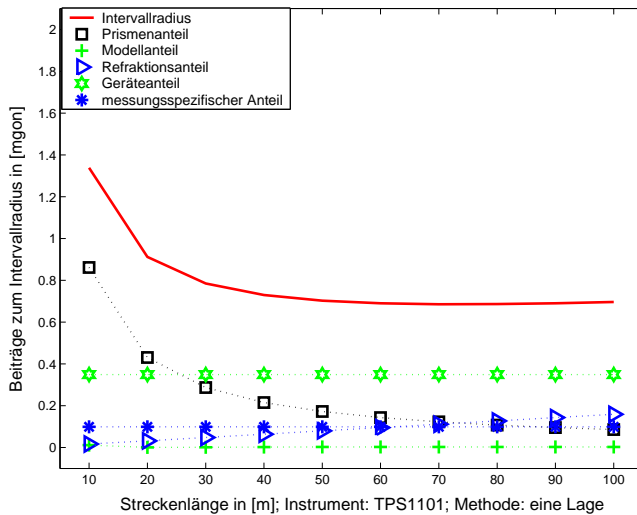


Abbildung 3.2: Beiträge zum Intervallradius der Zenitdistanzbeobachtung bei horizontaler Messung in einer Lage und Variation der Streckenlänge im Nahbereich bis 100 m

Durch Anwendung einer geeigneten Messmethode, nämlich der Messung in zwei Lagen, kann der Geräteanteil, der einen wichtigen Beitrag liefert, stark reduziert werden. Als Faustformel kann der Intervallradius für die korrigierte Zenitdistanz bei der Messung in einer Lage bzw. in zwei Lagen für dieses Beispiel folgendermaßen approximiert werden:

$$\begin{cases} \zeta_{r,I}^* & \approx (0.45 + \frac{8.7}{S} + 1.6 \cdot 10^{-3} S) [mgon], \\ \zeta_{r,II}^* & \approx (0.30 + \frac{8.7}{S} + 1.6 \cdot 10^{-3} S) [mgon]. \end{cases} \quad (3.2)$$

Hierbei wird eine Variation der Intervallradien mit der Zenitdistanz vernachlässigt. Der erste Koeffizient lässt sich geräteinternen Anteilen zuordnen, die durch Messung in zwei Lagen reduziert werden können. Der zweite

Abbildung 3.1(a) zeigt die Intervallradien der korrigierten Zenitdistanz bei Variation der Streckenlänge und horizontalen Visuren. Der streckenabhängige Intervallradius lässt sich in einen Nahbereich (in diesem Beispiel für Strecken kleiner 100 m) und in den Entfernungsbereich größer 100 m einteilen. Ein wichtiger Beitrag in beiden Bereichen ist der Geräteanteil mit ca. 0.35 mgon. Das Verhalten im Nahbereich wird durch den Prismenanteil gesteuert, der proportional  $\frac{1}{S}$  abklingt, vgl. nebenstehende Abbildung 3.2 und Gleichung 3.1. Für Zielweiten größer 200 m dominiert der linear streckenabhängige Anteil der Refraktion.

Abbildung 3.1(b) zeigt, dass der Intervallradius weitgehend zenitdistanzunabhängig ist. Für die 200 m lange Teststrecke nimmt er eine Größe von  $\zeta_r < 0.8 mgon$  an. Den Haupteinfluss bilden die Geräte- und Refraktionsanteile mit je ca. 0.3 – 0.4 mgon. Abbildung 3.2 zeigt die Beiträge zum Intervallradius für die Messung in einer Lage und Variation der Streckenlänge im Nahbereich. Der dominierende Anteil ist der Prismenanteil, der proportional  $\frac{1}{S}$  abklingt, vgl. (3.1).

Koeffizient beschreibt die Auswirkungen einer Schräganzielung des Prismas und der letzte den Einfluss der Refraktion. Eine ausführliche Diskussion der Koeffizienten der Faustformeln findet sich in Abschnitt 3.3. Dort wird untersucht, wie die Werte der Koeffizienten durch die Wahl geeigneter Instrumente oder Anwendung von Mess- und Auswertemethoden reduziert werden können und welche Werte eine untere Schranke bilden, die mit der aktuellen Technik nicht unterschritten werden kann.

### 3.1.2 Richtungsbeobachtungen

#### Korrektur für die Richtungsbeobachtung

Die korrigierte Richtung  $R$  lässt sich wie folgt darstellen:

$$R = \left( \rho + \frac{c_h}{\sin \zeta_A} + (Q(\rho) + i) \cot \zeta_A \right) - \frac{S}{2 n_A} \frac{\partial n_A}{\partial a^\perp} \varrho - \frac{\sin \alpha_A}{S} \left( \frac{n_A \cos \alpha_A d}{\sqrt{n_R^2 - n_A^2 \sin^2 \alpha_A}} - e \right) \varrho. \quad (3.3)$$

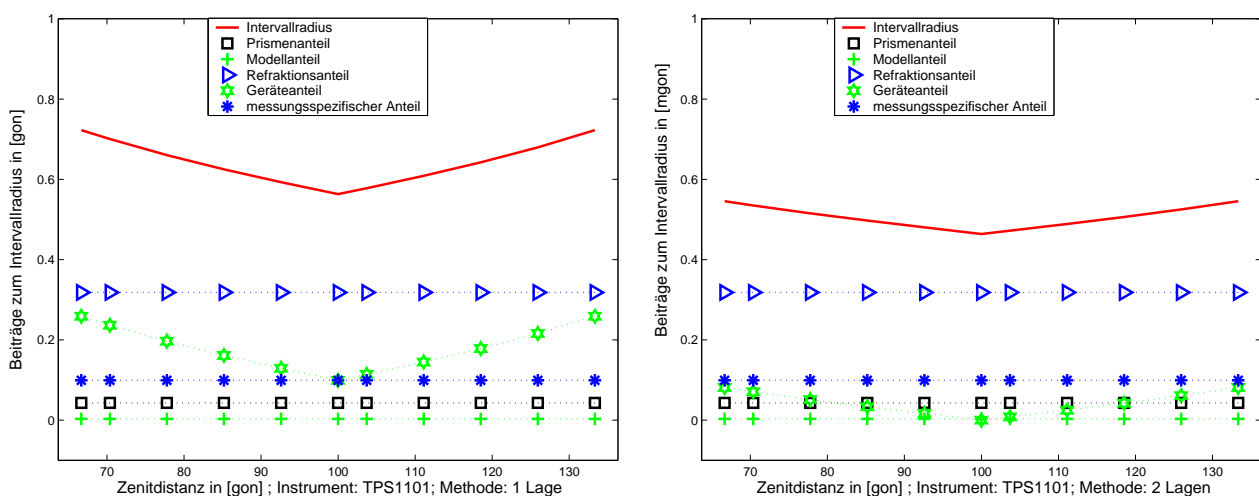
Der erste Term beschreibt die geräteinternen Korrekturen  $\mathbf{k}_S$ . Die einzelnen Einflussparameter sind nochmals aufgeführt:

- $\rho$  beschreibt die rohe Teilkreisablesung,
- $c_h$  den horizontalen Kollimationsfehler,
- $Q(\rho)$  die durch den Zweiachskompensator ermittelte Stehachsneigung in Kippachsrichtung,
- $i$  den Kippachsfehler und
- $\zeta_A$  die am Gerätedisplay angezeigte Zenitdistanz, vgl. (2.88) und
- $\frac{\partial n_A}{\partial a^\perp}$  den konstanten Brechungsindexgradienten senkrecht zur Messungslinie.

Der zweite Term beschreibt den Einfluss der Refraktion. Es wird ein konstanter Brechungsindexgradient  $\frac{\partial n_A}{\partial a^\perp} = -0.1 \pm 0.05 \frac{ppm}{m}$  senkrecht zur Messungslinie angenommen, so dass die Integrale in MORITZ (1961, 1962) bzw. BROCKS (1939) analytisch ausgewertet werden können. Der dritte Term beschreibt den Einfluss einer Schräganzielung des Prismas. Die Bezeichnungen sind identisch mit denen, die im vorherigen Abschnitt eingeführt wurden.

#### Intervallradien beim Messen in zwei Lagen

Die entsprechenden Werte sollen beispielhaft für das Instrument TPS1101 von Leica Geosystems abgeleitet werden. Die Intervalle der Einflussparameter des Instruments wurden in Beispiel 17 eingeführt. Für den Prismenanteil und die atmosphärischen Bedingungen gelten die Voraussetzungen aus Abschnitt 3.1.1. Damit erhält man die folgenden Ergebnisintervalle für Messungen bei Anwendung der Messmethode „Messung in zwei Lagen“.



(a) Messung in einer Lage

(b) Messung in zwei Lagen

Abbildung 3.3: Beiträge zum Intervallradius der Richtungsbeobachtung bei Variation der Zenitdistanz



Abbildung 3.3 zeigt, wie durch die Messung in zwei Lagen der Geräteanteil reduziert werden kann. Teilabbildung (a) stellt die Beiträge zum Intervallradius beim Messen in einer Lage unter Variation der Zenitdistanz und konstanter Streckenlänge von  $S = 200 \text{ m}$  dar. Neben dem Refraktionsanteil ist der zenitdistanzabhängige Geräteanteil ein Haupteinflussfaktor. Teilabbildung (b) zeigt die Beiträge für Messungen in zwei Lagen. Durch Anwendung dieser Messmethode werden die instrumentenspezifischen Beiträge eliminiert, alle anderen Einflussfaktoren können nicht reduziert werden. Die Intervallradien verhalten sich bei Variation der Streckenlänge ähnlich denen der Zenitdistanz, vgl. Abbildung 3.2. Für Zielweiten über  $300 \text{ m}$  bildet der streckenproportionale Refraktionsanteil den Haupteinflussfaktor. Dieser kann nur durch eine bessere Modellierung oder Erfassung der atmosphärischen Bedingungen reduziert werden. Als Faustformel lässt sich für das Messen in einer bzw. zwei Lagen angeben:

$$\begin{cases} R_{r,I} & \approx (0.2 + \frac{8.8}{S} + 1.6 \cdot 10^{-3} S + 0.3 |\cot \zeta|) [mgon], \\ R_{r,II} & \approx (0.1 + \frac{8.8}{S} + 1.6 \cdot 10^{-3} S + 0.2 |\cot \zeta|) [mgon]. \end{cases} \quad (3.4)$$

Der erste und vierte Koeffizient lassen sich geräteinternen Anteilen zuordnen, die durch Messung in zwei Lagen reduziert werden können. Der zweite Koeffizient beschreibt die Auswirkungen einer Schräganzielung des Prismas und der dritte Faktor den Einfluss der Refraktion. Die Faustformeln werden in Abschnitt 3.3 nochmals ausführlich diskutiert.

### 3.1.3 Streckenmessung

#### Korrektionsmodell

Im Gegensatz zu den bisher betrachteten Messungstypen Zenitdistanzen und Richtungen werden bei der Streckenmessung in der Regel keine besonderen Messmethoden angewandt. Die gesamte Reduktion stützt sich also auf durchgreifende Korrekturen und nicht zugängliche und damit nicht modellierbare geräteinterne Aufbereitungsschritte, vgl. RÜEGER (1996) oder BRUNNER UND WOSCHITZ (2001). Die korrigierte Messung ergibt sich mit Berücksichtigung der Refraktion (RÜEGER 1996, S.73ff) und der Streckenverlängerung durch ungenau ausgerichtete Prismen (RÜEGER 1996, S.160ff) zu:

$$\begin{aligned} S = & (1 + n_0 - n_A) \left( \frac{c_0}{2n_0 f_0} \left\{ N + \frac{\Delta\varphi}{2\pi} \right\} + d_0 \right) + \frac{S^3}{24} \left( \left( \frac{\partial n_A}{\partial \alpha^\perp} \right)^2 + \sin^2 \zeta \left( \frac{\partial n_A}{\partial z^*} \right)^2 \right) \\ & - d \left( \frac{n_B}{n_A} - \sqrt{\left( \frac{n_B}{n_A} \right)^2 - \sin^2 \alpha_A} \right) + e (1 - \cos \alpha_A). \end{aligned} \quad (3.5)$$

Der erste Term beschreibt die „erste Geschwindigkeitsreduktion“ angewandt auf die Rohstrecke. Diese ist identisch mit dem Ergebnis der Integration der Formel nach MORITZ (1961) für konstante Brechungsindexgradienten und entspricht der klassischen Geschwindigkeitsreduktion in der linearisierten Form, vgl. RÜEGER (1996, S.75). Folgende Parameter treten auf:

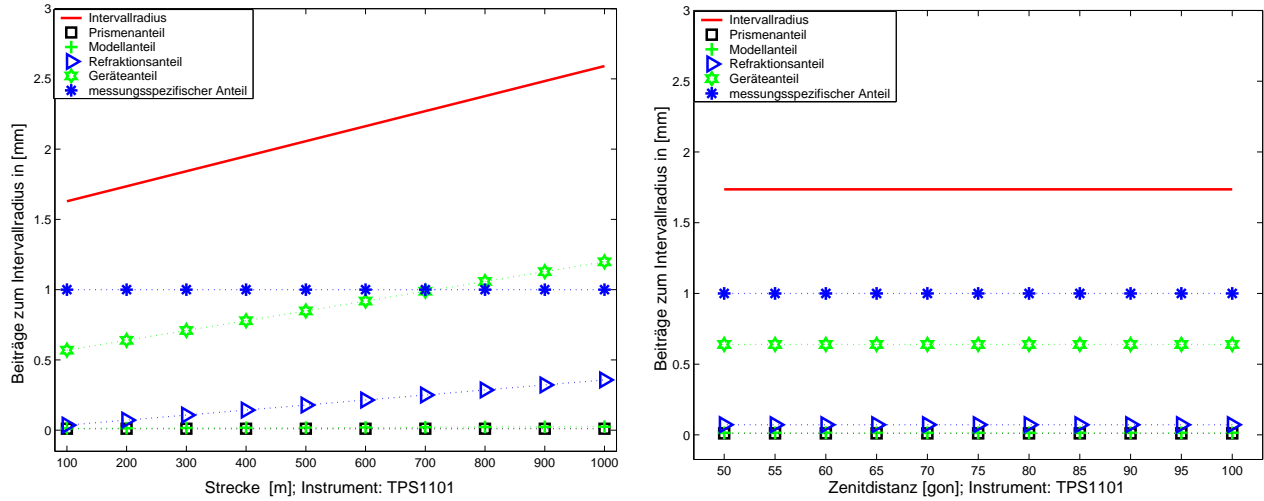
- $c_0$  bezeichnet die Vakuumlichtgeschwindigkeit,
- $n_0, n_A$  den geräteinternen Gruppenbrechungsindex zu den Bezugsdaten des Herstellers bzw. den Brechungsindex der Atmosphäre zu repräsentativen Umgebungsdaten,
- $f_0$  die geräteinterne Modulationsfrequenz,
- $N$  die Phasenmehrdeutigkeit,
- $\Delta\varphi$  das Phasenreststück und
- $d_0$  die Additionskonstante.

Der zweite Term beschreibt die Krümmungsreduktion und zweite Geschwindigkeitsreduktion für die doppelt gekrümmte Lichtkurve. Hierbei bezeichnen  $\frac{\partial n_A}{\partial z^*}$  den Gradienten des Brechungsindex in Richtung des Zenits und  $\frac{\partial n_A}{\partial \alpha^\perp}$  den Gradienten des Brechungsindex senkrecht zur Messungslinie. Der dritte Term erfasst die Streckenverlängerung durch Fehlausrichtung des Prismas, vgl. RÜEGER (1996, S.160ff). Die Bezeichnungen sind in den vorherigen Abschnitten eingeführt worden.

#### Intervallradien für die korrigierte Streckenmessung

Für die sensorspezifischen Einflussparameter werden wieder exemplarisch die Werte für das Instrument TPS1101 und das Prisma GPR1 von Leica Geosystems herangezogen. Für den geräteinterne Brechungsindex gilt

$n_0 = 1.0002830 \pm 0.05 \text{ ppm}$  und für die Wellenlänge  $[\lambda_0] = [770, 790]$ , vgl. RÜEGER (1996, S.31). Bei einer Frequenzstabilität von ca.  $df_0 = 0.5 \cdot 10^{-6} f_0$  (JOECKEL UND STÖBER 1999, S.82f) variiert die Modulationsfrequenz zwischen  $[f_0] = 100 \text{ MHz} \pm 50 \text{ Hz}$ . Die Additionskonstante sei auf  $\pm 0.5 \text{ mm}$  genau bekannt. Rundungsfehler werden zu  $\pm 0.5 \text{ mm}$  abgeschätzt. Diese Werte entsprechen der Hälfte der kleinsten Anzeigeeinheit am Gerätedisplay (LEICA GEOSYSTEMS 1998, S.124). Die Eingangsintervalle für die atmosphärischen Zustandsgrößen werden wie in den vorherigen Abschnitten angesetzt.



(a) Variation der Streckenlänge bei konstanter Zenitdistanz  $\zeta = 90 \text{ gon}$

(b) Variation der Zenitdistanz bei konstanter Strecke  $S = 200 \text{ m}$

Abbildung 3.4: Beiträge zum Intervallradius der Streckenmessung

Abbildung 3.4(a) stellt den Intervallradius bei Variation der Streckenlänge zwischen  $S = 100 \text{ m}$  und  $S = 1000 \text{ m}$  und konstanter Zenitdistanz von  $\zeta = 90 \text{ gon}$  dar. Der Intervallradius wächst linear mit der Streckenlänge. Den Hauptbeitrag liefern die von der Streckenlänge unabhängigen, messungsspezifischen Anteile wie Rest- und Rundungsfehler mit ca.  $1 \text{ mm}$ . Der Geräteanteil ist streckenproportional und stellt den zweitgrößten Beitrag dar.

In Abbildung 3.4(b) ist der Intervallradius bei konstanter Streckenlänge von  $S = 200 \text{ m}$  und variabler Zenitdistanz zwischen  $\zeta = 50 \text{ gon}$  und  $\zeta = 100 \text{ gon}$  aufgetragen. Alle Einflussfaktoren sind von der Zenitdistanz unabhängig. Als eine Faustformel lässt sich der Intervallradius für die Streckenmessung mit diesem spezifischen Instrument und unter den angenommenen äußeren Bedingungen abschätzen:

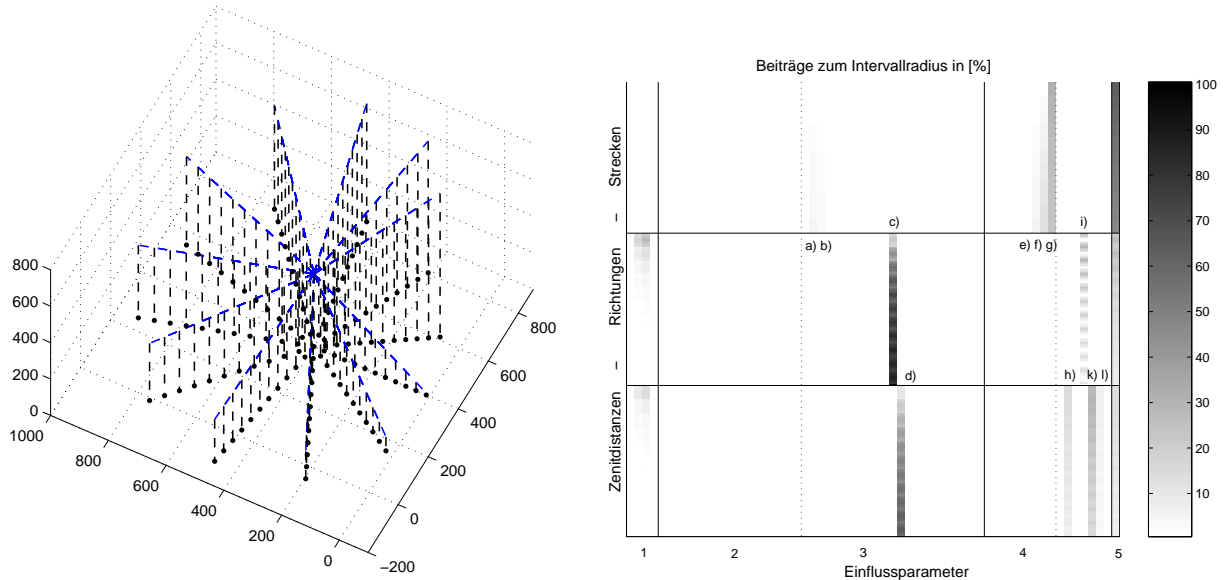
$$S_r \approx (1.5 + 1 \cdot 10^{-3} S) [\text{mm}]. \quad (3.6)$$

## 3.2 Analyse der Haupteinflussfaktoren

Im ersten Abschnitt dieses Kapitels wurden die Einflüsse nach Messungstyp getrennt dargestellt und analysiert. Eine weitere wichtige Fragestellung ist, ob unabhängig vom Messungstypen Haupteinflussfaktoren existieren. Diese Vermutung ist naheliegend, da in modernen Vermessungsinstrumenten Richtungen, Strecken und Zenitdistanzen gemeinsam gemessen werden und durch geräteinterne Korrektionsfunktionen miteinander verknüpft sind. Zusätzlich sind, wie in den Abschnitten 3.1.1 bis 3.1.3 dargelegt wurde, die Korrektionsmodelle für die Refraktion und Fehlausrichtung von Prismen bei allen Messungstypen ähnlich.

Als Analysegröße soll der relative Beitrag  $\psi_{i,j}^*$  eines Einflussparameters  $p_j$  zum Intervallradius  $l_{r,i}$  nach (2.87) herangezogen werden. In der folgenden Simulation wird ein Testgitter gerechnet, das aus elf gleichmäßig über den Horizont verteilten Messungslinien besteht. In dessen Mitte befindet sich der Instrumentenstandpunkt, von dem aus 121 Gitterpunkte polar angehängt werden, vgl. Abbildung 3.5(a). Die Zenitdistanz, unter der die Punkte einer Messungslinie bestimmt werden, ist konstant. Von Messungslinie zu Messungslinie variiert die Zenitdistanz um ca.  $9 \text{ gon}$ , so dass insgesamt ein Sektor von  $50 \text{ gon}$  bis  $150 \text{ gon}$  überdeckt wird. Der Abstand zweier Punkte einer Messungslinie beträgt  $50 \text{ m}$ , wobei durch die Messungslinie ein Entfernungsbereich von  $50 \text{ m}$

bis 550 m simuliert wird. Es wird das Instrument TPS1101 mit Prisma GPR1 von Leica Geosystems modelliert. Die Schräganzielung des Prismas soll 10 gon betragen. Weiter wird das Refraktionsmodell nach MORITZ (1961) mit einem konstanten Brechungsindexgradienten angesetzt. Die Richtungsbeobachtungen werden in zwei Lagen gemessen, die Zenitdistanzen nur in einer Lage.



(a) Testnetz

(b) Darstellung der Gruppen der Einflussparameter gemäß Tabelle 3.1: 1 Prismenanteil - 2 Modellanteil - 3 Refraktionsanteil - 4 Geräteanteil - 5 messungsspezifischer Anteil

Abbildung 3.5: Beiträge zum Intervallradius in [%] bei Messungen im Testfeld für Strecken, Richtungen und Zenitdistanzen.

Abbildung 3.5(b) zeigt den prozentualen Beitrag der Einflussparameter bezogen auf den Intervallradius. Die Strecken-, Richtungs- und Zenitdistanzanteile bilden je einen Block über alle Zeilen hinweg. Jede Spalte stellt den relativen Beitrag eines Einflussparameters dar, die in insgesamt fünf Gruppen zusammengefasst wurden, vgl. Tabelle 3.1. Unterschiedliche Zielweiten und Zenitdistanzen im Testnetz führen nach den Faustformeln (3.2), (3.4) und (3.6) zu unterschiedlich großen Intervallradien und bedingen die deutliche Variation der Grauwerte in Abbildung 3.5(b). Die folgende Tabelle gibt einen Überblick über die dominierenden Einflussparameter.

3. Refraktionsanteil	a) Trockentemperatur $\vartheta$
	b) Druck $p$
	c) Brechungsindexgradient $\frac{\partial n_A}{\partial \alpha \perp}$
	d) Brechungsindexgradient $\frac{\partial n_A}{\partial z^*}$
4. Geräteanteil	e) Trägerwellenlänge der EDM $\lambda$
	f) Frequenz $f_0$
	g) Additionskonstante $d_0$
	h) Vertikaler Kollimationsfehler $c_v$
	i), k) Quer- und Längsanteil des Stehachsfehlers $Q(\rho), L(\rho)$
	l) Indexfehler $i_{E_x}$

Tabelle 3.2: Übersicht über die dominierenden Einflussparameter

Eine Gruppe wichtiger Einflussparameter sind die messungsspezifischen Rest- und Rundungsfehler (Gruppe 5). Sie liefern 47 – 64% (durchschnittlich 54%) der durch den Intervallradius beschriebenen Impräzision für Strecken, für die in zwei Lagen gemessenen Richtungen 3 – 29% (durchschnittlich 15%) und für die in einer Lage gemessenen Zenitdistanzen ca. 7 – 15% (durchschnittlich 11%). Im Wesentlichen wird durch die Rest- und

Rundungsfehler die Instrumentenklasse bestimmt, so dass durch den Einsatz eines höherwertigen Instrumentes diese Anteile reduziert werden.

Weitere wesentliche Einflussparameter sind die gerätespezifischen Anteile (Gruppe 4): für Strecken 34 – 42% (durchschnittlich 39%), für Richtungen bis 29% (durchschnittlich 11%) und für Zenitdistanzen 26 – 54% (durchschnittlich 39%). Für Richtungen hat diese Gruppe den geringsten Einfluss, da die Beiträge - bis auf den Stehachsfehler - bereits durch eine geeignete Messmethode (Messung in zwei Lagen) eliminiert wurden. Der Beitrag für den Intervallradius der Zenitdistanz kann ebenso durch Messung in zwei Lagen reduziert werden. Lediglich bei der Streckenmessung werden in der Regel keine Messmethoden zur Verringerung angewandt. Im Hinblick auf die Reduktion der Impräzision wäre es sinnvoll, die Bildung von Streckenverhältnissen als Messmethode neu zu beleuchten.

Der Anteil der Refraktionsmodelle und der zusätzlichen meteorologischen Messungen (Gruppe 3) liegt bei Richtungen zwischen 16 – 96% (durchschnittlich 64%), bei Strecken 1 – 9%, (durchschnittlich 6%) und bei Zenitdistanzen bis maximal 65% (durchschnittlich 42%). Für die Streckenmessung wird der Einfluss durch die Impräzision der Temperatur- und Druckerfassung bestimmt, wohingegen bei Richtungs- und Zenitdistanzbeobachtungen die Impräzision der Bestimmung des Brechungsindexgradienten den dominanten Anteil ausmacht. Soll die Impräzision weiter reduziert werden, so ist das Hauptaugenmerk auf eine bessere Bestimmung des Brechungsindexgradienten zu richten. Konkrete Ansätze wurden hierzu beispielsweise von HENNES ET AL. (1999) vorgestellt.

Modellkonstanten (Gruppe 2) liefern für alle Messungstypen Anteile bis maximal 1%, d.h. eine Verbesserung oder Erweiterung der Modelle wird nur zu einer geringfügigen Reduktion der Impräzision via besser bestimmter Modellkonstanten führen. Als Sekundäreffekt wird sich allerdings der Fortpflanzungsmechanismus für alle weiteren Anteile ändern.

Schließlich liefert der Prismenanteil (Gruppe 1) weniger als 1% für Strecken, je nach Zielweite zwischen 0 – 48% für Richtungen und 1 – 25% für Zenitdistanzen. Diese Effekte können reduziert werden, wenn die Prismen mit höherer Präzision auf das Ziel ausgerichtet werden. Bei fest installierten Prismen wie in Überwachungsnetzen, wo eine individuelle Ausrichtung nicht möglich ist, können diese Effekte nicht reduziert werden.

## 3.3 Diskussion der Koeffizienten der Faustformeln

### 3.3.1 Vorbemerkungen

Im folgenden Abschnitt werden die Koeffizienten der Faustformeln (3.2), (3.4) und (3.6) diskutiert. Die Werte der Koeffizienten hängen vom Instrumentarium sowie von den verwendeten Messmethoden, Korrektionsmodellen und Auswerteverfahren ab. Die folgende Darstellung soll darüber einen exemplarischen Überblick geben und zeigen, welche „Variationsbreite“ für die einzelnen Koeffizienten zu erwarten ist. Die Beantwortung dieser Fragestellung ist für die Optimierung der Intervallradien (*Intervalldesign*) wesentlich, die in Abschnitt 5.4 vorgestellt wird. Sie liefert zum Einen Anhaltspunkte, wie die Werte der Koeffizienten der Faustformeln vom gewählten Instrumentarium, den Auswerte- und Messmethoden abhängen. Zum Anderen werden minimale Werte für die Koeffizienten angegeben, die durch kein Messverfahren oder Instrumentarium unterschritten werden können.

Diese unteren Grenzen für die Werte der Koeffizienten werden im Allgemeinen durch zwei begrenzende Faktoren bestimmt. Einerseits liefert die *Auflösung der Sensoren* eine physikalische Grenze, die mit dem aktuellen Stand der Technik nicht unterschritten werden kann. Andererseits kann bei hochaufgelösten Daten die *Repräsentativität der Größen* verloren gehen. Beispielsweise können mit hochauflösenden Temperatursensoren Werte bestimmt werden, die nur zum Moment der Erfassung gelten, da die Temperatur in der Größenordnung der hohen Auflösung von  $\pm 0.03, \dots, 0.1 K$  (HENNES ET AL. 1999) schnellen Fluktuationen unterworfen ist. Turbulenzen, kurzzeitige Variationen und im Vergleich dazu lange Ansprechzeiten der Sensoren lassen Schwankungen von mehreren Hundertstel bis Zehntel Grad zu (HENNES ET AL. 1999). Ebenso spielen weitere Effekte eine Rolle wie Erwärmungseffekte der Übertragungskabel ( $\pm 0.2K$ ), vgl. DEUSSEN (2000, S.85ff). Die allgemeine Problematik der Temperaturmessung, die auch für moderne Sensoren zu berücksichtigen ist, wird von GOTTWALD (1985, S.45-88) diskutiert.

Die Koeffizienten der Faustformeln für Strecken, Richtungen und Zenitdistanzen sollen folgendermaßen in einem Vektor  $\kappa$  geordnet sein:

$$\kappa = \begin{pmatrix} a_S [mm], & c_S [ppm], & \dots \\ a_R [mgon], & b_R [mgon\ m], & c_R [\frac{10^{-3} mgon}{m}], & d_R [mgon], \\ \dots & a_\zeta [mgon], & b_\zeta [mgon\ m], & c_\zeta [\frac{10^{-3} mgon}{m}] \end{pmatrix}^T. \quad (3.7)$$

### 3.3.2 Koeffizienten der Faustformel für die Streckenmessung

#### Konstanter Anteil: $a_S$

Der konstante Anteil der Faustformel für den Intervallradius der Streckenmessung wird durch die Impräzision der *Additionskonstanten* und durch die *Rest- und Rundungsfehler* des Instrumentes dominiert.

- Die Impräzision der *Additionskonstante* kann durch Kalibrierverfahren bestimmt und somit verbessert werden. Gegebenenfalls ist das Modell einer Additionskonstanten um nichtlineare Korrekturterme zu erweitern, wie dies beispielsweise von HENNES ET AL. (1994) diskutiert wurde.
- Für die *Rest- und Rundungsfehler* aller Messungstypen (Strecken, Richtungen, Zenitdistanzen) wurde in der vorliegenden Arbeit angenommen, dass diese die Hälfte der letzten Nachkommastelle des am Geräte-display anzeigbaren Wertes ausmachen. Um diese Hypothese für verschiedene Instrumententypen nachzuweisen, müsste die Auflösung der Instrumente kontrolliert werden. MEIER-HIRMER UND MEIER-HIRMER (1997) haben beispielsweise gezeigt, dass die letzte am Display anzeigbare Nachkommastelle in dieser Hinsicht nicht unbedingt für die Qualität des Instruments aussagekräftig ist, wenn sie durch gezielte geräteinterne Rundung auf eine gewisse Nachkommastelle bedingt ist. Im Folgenden wird dennoch angenommen, dass die Größe der Rest- und Rundungsfehler von der Instrumentenklasse abhängt, d.h. dass ein Unterschied in der Qualität der Sensoren besteht. Durch die Wahl eines Instruments, das einer höherwertigen Instrumentenklasse angehört, kann somit der Anteil der Rest- und Rundungsfehler reduziert werden.

#### Streckenlängenabhängiger Anteil: $c_S$

Der Koeffizient  $c_S$  beschreibt den Einfluss der Anteile, die von der Streckenlänge abhängen. Im einfachen Modell, vgl. Gleichung (3.5), das diesen Untersuchungen zu Grunde gelegt wurde, wird die Größe des Koeffizienten von zwei Hauptfaktoren beeinflusst: *Geräteanteile* und *Refraktionsanteile*.

- Die Geräteanteile sind durch die Impräzision der *Modulationsfrequenz* und der *Trägerwellenlänge* bestimmt. Betrachten wir zunächst die Impräzision der Modulationsfrequenz.
  - JOECKEL UND STOBER (1999, S.47f) geben die Frequenzstabilität auf  $\pm 10 ppm$  an. Nach Korrektur der Temperatureffekte werden Werte  $\pm 1 ppm$  erreicht. Zusätzlich spielen Einlaufeffekte und Alterungsprozesse eine Rolle, vgl. beispielsweise RÜEGER (1996, S.33f) und JOECKEL UND STOBER (1999, S.48f). Für das Mekometer 5000 wird im Handbuch eine Frequenzstabilität von  $\pm 0.3 ppm$  angegeben (KERN AG 1986, S.4). Für die folgenden Untersuchungen wird eine Impräzision von  $df = \pm 1 ppm f_0$  als untere Grenze angesetzt, falls keine anderen Zahlenwerte für die Instrumente vorliegen. Einer Reduktion dieses Beitrags wirkt entgegen, dass die Messfrequenz  $f_0$  bei vielen modernen Geräten erhöht wurde, um die Auflösung der Phasenmessung zu steigern, vgl. RÜEGER (1996, S.233ff). SCHWARZ (1993, S.95) gibt diese mit  $1 - 2 \cdot 10^{-4}$  an, was  $2 - 4 mm$  bei  $f_0 = 15 MHz$  und  $0.1 mm$  bei  $f_0 = 100 MHz$  entspricht. Mit einer Erhöhung der Messfrequenz nimmt die Auflösung zu, aber auch die Impräzision  $df$ .
  - Die *Trägerwellenlänge* wird zur Berechnung der Refraktionskorrektur benötigt (JOECKEL UND STOBER 1999, S.71ff); sie ist geräteabhängig. Die Impräzision die Trägerwellenlänge wird von der Güte der Diode bestimmt, die verwendet wird. RÜEGER (1996, S.54) nennt dabei Abweichungen von der Nominalwellenlänge zwischen  $10 - 20 nm$  für GaAlAs-IR-Dioden. Bei stabilisierten Gaslasern (HeNe-Laser) sind Abweichungen von  $\pm 5 nm$  zu erwarten (RÜEGER 1996, S.54). Des weiteren spielt beispielsweise die Bandbreite des Spektrums eine Rolle. Für Laserdioden liegen die Bandbreiten, bei denen die Leistung auf 50% abgesunken ist, bei  $2 - 4 nm$ , für Infrarot um  $40 nm$ . Weitere Einflüsse spielen Alterungs- und Erwärmungsprozesse. RÜEGER (1996, S.37) gibt hierfür  $0.25 \frac{nm}{K}$  an. Für die weiteren Untersuchungen soll daher angenommen werden, dass die Wellenlänge minimal um  $\pm 10 nm$  von der Nennwellenlänge abweicht.

- Die *Refraktionsanteile* wirken sich hauptsächlich bei der Bestimmung der sogenannten ersten Geschwindigkeitskorrektur aus, vgl. erster Term in (3.5). Neben der Trägerwellenlänge, die als Geräteanteil behandelt wurde, spielt die Impräzision der Temperatur und des Drucks eine Rolle, vgl. auch Abbildung 3.5 in Abschnitt 3.2. Diese Größen werden gebraucht, um die Ergebnisse der rohen Streckenmessung auf die aktuellen meteorologischen Rahmenbedingungen zu beziehen. Die Temperatur kann messtechnisch auf  $\pm 0.01 - 0.03 K$  (HENNES ET AL. 1999) bestimmt werden. Allerdings geht in diesem Bereich die Repräsentativität der Werte verloren, worauf eingangs schon hingewiesen wurde. Für die folgende Diskussion wird eine minimale Impräzision von  $\pm 0.05 K$  angesetzt. Eine weitere Steigerung ist auch durch das Modell des Brechungsindex limitiert. Modellfehler und Beschränkungen vermindern die Repräsentativität des Barrell und Sears Modells. vgl. DEICHL (1984) oder RÜEGER (1999). Die IAG-Resolution 3 (SCHWARZ 2000, S.66) nennt verbesserte Modellkonstanten für eine Berechnung des Brechungsindex auf  $\pm 0.5 ppm$ . Um eine noch höhere Repräsentativität zu erhalten, sollte das Owns-Modell in der sogenannten Sellmeier-Form verwendet werden, vgl. RÜEGER (1996, S.222ff).

Die folgende Tabelle gibt eine Übersicht über minimale Werte für die Koeffizienten der Faustformel der Streckenmessung bei einer exemplarischen Instrumentenauswahl.

Instrument	$f_0 \pm df$ [MHz ± Hz]	$\lambda \pm d\lambda$ [nm]	kleinste Einheit [mm]	$a_S$ [mm]	$c_S$ [ppm]
Leica TPS1101	$100 \pm 50$	$780 \pm 15$	1	1.5	1
Leica TCA2003	$50 \pm 25$	$850 \pm 15$	0.1	0.17	0.9
Mekometer 5000	$10 \pm 3$	$632.6 \pm 0.5$	0.1	0.15	0.75

Tabelle 3.3: Minimale Werte der Koeffizienten der Faustformel der Streckenmessung für eine exemplarische Instrumentenauswahl

### 3.3.3 Koeffizienten der Intervallradien der Richtungsbeobachtung

#### Konstanter Anteil: $a_R$

Der Koeffizient  $a_R$  beschreibt den konstanten Beitrag zum Intervallradius der Richtungsbeobachtung. Der Wert der Koeffizienten wird hauptsächlich von den modellierten *Geräteanteilen* (außer dem Stehachsfehler) und den *Rest- und Rundungsfehlern* beschrieben. Betrachten wir die Effekte im Einzelnen:

- Für die *Rest- und Rundungsfehler* wird angenommen, dass diese Effekte maximal bis zur Hälfte der letzten am Display anzeigbaren Nachkommastelle auflösbar sind. Restfehler können reduziert werden, falls durch Kalibrierung der Instrumente verbesserte Korrektionsmodelle anstelle des hier zugrunde gelegten einfachen Modells, vgl. Gleichung (3.3), bestimmt werden können. Ist dies nicht möglich, so bleibt nur die Wahl eines höherwertigen Instruments, um die Werte der Koeffizienten zu reduzieren. Diese Wahl ist mit entsprechend höheren Anschaffungskosten verbunden.
- Die *Geräteanteile* können durch geeignete Messverfahren (Messen in zwei Lagen) reduziert werden. Ein Vergleich der Faustformeln (3.4) für die Messung in einer Lage und die Messung in zwei Lagen verdeutlicht dies. Von Seiten der Messmethoden ist damit allerdings der gesamte Reduktionsspielraum ausgeschöpft.

Instrument	$dQ(\rho)$ [mgon]	$di$ [mgon]	$dc_h$ [mgon]	$a_{R,I}$ [mgon]	$a_{R,II}$ [mgon]
TPS1101	0.2	0.05	0.1	0.2	0.1
TCA2003	0.1	0.005	0.01	0.1	0.01

Tabelle 3.4: Minimale Werte des Koeffizienten  $a_R$  für eine exemplarische Instrumentenauswahl

Tabelle 3.4 gibt exemplarisch für zwei Instrumente die Koeffizienten bei zwei unterschiedlichen Messmethoden (Messung in einer Lage und Messung in zwei Lagen) in den letzten beiden Spalten an. Der Wert für die Impräzision der Stehachsneigung in Kippachsrichtung  $dQ(\rho)$  wurde als Auflösung des Stehachskompensators aufgefasst, die in den Gerätehandbüchern verzeichnet ist (LEICA GEOSYSTEMS 1998, S.126) bzw. (LEICA GEOSYSTEMS 2000, S.241). In der dritten und vierten Spalte sind die angenommenen Werte für die Impräzision des Kippachsfehlers  $di$  und des horizontalen Anteils des Zielachsfehlers  $dc_h$  angegeben.

**Prismenanteil:  $b_R$** 

Der Koeffizient  $b_R$  beschreibt den Beitrag, der sich antiproportional zu  $S$  verhält. Der Wert des Koeffizienten ist durch den Prismenanteil bestimmt. Hintergrund für diese Diskussion sind hauptsächlich Anwendungen der automatischen Überwachung von Bauwerken, bei denen Prismen fest eingebaut sind. Werden sie von mehreren Standpunkten aus leicht unterschiedlichen Richtungen angezielt, so können die Prismen nicht exakt auf alle Standpunkte ausgerichtet sein, so dass Schräganzielungen auftreten.

Von der instrumentellen Entwicklung her betrachtet, sind Reflektoren für die Streckenmessung entwickelt und optimiert worden. So lassen sich durch die spezielle Konstruktionsmaße der Prismen

$$d : e \approx 3 : 2 \approx n_R : n_A \quad (3.8)$$

Abweichungen von nicht senkrecht auftreffenden Messstrahlen minimieren, vgl. RÜEGER (1996, S.158ff). Dabei bezeichnet  $d$  den Abstand der Prismengrundfläche von der Prismenspitze und  $e$  den Abstand der Prismengrundfläche von der Prismenstehachse. Der Brechungsindex des Prismas wird mit  $n_R$  benannt und der Brechungsindex der Atmosphäre mit  $n_A$ . Bis zu Verschwenkungen von  $\pm 15 \text{ gon}$  ist der Einfluss auf die Streckenmessung für Prismen, die (3.8) erfüllen, zu vernachlässigen ( $< 0.1 \text{ mm}$ ).

Werden die selben Prismen auch für die Richtungsbeobachtung oder Zenitdistanzmessung verwendet, so hängt die Sensitivität ebenfalls vom Prismendesign (3.8) ab. Allerdings nimmt der Effekt durch die Streckenabhängigkeit größere Werte im Verhältnis zur Auflösbarkeit der Winkelmessung an, als dies der Fall bei der Streckenmessung ist, vgl. RÜEGER (1996, S.163) und Gleichungen (3.1) bzw. (3.3) mit (3.5). Betrachtet man die Impräzision der Korrektionsformel, vgl. (3.3) letzter Term, so spielen neben der Impräzision der geometrischen Abmessungen, die auf  $\pm 0.5 \text{ mm}$  (RÜEGER 1996, S.257) bzw. für das Mekometer-5000-Prisma auf  $\pm 0.05 \text{ mm}$  abgeschätzt werden, auch die Impräzision der Ausrichtung eine Rolle. RÜEGER (1996, S.163) gibt dafür  $\pm 0.7 \text{ mgon}$  an, KAHMEN (1993, S.207f) fordert  $\pm 0.3 \text{ gon}$  im Kontext der Ausrichtung von Basislatten. Alle genannten Effekte werden verstärkt, wenn der Einfallswinkel  $\alpha_A$  größer wird.

Folgende *Lösungsstrategien* stehen für eine Reduktion der Effekte zur Verfügung: Besteht die Möglichkeit, mehrere Prismen am Zielort anzubringen und diese auf die Standpunkte exakt auszurichten, so ist dieser Ansatz zu bevorzugen, da in diesem Fall die Impräzision minimiert wird. Für eine mögliche technische Umsetzung vergleiche LUTES (2001). Allerdings müssen dann die Exzentren zwischen den Prismenreferenzpunkten hinreichend genau bestimmbar sein, vgl. FLACH UND NATEROP (1999). Prinzipiell sollten Prismen verwendet werden, die dem Kriterium (3.8) genügen. Konkrete Fabrikate sind beispielsweise bei RÜEGER (1996, S.257) angegeben. Rundumprismen eignen sich für die hohen Anforderungen bei Überwachungsmessungen nur bedingt, da systematische Effekte je nach Ausrichtung des Prismas auftreten, vgl. FAVRE UND HENNES (2000). Alternativ zu Prismen können Zieltafeln oder Zielzeichen (Kegel) verwendet werden, die von allen Seiten aus anzielbar sind. Für automatische Überwachung ist diese Alternative jedoch ebenfalls wenig geeignet.

Die folgende Tabelle bietet eine Übersicht über die Größenordnung der Koeffizienten  $b_R$  bei verschiedenen Einfallswinkel (senkrechter Einfall:  $\alpha_A = 0$ ) und Prismengeometrien (RÜEGER 1996, S.257). Weiter wurde  $da = de = \pm 0.5 \text{ mm}$  und  $d\alpha_A = 0.7 \text{ gon}$  angenommen. In der zweiten und dritten Spalte sind die Kenngrößen der Prismengeometrie zusammengestellt. In der vierten bis siebten Spalte sind die Werte für die Koeffizienten bei entsprechenden Einfallswinkeln aufgelistet. In der letzten Spalte ist approximativ eine lineare Funktion für den Zusammenhang  $\underline{b}_R(\alpha_A)$  angegeben.

Prisma	d [mm]	e [mm]	Einfallswinkel [gon]				$\underline{b}_R(\alpha_a)$ [mgon m]
			0	4.1	8.3	12.5	
Leica GPR1	39.6	26.6	0.0	3.5	7.1	10.8	$0 + 0.9 \alpha_A$
TOPCON	48	72	28.7	53.8	78.8	103.3	$29 + 6 \alpha_A$
AGA NEW	40.5	62	24.6	46.4	68	89	$25 + 5 \alpha_A$
Zeiss TK 7	40	27	0.3	20.5	40.5	60.5	$0.3 + 5 \alpha_A$

Tabelle 3.5: Variation des Koeffizienten  $b_R$  in [mgon m] in Abhängigkeit der Prismengeometrie und des Einfallswinkels für eine exemplarische Instrumentenauswahl

**Streckenlängenabhängiger Anteil:  $c_R$** 

Der Koeffizient  $c_R$  beschreibt den Anteil der Faustformel, der linear von der Streckenlänge abhängt. Dieser gibt den Einfluss der *Seitenrefraktion* auf die Richtungsbeobachtung an. Mit der Approximation des Brechungsindex

dexgradienten (SCHWARZ 1993, S.51) erhält man die Impräzision des Brechungsindexgradienten in Funktion der Impräzision des Temperaturgradienten senkrecht zur Messungslinie:

$$\frac{\partial n_A}{\partial \mathbf{a}^\perp} \approx -1.10^{-6} \frac{\partial \vartheta}{\partial \mathbf{a}^\perp} \Rightarrow d \frac{\partial n_A}{\partial \mathbf{a}^\perp} \approx -1.10^{-6} d \frac{\partial \vartheta}{\partial \mathbf{a}^\perp}. \quad (3.9)$$

Die Impräzision des Temperaturgradienten und somit auch des Brechungsindexgradienten hängt zum Einen von der *messtechnischen Auflösbarkeit* und der *Repräsentativität der Temperaturmessungen* ab. Zum Anderen wird der Gradient als Differenzenquotient approximiert, so dass ein *Approximationsfehler* zu berücksichtigen ist (DEUSSEN 2000, S.87ff). WITTE (1990) nennt eine Bestimmbarkeit des Temperaturgradienten  $\frac{\partial \vartheta}{\partial \mathbf{a}^\perp}$  von  $\pm 0.1 \frac{K}{m}$ . Die prinzipielle messtechnische Realisierung ist beispielsweise in GOTTWALD (1985) vorgestellt. Für den Alptransit-Tunnel wurde ein Verfahren von HENNES ET AL. (1999) vorgestellt, das Werte bis  $\pm 0.06 \frac{K}{m}$  liefert. Diese geringen Werte für die Impräzision setzen konstante Rahmenbedingungen voraus und sind mit entsprechendem Messaufwand verbunden. Die Ergebnisse umfangreicher Untersuchungen für die Seitenrefraktion sind in WILHELM (1994) zusammengefasst. Die Auflösung des Temperaturgradienten bestimmt die Impräzision der Refraktionskorrektur, die nicht unterschritten werden kann, vgl. Abbildung 3.5. Ähnliches gibt auch für die Vertikalrefraktion. Gleichzeitig wird in gewisser Weise die Problematik der Bestimmbarkeit auf die Problematik der Repräsentativität verlagert.

Eine Alternative, die in Zukunft für Anwendungen zur Verfügung stehen dürfte, ist die Verwendung refraktionsfreier Messverfahren. Die entwickelten Verfahren nutzen entweder optische Turbulenzeffekte (Intensitätsfluktuationen beim Szintillometer bzw. Bildtanzern oder Bildflimmern bei Auswertungen mit CCD-Sensoren) oder Dispersionseffekte, vgl. beispielsweise für eine Übersicht BÖCKEM ET AL. (2000). Dadurch kann der Einfluss der Refraktion integral über den gesamten Messweg bestimmt werden, vgl. DEUSSEN (2000).

Für die weitere Diskussion wird davon ausgegangen, dass der Brechungsindexgradient nicht besser als  $\pm 0.05 \frac{ppm}{m}$  angegeben werden kann, was eine recht optimistische Größenordnung ist. Der Koeffizient nimmt dann den Wert  $c_R = 0.16 \frac{10^{-3} \text{mgon}}{m}$  an.

### Zenitdistanzabhängiger Anteil: $d_R$

Der Koeffizient  $d_R$  beschreibt die Geräteanteile (*Kippachsfehler* und *Stehachsfehler*), die mit dem Kotangens der Zenitdistanz variieren. In der vorliegenden Arbeit wird - abweichend von der klassischen Einteilung in der geodätischen Instrumentenkunde - der Stehachsfehler zu den Instrumentenfehlern gezählt, da er bei modernen Geräten maßgeblich durch die Auflösung und Korrektheit des Stehachskompensators bestimmt wird.

Der Anteil des *Stehachsfehlers* wird durch die Auflösung des Stehachskompensators bei elektronischen Geräten bzw. durch die Angabe der Röhrenlibelle bei mechanischen Instrumenten beschränkt, welche ungefähr die verbleibende Restschiefe der Stehachse angeben. Die Auflösung des Kompensators könnte instrumentell gesteigert werden. Allerdings ist dabei eine erhöhte Temperaturempfindlichkeit oder eine längere Einschwingzeit des Kompensators zu erwarten. Der zweite Effekt wirkt sich besonders störend bei Anwendungen aus, bei denen in schneller Abfolge Messungen in unterschiedlichen Azimuten ausgeführt werden sollen, beispielsweise bei der Fahrzeugsteuerung. Der Anteil des *Kippachsfehlers* kann durch eine geeignete Messmethode (Messung in zwei Lagen) eliminiert werden.

Die folgende Tabelle gibt exemplarisch für zwei Instrumente unterschiedlicher Leistungsklassen die Koeffizienten an. Die Auflösung des Kompensators ist den Handbüchern LEICA GEOSYSTEMS (1998, S.126) bzw. LEICA GEOSYSTEMS (2000, S.241) entnommen. Die Impräzision des Kippachsfehlers  $d_i$  wurde zur Hälfte der letzten, am Display anzeigbaren Stelle angesetzt.

Instrument	$dQ(\rho)$ [mgon]	$d_i$ [mgon]	$\underline{d}_{R,I}$ [mgon]	$\underline{d}_{R,II}$ [mgon]
TPS1101	0.2	0.05	0.3	0.2
TCA2003	0.1	0.005	0.2	0.1

Tabelle 3.6: Minimale Werte des Koeffizienten  $d_R$  für eine exemplarische Instrumentenauswahl





**Beispiel 18** Betrachten wir zunächst die Berechnung der Brechzahl  $N_G$  unter Berücksichtigung der Dispersion nach Barrell und Sears in der sogenannten *Cauchy Reihenentwicklung*, vgl. JOECKEL UND STOBER (1999, S.71ff):

$$N_G = A + 3\frac{B}{\lambda^2} + 5\frac{C}{\lambda^4} + \dots \quad (3.11)$$

Die Brechzahl des Medium hängt von der Trägerwellenlänge  $\lambda$  und den drei Konstanten  $A$ ,  $B$ ,  $C$  ab. Unter der Annahme, dass diese physikalisch korrekt angegeben wurden, kann die Impräzision durch folgende Werte abgeschätzt werden:  $[A] = [287.6035, 287.6045]$ ,  $[B] = [1.62875, 1.62885]$  und  $[C] = [0.01355, 0.01365]$ . Typische Angaben für Entfernungsmessgeräte sind  $[\lambda] = [770, 790]$  nm (RÜEGER 1996, S.54). Für die Funktion  $N_G = N_G(A, B, C, \lambda)$  stimmen Wertebereich und Intervallauswertung überein, und man erhält das Ergebnisintervall:

$$\mathcal{W}([N_G]) = [N_G] = \left[ \underline{A} + 3\frac{\underline{B}}{\underline{\lambda}^2} + 5\frac{\underline{C}}{\underline{\lambda}^4}, \quad \overline{A} + 3\frac{\overline{B}}{\overline{\lambda}^2} + 5\frac{\overline{C}}{\overline{\lambda}^4} \right]. \quad (3.12)$$

	$N_G$	$\overline{N_G}$	$N_{G,m}$	$w([N_G])$
Sensitivitätsanalyse	295.6007	296.0378	295.8192	0.4370
Intervallauswertung	295.6050	296.0423	295.8236	0.4372
Wertebereich	295.6050	296.0423	295.8236	0.4372

Tabelle 3.8: Vergleich der Approximation des Wertebereichs für die Brechzahl  $N_G$ , alle Werte in [ppm]

Tabelle 3.8 zeigt den Vergleich der beiden Approximationen des Wertebereichs. In der zweiten und dritten Spalte sind die untere bzw. obere Intervallgrenze angegeben. In der vierten und fünften Spalte wird die Intervallmitte und der Durchmesser des Intervalls dargestellt. Die Ergebnisse der Sensitivitätsanalyse stimmen mit der Intervallauswertung bzw. dem Wertebereich in der Größenordnung  $10^{-8}$  überein. Die Hausdorff-Distanz (2.17) zwischen beiden Mengen beträgt:  $d = 4.5 \cdot 10^{-9}$ . Diese Größenordnung stellt eine ausreichende Approximation für alle Messoperationen dar.

**Beispiel 19** Als zweites Beispiel wird die Bestimmung des Sättigungsdampfdrucks  $p_w$  mittels relativer Feuchte  $RF$  und Temperatur  $\vartheta$  nach der Formel von *Magnus-Tetens* (RÜEGER 1996, S.63ff) dargestellt

$$p_w = \frac{RF}{100} 10^{\frac{c_1\vartheta}{c_2+\vartheta}+c_3}. \quad (3.13)$$

Die Intervallauswertung wird untersucht bei  $[\vartheta] = [19.8, 20.2]^\circ\text{C}$  und  $RF = [77, 83]\%$ . Die Konstanten sind in JOECKEL UND STOBER (1999, S.76) bzw. RÜEGER (1996, S.63ff) angegeben, wobei die Impräzision zur Hälfte der letzten Nachkommastelle angesetzt wurde:  $[c_1] = [7.45, 7.55]$ ,  $[c_2] = [237.45, 237.55]$  und  $[c_3] = [0.785705, 0.785715]$ . Der Wertebereich kann bei dieser einfachen Funktion und den obigen Zahlenwerten wieder direkt angegeben werden:

$$\mathcal{W}([p_w]) = \left[ \frac{RF}{100} 10^{\frac{c_1\vartheta}{c_2+\vartheta}+c_3}, \frac{\overline{RF}}{100} 10^{\frac{\overline{c_1}\overline{\vartheta}}{\overline{c_2}+\overline{\vartheta}}+\overline{c_3}} \right]. \quad (3.14)$$

Da die Variable  $\vartheta$  im Funktionsausdruck zweimal vorkommt und bei der Intervallauswertung als unabhängig betrachtet wird, wird der Wertebereich überschätzt. Dies kann leicht gezeigt werden, indem man zwei Variable  $\vartheta_1$  und  $\vartheta_2$  einsetzt. Wir beschränken uns wegen der Monotonie der Exponentialfunktion auf die Betrachtung des Exponenten in (3.13):

$$Exp := \frac{c_1\vartheta_1}{c_2 + \vartheta_2} + c_3. \quad (3.15)$$

Die Intervallauswertung liefert dann:

$$[Exp] = \left[ \frac{\underline{c_1}\underline{\vartheta_1}}{\underline{c_2} + \underline{\vartheta_2}} + \underline{c_3}, \frac{\overline{c_1}\overline{\vartheta_1}}{\overline{c_2} + \overline{\vartheta_2}} + \overline{c_3} \right]. \quad (3.16)$$

Da  $\vartheta_1$  und  $\vartheta_2$  beide unabhängig voneinander variieren, können größere bzw. kleinere Funktionswerte angenommen werden als im Wertebereich. In vorherigen Beispiel 18, Gleichung (3.12) war dies nicht der Fall, da dort für die einzelnen Terme die Minima und Maxima für die zu betrachtenden Intervalle an der gleichen Stelle,

nämlich bei  $\bar{\lambda}$  bzw.  $\underline{\lambda}$  angenommen wurden. Die folgende Tabelle stellt die Zahlenwerte für die untere bzw. obere Intervallgrenze sowie die Intervallmitte und den Intervalldurchmesser für dieses Beispiel gegenüber:

	$p_w$	$\overline{p_w}$	$p_{w,m}$	$w([p_w])$
Sensitivitätsanalyse	17.5741	19.7814	18.6778	2.2073
Intervallauswertung	17.5587	19.8446	18.7017	2.2859
Wertebereich	17.5948	19.8025	18.6987	2.2077

Tabelle 3.9: Vergleich der Approximation des Wertebereichs des Partialwasserdampfdruckes, alle Werte in [hPa]

Die Hausdorff-Distanzen ergeben  $d(\mathcal{W}([p_w]), [p_w]) = 0.0632 \text{ hPa}$  für den Abstand zwischen Wertebereich und Intervallauswertung und  $d(\mathcal{W}([p_w]), p_{w,A}) = 0.0207 \text{ hPa}$  zwischen Wertebereich und der Approximation, die mit der Sensitivitätsanalyse berechnet wurde. Die Variation zwischen dem Wertebereich und seinen Approximationen durch die Intervallauswertung  $[p_w]$  bzw. durch die affine Form  $p_{w,A}$  liegt weit unterhalb der Auflösbarkeit der Messung, so dass die Sensitivitätsanalyse ohne Bedenken angewandt werden kann, um einen repräsentativen Bereich für den Wertebereich anzugeben.

**Beispiel 20** Mit diesen Werten kann jetzt der Brechungsindex  $n_A$  nach der von Kohlrausch vereinfachten Formel (RÜEGGER 1996, S.54ff) oder (JOECKEL UND STÖBER 1999, S.70ff) berechnet werden:

$$n_A = 1 + N_A \cdot 10^{-6} = 1 + \frac{p}{p_0} \frac{n_G - 1}{1 + \alpha \vartheta} - \frac{1}{1 + \alpha \vartheta} \frac{p_w}{p_{w,0}}. \quad (3.17)$$

Da sehr kleine Eingangsintervalle vorliegen und der zweite Term um drei Größenordnungen kleiner ist als der erste Term, kann der Wertebereich wieder direkt aus der Betrachtung von (3.17) angegeben werden. Um keine Unabhängigkeit bezüglich der Temperatur vorzutauschen, muss  $p_w$  nach (3.13) in seine einzelnen Komponenten zerlegt werden. Mit  $[\alpha] = [3.6605, 3.66015] \cdot 10^{-3}$ ,  $[p_0] = [1013.245, 1013.255]$ ,  $[p_{w,0}] = \frac{1}{4.1 \cdot 10^{-8}} \pm 5 \cdot 10^{-3}$  und den meteorologischen Werten von  $[\vartheta] = [19.8, 20, .2]^\circ\text{C}$ ,  $RF = [77, 83]\%$  und  $[p] = [999.5, 1000.5] \text{ hPa}$  ergibt sich dann:

	$N_A$	$\overline{N_A}$	$N_{A,m}$	$w([N_A])$
Sensitivitätsanalyse	270.7530	271.8850	271.3191	1.1321
Intervallauswertung	270.7519	271.8934	271.3227	1.1414
Wertebereich	271.1562	271.8911	271.5236	0.734

Tabelle 3.10: Vergleich der Approximation des Wertebereichs des Brechungsindex, alle Werte in [ppm]

Tabelle 3.10 stellt die Brechzahlen  $N_A = (n_A - 1) \cdot 10^6$  dar. Mit diesen Vorüberlegungen lässt sich der Wertebereich der Strecke nach erster Geschwindigkeitsreduktion (erster Term in (3.5)) folgendermaßen abschätzen:

Intervallauswertung		Sensitivitätsanalyse		Wertebereich	
$\underline{S}$	$\overline{S}$	$\underline{S}$	$\overline{S}$	$\underline{S}$	$\overline{S}$
99.9994	100.0029	99.9996	100.0028	100.0006	100.0016
550.0036	550.0093	550.0044	550.0085	550.0059	550.0069
1000.0077	1000.0157	1000.0092	1000.0143	1000.0112	1000.0121

Tabelle 3.11: Vergleich der Approximation des Wertebereichs durch die Intervallauswertung und die Sensitivitätsanalyse für drei unterschiedlich lange Strecken, alle Werte in [m]

Die Intervalle, die mit der Intervallauswertung und der Sensitivitätsanalyse berechnet werden, schließen beide den Wertebereich ein. Ihr Durchmesser nimmt mit zunehmender Streckenlänge zu, wohingegen der Durchmesser des Wertebereichs annähernd konstant bleibt, da die Additionskonstante  $d_0 = [-0.5, 0.5] \text{ mm}$  den größten Einfluss hat.

## 4. Analyse geodätischer Netze unter Berücksichtigung systematisch wirkender Restunsicherheiten

In den vorangegangenen Kapiteln wurden Intervalle als Spezialfälle von Fuzzy-Sets zur Impräzisionsbeschreibung für geodätische Messungen motiviert und konkrete Werte sowie Größenordnungen exemplarisch für gängige Messungstypen abgeleitet. Eine zentrale Aussage war dabei, dass die Gesamtungewissheit geodätischer Messungen adäquat durch eine Überlagerung von Impräzision und Stochastizität beschrieben werden kann.

In diesem Kapitel wird die Fortpflanzung der Unsicherheit der korrigierten Messungen auf die zu bestimmenden Parameter des geodätischen Modells hergeleitet. In einem ersten Schritt wird das Gauß-Markov-Modell als Standardansatz zur Parameterschätzung in der Geodäsie vorgestellt. Es bildet die Grundlage für die Fortpflanzung der Gesamtunsicherheit und liefert als Ergebnis die geschätzten Parameter und ihre Varianz-Kovarianz-Matrix (stochastische Komponente der Unsicherheit). In einem zweiten Schritt wird die Impräzision durch Intervallauswertung der Schätzvorschrift linear auf die bestimmten Parameter fortgepflanzt. Eine Wertebereichsbetrachtung der Intervallauswertung führt auf *Zonotope* als Impräzisionsmaße für Punktpositionen. In einem weiteren Schritt werden verschiedene Konzepte zur Überlagerung beider Komponenten der Unsicherheit der geschätzten Parameter, Stochastizität und Impräzision, vorgestellt, wobei eine stochastische Interpretierbarkeit der Gesamtunsicherheit der Ergebnisse im Vordergrund steht. Abschließend werden alternative Konzepte zur Intervallfortpflanzung diesem Ansatz gegenübergestellt.

### 4.1 Parameterschätzung im Gauß-Markov-Modell

#### 4.1.1 Konzept

Im geodätischen Modell werden die nichtlinearen Zusammenhänge zwischen den Beobachtungen und den zu bestimmenden Parametern vorgehalten. Die klassische Lösung der Parameterbestimmung in der Geodäsie ist die Schätzung der Parameter nach der Methode der kleinsten Quadrate. Dazu wird der ursprünglich nichtlineare Zusammenhang zwischen Messungen und Parametern durch eine Taylor-Reihenentwicklung an der Stelle der Näherungswerte  $\overset{\circ}{\mathbf{x}}$  für die Parameter linearisiert. Unter Berücksichtigung der stochastischen Eigenschaften der korrigierten Beobachtungen erhält man das *Gauß-Markov-Modell* (KOCH 1997, Kap.3.2). Gleichung (4.1) bezeichnet man als *mathematisches Modell*, (4.2) als *stochastisches Modell*.

$$E(\mathbf{l}) = \mathbf{f}(\overset{\circ}{\mathbf{x}}) + \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\overset{\circ}{\mathbf{x}}} = \mathbf{a}_0 + \mathbf{A} \mathbf{d}\overset{\circ}{\mathbf{x}}, \quad (4.1)$$

$$D(\mathbf{l}) = \mathbf{C}_{ll} = \sigma_0^2 \mathbf{Q}_{ll} = \sigma_0^2 \mathbf{P}_{ll}^{-1}, \quad (4.2)$$

wobei

$E(\cdot), D(\cdot)$	den Erwartungswert- bzw. Dispersionsoperator darstellt,
$\mathbf{l}$	den $n \times 1$ Vektor der Beobachtungen,
$\mathbf{A} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right _{\overset{\circ}{\mathbf{x}}}$	die $n \times u$ Jakobimatrix (Designmatrix) an der Stelle $\overset{\circ}{\mathbf{x}}$ mit $u < n$ ,
$\mathbf{P}_{ll}$	die $n \times n$ Gewichtsmatrix der Beobachtungen,
$\mathbf{a}_0$	den $n \times 1$ Vektor des Gliedes 0. Ordnung der Taylorentwicklung,
$\mathbf{f}$	den funktionalen Zusammenhang zwischen Messungen und zu schätzenden Parametern,
$\mathbf{d}\overset{\circ}{\mathbf{x}}$	den $u \times 1$ Vektor der geschätzten Zuschläge zu den Parametern,
$\overset{\circ}{\mathbf{x}}$	den $u \times 1$ Vektor der Näherungswerte der Parameter und
$\sigma_0^2$	den a priori Varianzfaktor.

Gleichung (4.1) lässt sich in die sogenannten Verbesserungsgleichungen überführen:

$$\mathbf{v} = \mathbf{A} \mathbf{d}\overset{\circ}{\mathbf{x}} - (\mathbf{l} - \mathbf{a}_0). \quad (4.3)$$

Mit der Minimumbedingung für den Vektor der Verbesserungen  $\mathbf{v}$

$$\Omega = \mathbf{v}^T \mathbf{P}_{ll} \mathbf{v} \rightarrow \min \quad (4.4)$$

ergeben sich bei spaltenregulärer Designmatrix die geschätzten Zuschläge zu den Näherungswerten der Parameter:

$$\begin{aligned} \mathbf{d}\hat{\mathbf{x}} &= (\mathbf{A}^T \mathbf{P}_{ll} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P}_{ll} (\mathbf{1} - \mathbf{a}_0) \\ &= \mathbf{N}^{-1} \mathbf{A}^T \mathbf{P}_{ll} (\mathbf{1} - \mathbf{a}_0), \end{aligned} \quad (4.5)$$

wobei  $\mathbf{N}$  die Normalgleichungsmatrix darstellt. Die ausgeglichenen Parameter und Beobachtungen, die Verbesserungen sowie ihre Kofaktormatrizen erhält man schließlich über:

$$\hat{\mathbf{x}} = \overset{\circ}{\mathbf{x}} + \mathbf{d}\hat{\mathbf{x}} \quad , \quad \mathbf{Q}_{\hat{x}\hat{x}} = (\mathbf{A}^T \mathbf{P}_{ll} \mathbf{A})^{-1}, \quad (4.6)$$

$$\hat{\mathbf{l}} = \mathbf{Q}_{\hat{l}\hat{l}} \mathbf{P}_{ll} \mathbf{l} + \mathbf{Q}_{\hat{v}\hat{v}} \mathbf{P}_{ll} \mathbf{a}_0 \quad , \quad \mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{A} \mathbf{Q}_{\hat{x}\hat{x}} \mathbf{A}^T, \quad (4.7)$$

$$\hat{\mathbf{v}} = -\mathbf{Q}_{\hat{v}\hat{v}} \mathbf{P}_{ll} (\mathbf{1} - \mathbf{a}_0) \quad , \quad \mathbf{Q}_{\hat{v}\hat{v}} = \mathbf{Q}_{ll} - \mathbf{Q}_{\hat{l}\hat{l}}. \quad (4.8)$$

Eine Überprüfung des gesamten Modellansatzes ist durch den *Globaltest* möglich, vgl. NIEMEIER (2002, S.147ff). Mit der quadratischen Form  $\Omega$  (4.4) wird der a posteriori Varianzfaktor erwartungstreu geschätzt:

$$\hat{\sigma}_0^2 := \frac{\hat{\mathbf{v}}^T \mathbf{P}_{ll} \hat{\mathbf{v}}}{n - u}. \quad (4.9)$$

Daraus wird mit dem a priori Varianzfaktor die Testgröße  $T$  des Globaltests aufgestellt. Sie ist  $\chi_{n-u}^2$  verteilt für der Nullhypothese der Identität der Varianzfaktoren  $\hat{\sigma}_0^2 = \sigma_0^2$ :

$$T = \frac{\hat{\sigma}_0^2}{\sigma_0^2} \sim \chi_{n-u}^2 \mid H_0 : \hat{\sigma}_0^2 = \sigma_0^2. \quad (4.10)$$

Weicht der a posteriori Varianzfaktor im Rahmen des Globaltests signifikant vom a priori Varianzfaktor ab, so liegen Diskrepanzen zwischen Modell und Daten vor, die durch ein unzureichendes stochastisches oder mathematisches Modell oder durch grobe Fehler begründet sind. Liegen keine signifikanten Abweichungen vor, so kann davon ausgegangen werden, dass Modell und Realität „gut“ übereinstimmen. In diesem Fall können die Varianz-Kovarianz-Matrizen mit dem a priori Varianzfaktor  $\sigma_0^2$  berechnet werden:

$$\mathbf{C}_{\hat{x}\hat{x}} = \sigma_0^2 (\mathbf{A}^T \mathbf{P}_{ll} \mathbf{A})^{-1}, \quad \mathbf{C}_{\hat{l}\hat{l}} = \mathbf{A} \mathbf{C}_{\hat{x}\hat{x}} \mathbf{A}^T \quad \text{und} \quad \mathbf{C}_{\hat{v}\hat{v}} = \mathbf{C}_{ll} - \mathbf{C}_{\hat{l}\hat{l}}. \quad (4.11)$$

### Besondere Aspekte bei der Ausgleichung geodätischer Netze

Bei der Modellierung der Messung werden neben den Koordinaten als eigentlichen Parametern noch Zusatzparameter wie Orientierungsunbekannte bei Richtungen oder Maßstabsfaktoren bei Streckenmessungen eingeführt. Diese können, da sie nicht von direktem Interesse sind, aus den Satz der zu schätzenden Parameter eliminiert werden. FUNCKE (1982) stellt dazu drei Verfahren zusammen. Im Folgenden soll kurz auf die *Methode der reduzierten Verbesserungsgleichungen* eingegangen werden. Hierbei wird der Parametervektor in einen Koordinatenanteil  $\mathbf{d}\hat{\mathbf{x}}_K$  und einen Anteil der Zusatzparameter  $\mathbf{d}\hat{\mathbf{x}}_Z$  unterteilt:

$$\mathbf{v} = (\mathbf{A}_K \quad \mathbf{A}_Z) \begin{pmatrix} \mathbf{d}\hat{\mathbf{x}}_K \\ \mathbf{d}\hat{\mathbf{x}}_Z \end{pmatrix} - (\mathbf{1} - \mathbf{a}_0). \quad (4.12)$$

Durch eine entsprechende Partitionierung der Normalgleichung kann  $\mathbf{d}\hat{\mathbf{x}}_Z$  formal gelöst werden. Wird dieses Ergebnis in (4.12) eingesetzt, so erhält man die Verbesserungsgleichung:

$$\begin{aligned} \mathbf{v} &= \mathbf{R} \mathbf{A}_K \mathbf{d}\hat{\mathbf{x}}_K - \mathbf{R} (\mathbf{1} - \mathbf{a}_0), \quad \text{mit} \quad \mathbf{R} := \mathbf{I} - \mathbf{A}_Z (\mathbf{A}_Z^T \mathbf{P}_{ll} \mathbf{A}_Z)^{-1} \mathbf{A}_Z^T \mathbf{P}_{ll} \\ &= \bar{\mathbf{A}}_K \mathbf{d}\hat{\mathbf{x}}_K - (\bar{\mathbf{1}} - \bar{\mathbf{a}}_0). \end{aligned} \quad (4.13)$$

Werden in den oben aufgeführten Ausdrücken (4.5)-(4.8)  $\mathbf{A}$ ,  $\mathbf{l}$  und  $\mathbf{a}_0$  durch  $\bar{\mathbf{A}}$ ,  $\bar{\mathbf{l}}$  bzw.  $\bar{\mathbf{a}}_0$  ersetzt, so enthält man entsprechende Formeln mit reduzierten Größen. Die Kofaktormatrix der ausgeglichenen Koordinaten und Beobachtungen sowie der Verbesserungen berechnen sich zu:

$$\mathbf{Q}_{\hat{x}\hat{x}} = (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1}, \quad \mathbf{C}_{\hat{l}\hat{l}} = \bar{\mathbf{A}} \mathbf{Q}_{\hat{x}\hat{x}} \bar{\mathbf{A}}^T + \mathbf{A}_Z (\mathbf{A}_Z \mathbf{P}_{ll} \mathbf{A}_Z)^{-1} \mathbf{A}_Z^T \mathbf{P}_{ll} \quad \text{und} \quad \mathbf{Q}_{\hat{v}\hat{v}} = \mathbf{Q}_{ll} - \mathbf{Q}_{\hat{l}\hat{l}}. \quad (4.14)$$

Zur Berechnung der Koordinatenzuschläge  $\mathbf{d}\hat{\mathbf{x}}_K$  reicht es jedoch aus, in (4.5) die Matrix  $\mathbf{A}$  durch  $\bar{\mathbf{A}}$  zu ersetzen, da die folgende Identität gibt:

$$\begin{aligned} \bar{\mathbf{A}}^T \mathbf{P}_{ll} (\bar{\mathbf{1}} - \bar{\mathbf{a}}_0) &= \bar{\mathbf{A}}^T \mathbf{P}_{ll} \mathbf{R} (\mathbf{1} - \mathbf{a}_0) \\ &= \bar{\mathbf{A}}^T \mathbf{P}_{ll} (\mathbf{1} - \mathbf{a}_0) - \mathbf{A}_K^T \mathbf{R}^T \mathbf{P}_{ll} \mathbf{A}_Z (\mathbf{A}_Z \mathbf{P}_{ll} \mathbf{A}_Z)^{-1} \mathbf{A}_Z^T \mathbf{P}_{ll} (\mathbf{1} - \mathbf{a}_0) \\ &= \bar{\mathbf{A}}^T \mathbf{P}_{ll} (\mathbf{1} - \mathbf{a}_0). \end{aligned} \quad (4.15)$$

Ein *geodätisches Netz* ist durch seine Geometrie, Topologie und durch sein Datum definiert. Ziel ist die Bestimmung zwei- oder dreidimensionaler kartesischer Punktkoordinaten ( $y_i, x_i$  im Zweidimensionalen bzw.  $x_i, y_i, z_i$  im Dreidimensionalen) aus den geodätischen Messungen wie Strecken, Richtungen, Zenitdistanzen oder GPS-Basislinien. Die *Netzgeometrie* wird durch die konkreten Werte der Messungen bestimmt. Die *Netztopologie* beschreibt, welche Punkte durch Messungen miteinander verbunden sind. Das *Netzdatum* muss festgelegt werden, um eindeutige Koordinaten bestimmen zu können, da die Messungen nicht oder bei GPS nur teilweise sensitiv bezüglich des Koordinatenursprung, der Orientierung des Koordinatensystems oder des Maßstabes sind. Wird nicht a priori über das Netzdatum verfügt, so weist die Designmatrix einen Spaltenrangdefekt  $d$  (*Datumsdefekt*) auf und die resultierende Normalgleichungsmatrix ist singulär und kann nicht invertiert werden. Es soll hier vorausgesetzt werden, dass alle Koordinaten durch die Netzkonfiguration bestimmbar sind und kein *Konfigurationsdefekt* auftritt (NIEMEIER 2002, S.214).

Singuläre Varianz-Kovarianz-Matrizen der Beobachtungen können beispielsweise bei der Weiterverarbeitung von GPS-Ergebnissen oder der Deformationsanalyse auftreten. Hier bilden ggf. frei ausgeglichene Punktkoordinaten die Eingangsgrößen, so dass die Gewichtsmatrix nicht nach (4.2) berechnet werden kann. Als Folge tritt auch bei spaltenregulärer Designmatrix eine singuläre Normalgleichungsmatrix auf. In dieser Arbeit soll auf diesen Aspekt nicht weiter eingegangen und vorausgesetzt werden, dass die Varianz-Kovarianz-Matrix der Beobachtungen regulär ist und sich die Gewichtsmatrix nach (4.2) berechnen lässt.

Um bei singulärer Normalgleichungsmatrix dennoch Koordinaten bestimmen zu können, haben sich in der Geodäsie verschiedene Methoden durchgesetzt, vgl. GRAFAREND UND SCHAFFRIN (1993), CASPARY UND WICHMANN (1994), KOCH (1997) oder NIEMEIER (2002). Ziel ist es, durch die Festlegung des Netzdatums die Netzgeometrie (*innere Geometrie*) nicht zu verzerren. Es dürfen somit nur  $d$  unabhängige Restriktionen an das Ausgangssystem (4.1) gestellt werden. Werden mehr als  $d$  Restriktionen eingeführt oder mehr als  $d$  stochastische Bedingungen gestellt, so wird ein Zwang auf das Netz ausgeübt und die innere Geometrie verzerrt.

Ein einfaches und pragmatisches Vorgehen ist, über  $d$  Koordinaten zu verfügen, d.h. sie auf feste Werte zu fixieren, vgl. NIEMEIER (2002, S.230f). Eine allgemeine Lösung kann mittels symmetrischer reflexiver generalisierter Inversen berechnet werden, vgl. KOCH (1997, Kap.33). In der Menge dieser Matrizen zeichnet sich die Pseudoinverse dadurch aus, dass sie eindeutig ist und minimale Spur hat (KOCH 1997, S.67f). Darüber hinaus ist die Norm des Vektors der Koordinatenzuschläge, der mit der Pseudoinversen geschätzt wird, minimal. Die Lösungsmethode wird daher auch als *freie Netzausgleichung mit Gesamtspurminimierung* oder *Gesamtnormminimierung* bezeichnet. Äquivalent kann diese Lösung durch Hinzufügen einer Restriktion zum Gleichungssystem (4.13) erhalten werden:

$$\mathbf{G}^T \mathbf{d}\hat{\mathbf{x}} = \mathbf{0}. \quad (4.16)$$

Die Spaltenvektoren der  $u \times d$  Matrix  $\mathbf{G}$  bilden eine Basis des Nullraums der Matrix  $\bar{\mathbf{A}}$ . Die Matrix  $\mathbf{G}$  lässt sich auch als Designmatrix einer Ähnlichkeitstransformation darstellen, vgl. ILLNER (1983) oder NIEMEIER (2002, S.234f), wobei gilt:

$$\bar{\mathbf{A}}\mathbf{G} = \mathbf{0}, \quad \mathbf{N}\mathbf{G} = \mathbf{0}. \quad (4.17)$$

Sollen nicht alle Netzpunkte auch Datumspunkte sein, so kann eine Restriktionsmatrix in der Form:  $\mathbf{D} = \mathbf{\Gamma}\mathbf{G}$  verwendet werden. Mit der  $u \times u$  Selektionsmatrix  $\mathbf{\Gamma}$  werden die entsprechenden Koordinaten der Datumspunkte ausgewählt. Die so bestimmte Varianz-Kovarianz-Matrix der geschätzten Koordinaten ist eine symmetrische reflexive generalisierte Inverse, deren Teilspur (für die Datumspunkte) minimal wird. Entsprechend ist die Teilnorm des geschätzten Koordinatenvektor für die Datumspunkte minimal, vgl. KOCH (1997, S.69f). Durch eine *S-Transformation* können die geschätzten Koordinatenvektoren und entsprechenden Varianz-Kovarianz-Matrizen ohne erneute Ausgleichung in ein anderes (inneres) Datum transformiert werden, vgl. ILLNER (1983) und die Ausführungen in Abschnitt 4.1.3.

#### 4.1.2 Mathematisches Modell der dreidimensionalen geometrischen Ausgleichung

Der klassische Ansatz zur Weiterverarbeitung von Messungen mit dem Ziel der Bestimmung von Punktkoordinaten ist eine getrennte Korrektur der Messungen einzelner Messungstypen und Aufspaltung der Koordinaten in eine Lage- und Höhenkomponente. Im Bereich der klassischen Landesvermessung war diese Trennung sinnvoll, da einerseits die einzelnen Messungstypen zeitlich getrennt mit verschiedenen Instrumenten zu verschiedenen Zielmarken gemessen wurden. Andererseits wurden hauptsächlich großräumige Netze (Triangulationsnetz 1. Ordnung) gemessen. Bei der Auswertung solcher annähernd ebener Netze zerfällt das Normalgleichungssystem in einen Lage- und einen Höhenanteil. Daher führt eine gemeinsame Auswertung auf ein schlecht konditioniertes

Gleichungssystem für die Höhe. Aus mess- wie rechentechnischen Gesichtspunkten ist somit diese Aufspaltung gerechtfertigt.

Im Bereich der Ingenieurvermessung und bei Verwendung moderner Instrumente ist die Sachlage anders. Die Netze sind eher kleinräumig, so dass eine Messung zur Bestimmung aller drei Koordinatenrichtungen beitragen kann. Eine Verwendung von GPS-Messungen liefert zusätzlich übergreifende dreidimensionale Information. Werden moderne Instrumente eingesetzt, so werden in der Regel alle drei Messungstypen (Richtungen, Strecken und Zenitdistanzen) vom selben Instrument (Totalstation) ausgeführt und erste Korrekturen mit der spezifischen Sensorsoftware angebracht. Die Messungen können weitgehend zeitgleich durchgeführt werden, so dass sie von gleichen äußeren (atmosphärischen) Bedingungen beeinflusst werden. Es ist daher sinnvoll, eine gemeinsame Korrektur der Messungen durchzuführen und direkt dreidimensionale Koordinaten zu berechnen. Daraus lassen sich jeder Zeit Lage- und Höhenkomponenten bezüglich beliebiger Bezugsflächen berechnen. Zusätzlich wird diese Transformation konsequent für das gesamte Netz im gleichen Modell durchgeführt und nicht mit unterschiedlichen Ansätzen für einzelne Messungstypen. Aus diesen Gründen wird im Folgenden eine gemeinsame Modellierung aller Messungstypen in einem einheitlichen Koordinatensystem und eine gemeinsame Reduktion bzw. Korrektur wegen äußerer Einflüsse vorgeschlagen.

In einem ersten Schritt werden unter Berücksichtigung von Exzentren und Windschiefe der Lotlinien zwischen Stand- und Zielpunkten die Koordinaten der Bodenpunkte oder Referenzmarken berechnet. Hierbei fließt die Impräzision der Zentrierung bzw. der Bestimmung der Exzentrizitäten mit ein. Als Ergebnis erhält man die Koordinaten des interessierenden Objektes. In einem zweiten Schritt werden daraus die eigentlichen Zielgrößen abgeleitet. Im Fall der Deformationsanalyse können diese beispielsweise dreidimensionale Verschiebungsvektoren oder Straintensoren sein. Für ebene Darstellungen von Objekten werden zweidimensionale Koordinaten auf einer entsprechenden Bezugsfläche die Zielgrößen sein, die unter Anwendung von Abbildungsvorschriften der Kartenprojektion erhalten werden. Im Bereich der Industrievermessung oder Qualitätsprüfung muss von den Koordinaten auf Formparameter geschlossen werden, mit denen beispielsweise ein Soll-Ist-Vergleich des hergestellten Objektes mit einem CAD-Plan ermöglicht wird (DRIXLER 1993). Zusammenfassend sind diese Zielgrößen beliebige Funktionen der Koordinaten.

Das funktionale Modell für Messungen zwischen zwei Punkten  $P$  und  $Q$  kann in beliebigen Koordinatensystemen dargestellt werden (HECK 1995, S.35ff). Es ist jedoch naheliegend, die Messungen als Polarkoordinaten bezüglich des Standpunktes aufzufassen. Durch die topozentrischen Koordinatendifferenzen  $\Delta \mathbf{x}_I^* = (\Delta x_I^*, \Delta y_I^*, \Delta z_I^*)^T$  zwischen den Instrumentenreferenzpunkten (z.B. den Schnittpunkten der Instrumentenachsen) lassen sich Strecke, Richtung und Zenitdistanz folgendermaßen ausdrücken:

$$S = \sqrt{(\Delta x_I^*)^2 + (\Delta y_I^*)^2 + (\Delta z_I^*)^2}, \quad (4.18)$$

$$r = \arctan \frac{\Delta y_I^*}{\Delta x_I^*} + o_i, \quad (4.19)$$

$$\zeta = \arctan \frac{\sqrt{(\Delta x_I^*)^2 + (\Delta y_I^*)^2}}{\Delta z_I^*}, \quad (4.20)$$

wobei  $o_i$  die Orientierung des Teilkreises bezeichnet. Um mehrere Standpunkte miteinander zu verknüpfen, muss auf ein übergeordnetes Koordinatensystem übergegangen werden. Hierzu eignet sich ein kartesisches, global geozentrisches System (HECK 1995, S.37ff). Der Zusammenhang zwischen Instrumentenreferenzpunkt und Bodenpunkt in diesem System kann folgendermaßen hergestellt werden:

$$\mathbf{x}_{P_0}^\bullet = \mathbf{x}_{P_I}^\bullet - \mathbf{R}^T(\Lambda_P, \Phi_P) \mathbf{e}_{P_I}^*, \quad \mathbf{x}_{Q_0}^\bullet = \mathbf{x}_{Q_I}^\bullet - \mathbf{R}^T(\Lambda_Q, \Phi_Q) \mathbf{e}_{Q_I}^*, \quad (4.21)$$

wobei  $\mathbf{x}_{P_0}^\bullet, \mathbf{x}_{Q_0}^\bullet$  die geozentrischen Ortsvektoren zu den Bodenpunkten  $P_0$  bzw.  $Q_0$  und  $\mathbf{x}_{P_I}^\bullet, \mathbf{x}_{Q_I}^\bullet$  die Ortsvektoren zu den Instrumentenreferenzpunkten  $P_I$  bzw.  $Q_I$  sind. Die Vektoren der Exzentrizitäten im Stand- und Zielpunkt,  $\mathbf{e}_{P_I}^*, \mathbf{e}_{Q_I}^*$  sind im jeweiligen topozentrischen System gegeben. Die Verknüpfung vom global geozentrischen System ins kartesische topozentrische System wird durch eine Rotation realisiert. Die Rotationsmatrix  $\mathbf{R}(\Lambda, \Phi)$  hängt dabei von den astronomischen Koordinaten  $\Lambda, \Phi$  ab, vgl. z.B. HECK (1995, S.39).

Im klassischen Modell tauchen Exzentren nur in Form von Instrumenten- und Ziel- bzw. Reflektorhöhe beispielsweise beim trigonometrischen Nivellement auf, vgl. KAHMEN (1993, Kap.12). Hier soll der Begriff der Exzentrizität weiter gefasst werden, um einerseits Aufbauten in mehreren Höhen (verschiedene Instrumente oder Zielmarken) auf einen gemeinsamen Bodenpunkt beziehen zu können. Beispielsweise sind bei Kunstbauwerken mit automatischer Überwachung oftmals Instrumentenstandpunkte auch Zielpunkte, so dass die Reflektoren exzentrisch angebracht werden müssen oder Zielpunkte sind von mehreren Standpunkten aus anzuzielen (FLACH UND NATEROP 1999). Bei anderen Anwendungen sollen sich die Messergebnisse verschiedener Sensoren

(GPS-Antenne und Totalstation) auf einen Punkt beziehen, vgl. LUTES (2001). Bei globalen Positionierungsverfahren des IERS sind solche Exzentrizitäten unvermeidbar, da es nicht möglich ist, Sensoren verschiedener Satellitenverfahren wie GPS-Antenne und Laser physikalisch auf einem Punkt zu vereinen. Die dabei verwendeten Exzentren nehmen Ausdehnung von kleinen Verknüpfungsnetzen zwischen den einzelnen Instrumenten auf einer Station an wie beispielsweise auf der Fundamentalstation Wettzell.

Andererseits erlaubt eine solche Modellierung erst die umfassende Beschreibung von Impräzision, die durch die Zentrierung und Sensorhöhenbestimmung in das Messergebnis mit einfließen. Ebenso lässt sich der Einfluss verschiedener Schwerefeldmodelle berücksichtigen bzw. die Vernachlässigung der Konvergenz der Lotlinien zwischen zwei Punkten modellieren.

Für den Zusammenhang zwischen den Koordinatendifferenzen der Bodenpunkte  $\Delta \mathbf{x}^\bullet$  im global geozentrischen System und den Koordinatendifferenzen  $\Delta \mathbf{x}_I^*$  der Instrumentenreferenzpunkte im topozentrischen System gilt:

$$\Delta \mathbf{x}_I^* = \mathbf{R}(\Lambda_P, \Phi_P) \Delta \mathbf{x}_0^\bullet - \mathbf{e}_P^* + \mathbf{R}(\Lambda_P, \Phi_P) \mathbf{R}^T(\Lambda_Q, \Phi_Q) \mathbf{e}_Q^*. \quad (4.22)$$

Die linearen Zusammenhänge zwischen Messungen  $\mathbf{l} = \{S, R, \zeta\}$  und Parametern  $\Pi = \{x_i, y_i, z_i, \Phi_i, \Lambda_i, e_{N,i}, e_{E,i}, e_{H,i}\}$  werden durch Differentiation der Ausdrücke (4.18)-(4.20) unter Berücksichtigung von (4.22) erhalten:

$$\frac{\partial \mathbf{l}}{\partial \Pi} = \frac{\partial \Delta \mathbf{x}_I^*}{\partial \Pi} \frac{\partial \mathbf{l}}{\Delta \mathbf{x}_I^*}. \quad (4.23)$$

Der Satz der zu schätzenden Parameter kann in vier Gruppen unterteilt werden: Punktkoordinaten, astronomische Koordinaten, Exzentrizitäten und Zusatzparameter, vgl. (4.24). Das resultierende Gauß-Markov-Modell hat die folgende Form:

$$\mathbf{l} + \mathbf{v} = \mathbf{A}_1 \mathbf{d}\mathbf{x}_1 + \mathbf{A}_2 \mathbf{d}\mathbf{x}_2 + \mathbf{A}_3 \mathbf{d}\mathbf{x}_3 + \mathbf{A}_4 \mathbf{d}\mathbf{x}_4 + \mathbf{a}_0. \quad (4.24)$$

Je nach Aufgabenstellung oder Vorinformation kann der Satz der Parameter variieren: Liegt stochastische Vorinformation für die Parameter vor, so können diese als Beobachtungen aufgefasst werden. Sind feste Werte für Parameter bekannt und sollen diese deterministisch behandelt werden, so wird der entsprechende Anteil dem Absolutgliedvektor  $\mathbf{a}_0$  zugeschlagen. In diesem Fall kann ein Intervallvektor  $[\mathbf{a}_0]$  angegeben werden, der die Impräzision des entsprechenden Anteils ausdrückt. Zusätzlich können noch weitere Bedingungen oder Restriktionen zwischen Parametern eingeführt werden. Das Gauß-Markov-Modell in der Form (4.24) weist somit eine große Flexibilität auf, um den einzelnen Aufgabenstellungen geeignet angepasst zu werden. Für alle folgenden Untersuchungen in dieser Arbeit wird, wenn nicht anders erwähnt, davon ausgegangen, dass das geodätische Datum durch Festpunkte bestimmt sei, die astronomischen Koordinaten und Exzentren deterministisch bekannt seien und die Zusatzparameter (wie Maßstab und Orientierung) nach bekannten Verfahren eliminiert wurden (FUNCKE 1982), so dass mit der Schreibweise aus (4.13) das resultierende Gauß-Markov-Modell folgende Form hat, wobei der Vektor der zu schätzenden Koordinaten mit  $\mathbf{d}\mathbf{x}$  bezeichnet wird:

$$\bar{\mathbf{l}} + \mathbf{v} = \bar{\mathbf{A}} \mathbf{d}\mathbf{x} + \bar{\mathbf{a}}_0. \quad (4.25)$$

### 4.1.3 Analyse geodätischer Netze nach Genauigkeit und Zuverlässigkeit

Im Folgenden werden statistische Kenngrößen zusammengestellt, die die Genauigkeit und Zuverlässigkeit der ausgeglichenen Koordinaten und des gesamten Netzes charakterisieren. Die Varianz-Kovarianz-Matrix der ausgeglichenen Koordinaten ist Träger der Genauigkeitsinformation des Netzes. Wichtige Kriterien für die Beurteilung der Genauigkeit der zwei- bzw. dreidimensionalen Punktpositionen sind *lokale Maße* wie Punktconfidenzellipsen bzw. -ellipsoide. Bezüglich einer Konfidenzwahrscheinlichkeit  $1 - \gamma$  sind diese Bereiche bei bekanntem Varianzfaktor durch die quadratischen Formen beschrieben:

$$\mathcal{K}_{1-\gamma} : (\hat{\mathbf{x}}_P - \mathbf{x}_P)^T \mathbf{C}_{\hat{\mathbf{x}}_P}^{-1} (\hat{\mathbf{x}}_P - \mathbf{x}_P) \leq \chi_{2,1-\gamma}^2 \quad \text{bzw.} \quad (\hat{\mathbf{x}}_P - \mathbf{x}_P)^T \mathbf{C}_{\hat{\mathbf{x}}_P}^{-1} (\hat{\mathbf{x}}_P - \mathbf{x}_P) \leq \chi_{3,1-\gamma}^2, \quad (4.26)$$

wobei  $\mathbf{x}_P$  die zwei- bzw. dreidimensionalen Koordinaten eines Punktes und  $\mathbf{C}_{\hat{\mathbf{x}}_P}$  die entsprechende  $2 \times 2$  bzw.  $3 \times 3$  Submatrix von  $\mathbf{C}_{\hat{\mathbf{x}}}$  darstellt. Die Orientierung der Konfidenzellipse wird durch die Eigenvektoren  $\mathbf{m}_i$  von  $\mathbf{C}_{\hat{\mathbf{x}}_P}$ , die Halbachslängen durch die Eigenwerte  $\lambda_i$  bestimmt:

$$\mathbf{C}_{\hat{\mathbf{x}}_P} = \mathbf{M} \mathbf{M}^T \quad (4.27)$$



mit der orthonormalen Modalmatrix  $\mathbf{M}$  in der Dimension der Matrix  $\mathbf{C}_{\hat{x}\hat{x}P}$  und der Diagonalmatrix  $\mathbf{\Lambda}$  der Eigenwerte  $\lambda_i$ . Relative Punktfehlermaße, die die Nachbarschaftsbeziehung zwischen zwei Punkten  $P, Q$  kennzeichnen, werden mit den Submatrizen  $\mathbf{C}_{\hat{x}\hat{x}PQ}$  bzw.  $\mathbf{C}_{\hat{x}\hat{x}QP}$  der Nebendiagonalen bestimmt (NIEMEIER 2002, S.261f)

$$\mathbf{C}_{\Delta\hat{x}\Delta\hat{x}} = \mathbf{C}_{\hat{x}\hat{x}P} + \mathbf{C}_{\hat{x}\hat{x}Q} - \mathbf{C}_{\hat{x}\hat{x}PQ} - \mathbf{C}_{\hat{x}\hat{x}QP}. \quad (4.28)$$

Neben diesen lokalen Maßen, die sich auf einen Punkt oder die Beziehung eines Punktes zu einem anderen beziehen, stehen *globale Maße* zur Beschreibung der gesamten Situation des Netzes zur Verfügung. Diese werden als mittlere oder durchschnittliche Werte angegeben. Durchschnittliche Punktfehler lassen sich aus der regulären Gesamtmatrix  $\mathbf{C}_{\hat{x}\hat{x}}$  ableiten:

$$\sigma_H^2 = \frac{\text{spur}\mathbf{C}_{\hat{x}\hat{x}}}{u} = \frac{1}{u} \sum_{i=1}^u \lambda_i, \quad (4.29)$$

$$\sigma_W^2 = \frac{\det\mathbf{C}_{\hat{x}\hat{x}}}{u} = \frac{1}{u} \prod_{i=1}^u \lambda_i, \quad (4.30)$$

wobei  $\sigma_H^2$ ,  $\sigma_W^2$  den mittleren Punktfehler nach Helmert bzw. Werkmeister beschreiben. Für singuläre  $\mathbf{C}_{\hat{x}\hat{x}}$  siehe beispielsweise JÄGER UND BILL (1985, S.78).

Prinzipiell enthält die Varianz-Kovarianz-Matrix der geschätzten Koordinaten Informationen über die Bestimmungsgüte der Netzgeometrie und des Netzdatums. Durch eine *S-Transformation* (ILLNER 1983) lassen sich der Geometrie- und Datumsanteil aus der Varianz-Kovarianz-Matrix bestimmen. Die allgemeine Form der  $u \times u$  S-Matrix lautet:

$$\begin{aligned} \mathbf{S} &= \mathbf{I} - \mathbf{G}(\mathbf{G}^T\mathbf{W}\mathbf{G})^{-1}\mathbf{G}^T\mathbf{W} \\ &= \mathbf{I} - \mathbf{R}. \end{aligned} \quad (4.31)$$

In der klassischen Vorgehensweise wird für die positiv definite Gewichtsmatrix  $\mathbf{W}$  die Einheitsmatrix gewählt (ILLNER 1983). Dadurch hat die Matrix  $\mathbf{C}_{\hat{x}\hat{x},G}$  des Geometrieanteils minimale Spur:

$$\begin{aligned} \mathbf{C}_{\hat{x}\hat{x}} &= \mathbf{S}\mathbf{C}_{\hat{x}\hat{x}}\mathbf{S}^T - \mathbf{R}\mathbf{C}_{\hat{x}\hat{x}}\mathbf{R}^T + 2\mathbf{R}\mathbf{C}_{\hat{x}\hat{x}} \\ &= \mathbf{C}_{\hat{x}\hat{x},G} - \mathbf{C}_{\hat{x}\hat{x},D} + 2\mathbf{C}_{\hat{x}\hat{x},D,G}. \end{aligned} \quad (4.32)$$

Nachteil dieser Vorgehensweise ist jedoch, dass bei der Wahl der Einheitsmatrix die Varianz-Kovarianz-Matrix nicht in zwei komplementäre Anteile zerlegt werden kann. Dies kann erreicht werden, wenn die Metrik der (regulären) Varianz-Kovarianz-Matrix selbst zugrunde gelegt wird, vgl. SILLARD UND BOUCHER (2001):

$$\begin{aligned} \mathbf{C}_{\hat{x}\hat{x}} &= (\mathbf{C}_{\hat{x}\hat{x}} - \mathbf{G}(\mathbf{G}^T\mathbf{C}_{\hat{x}\hat{x}}^{-1}\mathbf{G})^{-1}\mathbf{G}^T) + \mathbf{G}(\mathbf{G}^T\mathbf{C}_{\hat{x}\hat{x}}^{-1}\mathbf{G})^{-1}\mathbf{G}^T \\ &= \mathbf{C}_{\hat{x}\hat{x},G}^* + \mathbf{C}_{\hat{x}\hat{x},D}^*. \end{aligned} \quad (4.33)$$

Ein weiteres Qualitätskriterium eines Netzes, das auf der Eigenwertzerlegung der Matrix  $\mathbf{C}_{\hat{x}\hat{x}}$  aufbaut, sind die sogenannten *Hauptschwachformen geodätischer Netze*. Erste Entwicklungen wurden von DUPRAZ UND NIEMEIER (1979) vorgestellt. JÄGER (1988) verfeinerte diese Arbeiten und wies auf Analogien zu Schwachformen bei Tragwerken hin. Die Hauptschwachform beschreibt die Linearkombination des Koordinatenvektors, die die größte Unsicherheit aufweist. Diese Form lässt sich durch das Produkt aus der Wurzel des größten Eigenwerts und dem dazugehörigen Eigenvektor der Varianz-Kovarianz-Matrix ausdrücken. Die Vektorkomponenten werden als Verschiebungsvektoren in den einzelnen Punktpositionen aufgefasst, vgl. JÄGER UND KALTENBACH (1990), HECK ET AL. (1996) oder KUTTERER (1998). Dies kann durch die Betrachtung der spektralen Synthese der Varianz-Kovarianz-Matrix motiviert werden:

$$\mathbf{C}_{\hat{x}\hat{x}} = \sum_{i=1}^u \lambda_i \mathbf{m}_i \mathbf{m}_i^T. \quad (4.34)$$

Sind die Eigenwerte der Größe nach absteigend geordnet und dominieren die ersten Eigenwerte das Spektrum, so tragen die diesen Eigenwerten zugeordneten Matrizen  $\mathbf{m}_i \mathbf{m}_i^T$  den größten Informationsanteil der Varianz-Kovarianz-Matrix, d.h. die größte Unsicherheit, da sie mit den entsprechend großen Eigenwerten gewichtet werden. Die Genauigkeit der bestimmten Punktlage und damit auch die Matrix  $\mathbf{C}_{\hat{x}\hat{x}}$  setzt sich wie oben angesprochen aus den zwei Anteilen zusammen: der Bestimmbarkeit der Punktposition durch die Netzgeometrie und dem Datumsanteil. Die Analyse als Hauptschwachform gilt also nur, wenn das Netzdatum gut bestimmt wurde, da sonst die spektralen Kenngrößen diesen Datumsanteil aufdecken und nicht die eigentliche Hauptschwachform.

Umgekehrt kann die Spektralzerlegung gezielt genutzt werden, um festzustellen, welche Datumsanteile schlecht bestimmt sind, vgl. HECK ET AL. (1996), KUTTERER (1998) oder SCHÖN ET AL. (2001).

Neben der Genauigkeit spielt die *Zuverlässigkeit* der ausgeglichenen Ergebnisse eine wichtige Rolle. Zugeordnete Maße sind primär Planungsmaße auf Grundlage der a priori Varianzen. Allerdings können auch bestehende Netze auf ihre Zuverlässigkeit hin untersucht werden. Träger der Information über die Zuverlässigkeit ist die Varianz-Kovarianz-Matrix der Verbesserungen  $\mathbf{C}_{\hat{v}\hat{v}}$ . Ziel ist es, grobe Fehler in den Beobachtungen aufzudecken und die Verzerrung des Netzes durch nicht aufdeckbare Fehler abzuschätzen. Grundidee ist somit, die Kontrollierbarkeit der Messungen zu verbessern. Für unkorrelierte Messung wird diese durch die Redundanzanteile  $r_i$  beschrieben, die die Diagonalelemente der idempotenten Matrix  $\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_{ll}$  sind:

$$r_i = (\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_{ll})_{i,i}. \quad (4.35)$$

Für korrelierte Beobachtungen, die hier nicht weiter betrachtet werden, vergleiche beispielsweise YONG-QI UND WANG (1996) und bei singulären Ausgleichungsproblemen WANG UND CHEN (1999). Die Summe der Redundanzanteile ergibt die Gesamtredundanz  $r = n - u + d$ . Das Komplement der Redundanzanteile sind die Koordinatenbeiträge  $u_{x,i}$  bzw. der Zusatzparameteranteil  $u_{z,i}$  der Beobachtungen, wobei gilt:

$$1 = r_i + u_{x,i} + u_{z,i}. \quad (4.36)$$

Die *innere Zuverlässigkeit* gibt den Grenzwert an eines mit der Wahrscheinlichkeit  $\beta_0$  beim Hypothesentest mit  $\gamma_0$  aufdeckbaren groben Fehlers. Im eindimensionalen Fall gilt bei unkorrelierten Beobachtungen für den Grenzwert  $\nabla_0 l_i$  der  $i$ -ten Beobachtung:

$$\nabla_0 l_i := \delta_{0,i}(\gamma_0, \beta_0) \frac{\sigma_{l_i}}{\sqrt{r_i}}, \quad (4.37)$$

wobei  $\delta_{0,i}(\gamma_0, \beta_0)$  den Nichtzentralitätsparameter der nichtzentralen Normalverteilung bezeichnet, der eine Beobachtung beim Vorliegen eines groben Fehlers folgt, vgl. BAARDA (1968), HECK (1981) oder JÄGER UND BILL (1985). Im mehrdimensionalen Fall (GPS-Basislinie) wird ein Grenzwertellipsoid angegeben:

$$\frac{1}{\sigma_0^2} (\nabla_0 \mathbf{l})^T \mathbf{\Gamma}^T \mathbf{P}_{ll} \mathbf{Q}_{\hat{v}\hat{v}} \mathbf{P}_{ll} \mathbf{\Gamma} (\nabla_0 \mathbf{l}) = \lambda_0(\gamma_0, \beta_0), \quad (4.38)$$

wobei  $m = 2$  oder  $m = 3$  gilt. Der Nichtzentralitätsparameter der  $\chi_m^2$ -Verteilung wird mit  $\lambda_0(\gamma_0, \beta_0)$  bezeichnet, vgl. BAARDA (1968), HECK (1981) oder JÄGER UND BILL (1985). Je größer der Redundanzanteil einer Beobachtung, desto kleiner ist bei ihr der grobe Fehler, der mit einer Wahrscheinlichkeit von  $1 - \beta_0$  gerade nicht mehr aufdeckbar ist. Gleichzeitig nimmt mit zunehmender Größe des Redundanzanteils der Beitrag der Beobachtung zur Bestimmung der Koordinaten ab. JÄGER UND BILL (1985, S.87) geben an, welche Werte die Redundanzanteile für terrestrische Beobachtungen in einem zuverlässigen zweidimensionalen Netz annehmen sollten.

Unter der *äußeren Zuverlässigkeit* versteht man die Auswirkungen eines nicht aufgedeckten groben Fehlers der  $i$ -ten Beobachtung auf die geschätzten Parameter.

$$\nabla_{\mathbf{0},i} \hat{\mathbf{x}} = \frac{\delta_{0,i}(\gamma_0, \beta_0)}{\sqrt{r_i}} \mathbf{C}_{\hat{x}\hat{x}} \bar{\mathbf{a}}_i^T = \mathbf{Q}_{\hat{x}\hat{x}} \bar{\mathbf{a}}_i^T p_i \nabla_0 l_i. \quad (4.39)$$

Als globales Maß bietet sich die Norm des Vektors  $\nabla_{\mathbf{0},i} \hat{\mathbf{x}}$  an (*Netzverzerrung*) (JÄGER UND BILL 1985, S.89):

$$\|\nabla_{\mathbf{0}} \hat{\mathbf{x}}\|_{\mathbf{C}_{\hat{x}\hat{x}}^{-1}} = \frac{\nabla_0 l_i}{\sigma_{l_i}} \sqrt{\bar{\mathbf{a}}_i \mathbf{Q}_{\hat{x}\hat{x}} \bar{\mathbf{a}}_i^T p_i} = \frac{\nabla_0 l_i}{\sigma_{l_i}} \sqrt{u_{x,i}} = \delta_{0,i} \sqrt{\frac{u_{x,i}}{r_i}}. \quad (4.40)$$

Soll das gleichzeitige Auftreten mehrerer grober Fehler betrachtet werden, so gilt:

$$\|\nabla_{\mathbf{0}} \hat{\mathbf{x}}\|_{\mathbf{C}_{\hat{x}\hat{x}}^{-1}} = \frac{1}{\sigma_0^2} \sqrt{\nabla_0 \mathbf{l}^T \mathbf{\Gamma}^T \mathbf{P}_{ll} \bar{\mathbf{A}} \mathbf{Q}_{\hat{x}\hat{x}} \bar{\mathbf{A}}^T \mathbf{P}_{ll} \mathbf{\Gamma} \nabla_0 \mathbf{l}}, \quad (4.41)$$

wobei  $\mathbf{\Gamma}$  eine Selektionsmatrix darstellt, die die entsprechenden groben Fehler auswählt, die betrachtet werden sollen. Ein Netz ist gut angelegt, wenn die Netzverzerrung und der Grenzwert für einen aufdeckbaren groben Fehler gering sind.

## 4.2 Fortpflanzung der Impräzision der korrigierten Messungen auf die geschätzten Parameter

Im vorherigen Abschnitt wurde im Rahmen des klassischen Konzeptes die Fortpflanzung der stochastischen Anteile der Gesamtunsicherheit der korrigierten Beobachtungen auf die geschätzten Parameter aufgezeigt. Der folgende Abschnitt beschreibt die Fortpflanzung der Impräzision. Grundsätzlich bestehen dazu zwei Ansatzmöglichkeiten. Zum Einen kann das bestehende Schätzverfahren, die Methode der kleinsten Quadrate, intervallausgewertet werden. Dieser Ansatz wurde von KUTTERER (1994) in seiner Dissertation erarbeitet und soll hier zur Anwendung kommen. Zum Anderen können für die Herleitung des Algorithmus zur Fortpflanzung der Impräzision bestimmte Zielfunktionen gewählt werden. Zwei dieser Ansätze werden in Abschnitt 4.5 exemplarisch vorgestellt.

### 4.2.1 Intervallauswertung des Parametervektors $\hat{\mathbf{x}}$

Eine sehr naheliegende Lösung zur Übertragung der Impräzision der korrigierten Messung auf die ausgeglichenen Parameter ist die Intervallauswertung der Gleichung (4.5). Hierbei werden a priori keine speziellen Kriterien an die Intervallfortpflanzung gestellt, sondern man bewegt sich weiter im Konzept der Ausgleichung nach der Methode der kleinsten Quadrate. Fasst man das lineare Modell der Ausgleichung (4.5) als vektorwertige Funktion auf, so ist der Begriff der Intervallauswertung und des Wertebereichs definiert:

$$\mathbf{d}\hat{\mathbf{x}} = f(\mathbf{l}, \hat{\mathbf{x}}) = (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} (\mathbf{l} - \mathbf{a}_0). \quad (4.42)$$

Die Intervallauswertung des Schätzwertes (4.5) bzw. (4.42) liefert einen Intervallvektor für die geschätzten Zuschläge  $\mathbf{d}\hat{\mathbf{x}}$ :

$$[\mathbf{d}\hat{\mathbf{x}}] = ((\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll}) \cdot ([\mathbf{l}] - [\mathbf{a}_0]), \quad (4.43)$$

wobei der Intervallvektor  $[\mathbf{l}]$  die Impräzision der korrigierten Beobachtungen und der Intervallvektor  $[\mathbf{a}_0]$  die Impräzision der Modellannahmen des geodätischen Modells (Linearisierung, Schwerefeldmodell, Exzentren, etc.) beschreiben, vgl. Abschnitt 4.1.2. Mit der in (4.43) angegebenen Klammerung liefert  $[\mathbf{d}\hat{\mathbf{x}}]$  sogar die engste Intervalleinschließung des Vektors  $\mathbf{d}\hat{\mathbf{x}}$ , vergleiche hierzu Abschnitt 2.2.3 bzw. KUTTERER (1994, S.27). Die ausgeglichenen Parameter  $\hat{\mathbf{x}}$  erhält man mit (4.6)

$$[\hat{\mathbf{x}}] = \hat{\mathbf{x}} + [\mathbf{d}\hat{\mathbf{x}}], \quad (4.44)$$

wobei die Näherungswerte als exakt angenommen werden. Die Gleichungen (4.43) bzw. (4.44) lassen sich nach (2.22) in einen Mittelpunkt- und einen Radiusanteil zerlegen:

$$\hat{\mathbf{x}}_{\mathbf{m}} = \hat{\mathbf{x}} + (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} (\mathbf{l}_{\mathbf{m}} - \mathbf{a}_{0\mathbf{m}}), \quad (4.45)$$

$$\hat{\mathbf{x}}_{\mathbf{r}} = \left| (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} \right| (\mathbf{l}_{\mathbf{r}} + \mathbf{a}_{0\mathbf{r}}), \quad (4.46)$$

wobei  $|\cdot|$  den Absolutbetrag jedes Matrixelements der Produktmatrix bezeichnet. Die Intervallauswertung weist folgende Eigenschaften auf: Der Vektor der Intervallmitten ist identisch mit dem klassischen Schätzwert der Parameter (4.5), falls die Intervallmitten des Beobachtungsvektors  $\mathbf{l}_{\mathbf{m}}$  und des Absolutgliedvektors  $\mathbf{a}_{0\mathbf{m}}$  mit dem Vektor der korrigierten Messungen bzw. mit  $\mathbf{a}_0$  übereinstimmen. Bei einer Bestimmung der Intervalle mittels Sensitivitätsanalyse ist diese Forderung erfüllt, vgl. (2.74), Abschnitt 2.3.5. Diese Eigenschaften erlauben eine elegante Trennung in die eigentliche Parameterschätzung und Betrachtung der Stochastizität im Gauß-Markov-Modell einerseits und andererseits die Betrachtung der Impräzision durch Intervallradien. Diese beschreiben für jede Koordinate ihren maximalen symmetrischen Variationsbereich, der durch die Impräzision der Eingangsgrößen bestimmt ist. KUTTERER (1994) hat weiter gezeigt, dass die Intervalle  $[d\hat{x}_i]$  komponentenweise für jedes  $d\hat{x}_i$  eine exakte Intervalleinschließung liefern, d.h. das Intervall  $[d\hat{x}_i]$  gibt den Wertebereich der  $i$ -ten Komponente an, vgl. auch Abschnitt 2.3.5.

Zusammenfassend lassen sich folgende Eigenschaften der Intervallboxen hervorheben:

1. Die Seiten der Intervallbox sind parallel zu den Koordinatenachsen.

2. Die Intervallbox ist auf der Position der geschätzten Koordinaten des Neupunktes zentrisch gelagert.
3. Die Intervallbox stellt einen maximalen nichtstochastischen Variationsbereich für die Punktposition des Neupunktes dar.
4. Die Größe der Intervallbox (Flächeninhalt oder Volumen im Dreidimensionalen) oder die Seitenlängen können benutzt werden, um quantitative Maße für die Impräzision der Punktposition abzuleiten.

**Beispiel 21** Betrachten wir den ebenen Vorwärtsschnitt als einführendes Beispiel: Von zwei Festpunkten aus werden durch Messung von Richtungen zu einem Fernziel und zum Neupunkt die Koordinaten des Neupunktes bestimmt. Das Problem ist eindeutig bestimmt, da neben den zwei Koordinaten des Neupunktes auch zwei Orientierungsunbekannte mitbestimmt werden müssen.

Abbildung 4.1(a) zeigt die Intervallbox für den Neupunkt beim Vorwärtsschnitt. Sie liefert den maximalen Variationsbereich, in dem die Punktposition aufgrund der durch Intervalle beschriebenen Impräzision variieren kann. Abbildung 4.1(b) stellt die 95%-Konfidenzellipse dar, die eine Konfidenzregion für die Punktposition angibt.

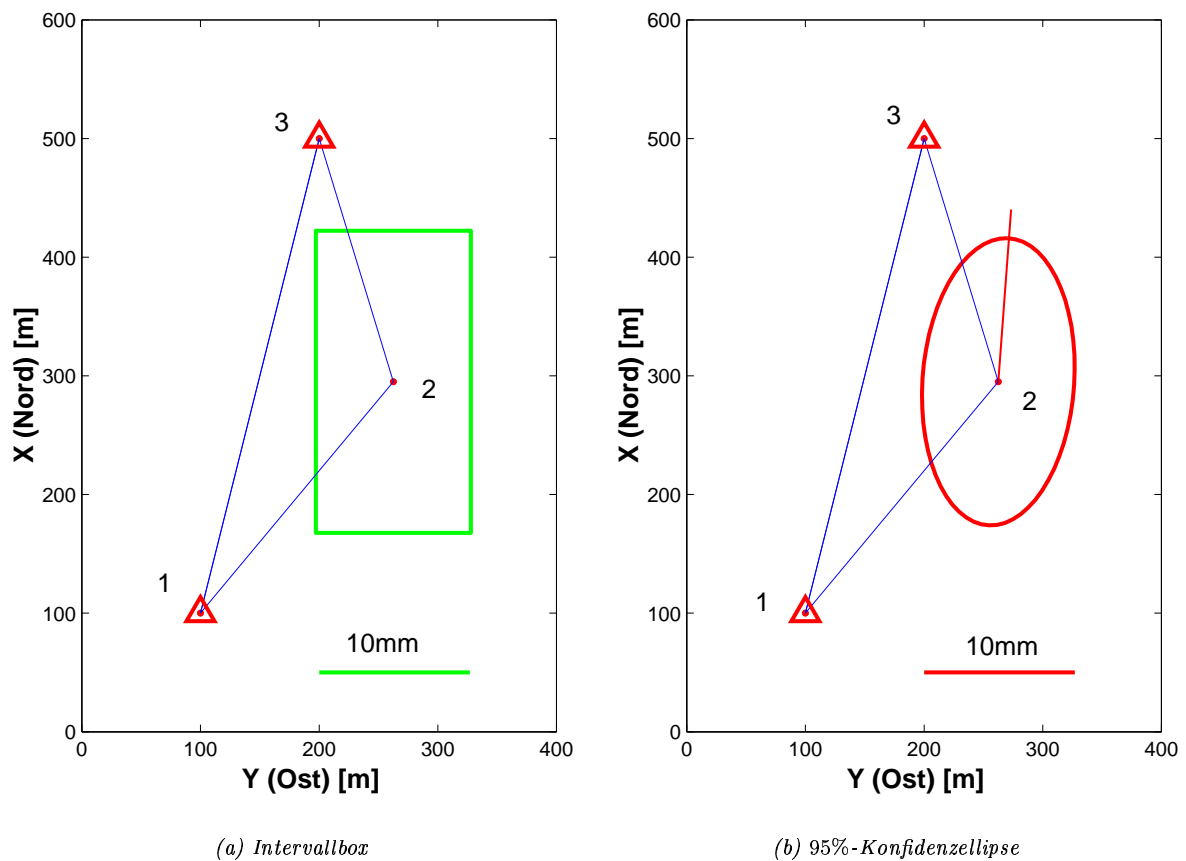


Abbildung 4.1: Darstellung der Intervallbox und der 95%-Konfidenzellipse. Intervallradien der in zwei Lagen gemessenen Richtungen nach Faustformel (3.4), Standardabweichung der Richtungsbeobachtung  $\sigma_R = 0.5$  mgon.

## 4.2.2 Wertebereich des Parametervektors $\hat{\mathbf{x}}$

Die Menge der Intervallvektoren ist allerdings nicht abgeschlossen gegenüber linearen Abbildungen, vgl. die Diskussion in Abschnitt 2.2.2. Folglich wird der tatsächliche Variationsbereich des Vektors aller Koordinaten nicht durch den Intervallvektor  $[\hat{\mathbf{x}}]$ , sondern durch den Wertebereich

$$\mathcal{W}_{\hat{\mathbf{x}}} := \left\{ \hat{\mathbf{x}} \in \mathbb{R}^u \mid \hat{\mathbf{x}} = (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} (\mathbf{1} - \mathbf{a}_0) + \overset{\circ}{\hat{\mathbf{x}}}, \mathbf{1} \in [\mathbf{1}], \mathbf{a}_0 \in [\mathbf{a}_0] \right\} \quad (4.47)$$

beschrieben. Die Intervallbox  $[\hat{\mathbf{x}}]$  stellt die engste Intervalleinschließung des Wertebereichs in der gegebenen Koordinatenbasis dar (KUTTERER 1994, S.26ff). Die Intervallbox ist somit die Intervallhülle des Wertebereichs

$$[\mathbf{d}\hat{\mathbf{x}}] = \square \mathcal{W}_{\hat{\mathbf{x}}}([\mathbf{l}], [\mathbf{a}_0]). \quad (4.48)$$

Sollen jedoch zwei oder dreidimensionale Punktkoordinaten gemeinsam betrachtet werden, so wird der Wertebereich der Punktposition überschätzt. Dies kann folgendermaßen verdeutlicht werden: Da die Punktkoordinaten durch (4.43) bei der Intervallauswertung immer getrennt betrachtet werden, werden beispielsweise die Abhängigkeiten zwischen der  $x$ -,  $y$ - oder  $z$ -Koordinate eines Punktes, die durch die lineare Abbildung erzeugt werden, nicht berücksichtigt, und der Wertebereich wird überschätzt. Weiter sind die Seiten der Intervallboxen jeweils parallel zu den Koordinatenachsen des gerade gegebenen Koordinatenrahmens, der je nach Anwendung gegebenenfalls willkürlich gewählt werden kann, wie in Abschnitt 4.1.1 begründet wurde. Durch eine Rotation des Koordinatenrahmens verändert sich die Form und Lage der Intervallboxen in Bezug auf die Geometrie des geodätischen Netzes, die ihrerseits unverändert bleibt. Das folgende Beispiel veranschaulicht diesen Effekt.

**Beispiel 22** Betrachten wir dazu ein Viereck, bei dem alle sechs Strecken gemessen wurden. Die vier Punktkoordinatenpaare des Lagenetzes werden mit freier Netzausgleichung unter Verwendung der Pseudoinversen gelöst. Die folgende Abbildung zeigt die Intervallboxen für die Intervallauswertung bezüglich zweier gegeneinander gedrehter Koordinatenrahmen.

Die in Abbildung 4.2(a) heller gezeichneten Intervallboxen beziehen sich auf den dargestellten Koordinatenrahmen. Sie weisen eher quadratische Form auf. Die dunkler dargestellten Intervallboxen beziehen sich auf einen um  $30^\circ$  gedrehten Rahmen. Das Verhältnis der Seitenlängen weicht für die Intervallboxen der Punkte 1, 3 und 4 in der gedrehten Basis stärker von Eins ab als in der Ausgangsbasis. Die Intervallbox des Punktes 2 ähnelt dagegen in der gedrehten Basis einem Quadrat.

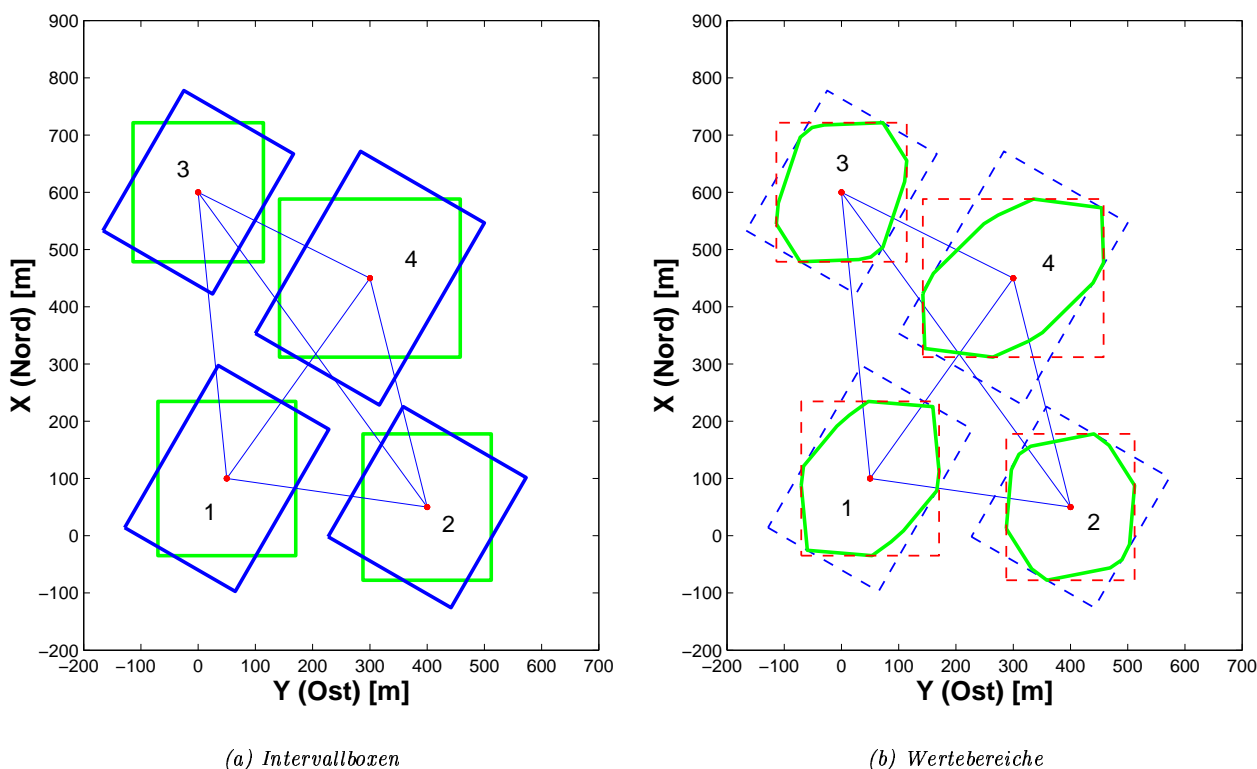


Abbildung 4.2: Vergleich der Intervallauswertung und der Wertebereiche eines Streckennetzes bezüglich zweier zueinander gedrehter Koordinatenrahmen. Die Intervallradien lassen sich nach der Faustformel (3.6) angeben.

Abbildung 4.2(b) stellt die Wertebereiche dar. Für jeden Punkt ist der Wertebereich ein 12-Eck, das in der Schnittmenge der Intervallboxen liegt, die für beide Koordinatenbasen gestrichelt eingezeichnet wurden. Da jede Intervallauswertung in ihrem Bezugssystem jeweils die engste Intervalleinschließung des Wertebereichs darstellt, muss mindestens eine Ecke des Wertebereichs auf jeder Seite der Intervallbox liegen. Die Schnittmenge

von Intervallboxen bezüglich verschiedener gedrehter Koordinatenrahmen lässt sich deshalb für eine einfache Approximation des Wertebereichs verwenden, vergleiche hierzu die Diskussion in Abschnitt 2.1.3.

### 4.2.3 Wertung

Die Wertung des Konzeptes der Intervallauswertung hängt von der verfolgten spezifischen Fragestellung ab, die effizient gelöst werden soll und kann somit von Anwendung zu Anwendung variieren. Für die Intervallauswertung der Schätzwerte im Sinne der Methode kleinsten Quadrate einer beliebigen Problemstellung weist das Konzept folgende Vorteile auf.

- Die Intervallbox schließt den Wertebereich in der gewählten Parameterbasis engstmöglich ein. Sind nur einzelne Parameter von Interesse und nicht der gesamte Parametervektor, so entspricht der Wertebereich eines einzelnen Parameters gerade dem Ergebnisintervall (komponentenweise exakte Einschließung), vgl. KUTTERER (1994, S.27).
- Die Intervallbox liefert somit einen sicheren oder garantierten Bereich, in dem die Schätzwerte maximal variieren können, wenn die Intervalle der Eingangsgrößen realistisch angenommen wurden.
- Durch die Zerlegung des Intervalls in Intervallmitte und -radius ermöglicht das Konzept eine elegante Trennung in die eigentliche Parameterschätzung (Intervallmitte) und die Betrachtung der Impräzision (Intervallradius).

Betrachten wir als Spezialfall der Parameterschätzung die geodätische Netzausgleichung. Hier sind ausgeglichene Punktpositionen, also Koordinatenpaare oder -tripel, und ihre Bewertung von Interesse. In diesem Kontext weist das Konzept einige Nachteile auf.

- Die Intervallbox schließt den Wertebereich zwar engstmöglich von außen ein, überschätzt ihn aber. Der Grad der Überschätzung kann nicht ohne Bestimmung des Wertebereichs angegeben werden, vgl. Abschnitt 2.2.2. Um ein Aufblähen der Intervalle von Größen, die aus Koordinaten abgeleitet werden, einzudämmen, sollte eine Intervallauswertung bei der Weiterverarbeitung erst beim letzten Rechenschritt vorgenommen werden, vgl. Abschnitt 2.2.2, Gleichung (2.24).
- Kenngrößen zur Qualifizierung der Impräzision, die sich auf die Intervallboxen beziehen, variieren mit der Orientierung des Koordinatenrahmens. Dadurch ist die Aussagekraft dieser Maße reduziert, da sie nicht allgemeingültig, also unabhängig von der Orientierung des Koordinatenrahmens sind. Da die Impräzision der Messung immanent ist, die ihrerseits nicht sensitiv bezüglich des Koordinatenrahmens ist, kann es zu Fehlinterpretationen dieser Maßzahlen kommen. So darf aus dem Vergleich der Intervalleinschließungen für die Netzpunkte in Abbildung 4.2 nicht geschlossen werden, dass die Impräzision der bestimmten Punktposition im einen Fall geringer sei als im anderen. Die Bestimmungsgüte hängt, wie in Abschnitt 4.1.1 beschrieben, nur von der Geometrie des Netzes, seinem Datum und der verwendeten Messverfahren ab.

Der erste Kritikpunkt ist hauptsächlich für geodätische Netze von Bedeutung. Hier sind Punktpositionen, also Koordinatenpaare oder -tripel, von Interesse. Bei allen anderen Ausgleichungsproblemen, bei denen einzelne, unabhängige Parameter bestimmt werden, eignen sich Intervalle direkt sehr gut zur Beschreibung der Impräzision.

Es muss an dieser Stelle ausdrücklich unterstrichen werden, dass der zweite Kritikpunkt nur von Bedeutung ist, wenn die Orientierung des Koordinatenrahmens frei wählbar ist wie beispielsweise bei der Planung, Optimierung und Analyse von Netzen aus dem Bereich der Ingenieurgeodäsie. Arbeitet man mit konventionellen Referenzrahmen wie dem ITRF oder Systemen der Landesvermessung, so ist der Koordinatenrahmen festgelegt und dieser Freiheitsgrad entfällt.

In anderen Disziplinen, die ebenfalls mit Intervallmathematik arbeiten, spielen diese Kritikpunkte eine eher untergeordnete Rolle. Man wählt sich in der Regel ein Bezugssystem, in dem für diese Anwendung die Berechnung und Beurteilung durchgeführt werden sollen. Hier steht oftmals der Aspekt der garantierten Parameterbestimmung im Vordergrund, vgl. KIEFFER ET AL. (2000). Die Intervallauswertung für den Schätzwert nach der Methode der kleinsten Quadrate ist sehr einfach. Für andere Fragestellungen wie das Lösen von Differentialgleichungen gestaltet sich die Berechnung der Ergebnisintervalle so kompliziert, so dass man sich mit einer

Einschließung der Lösung für eine konkrete Aufgabenstellung zufrieden gibt. In anderen Anwendungsbereichen wird einfach auf die komponentenweise Betrachtung der Intervallvektoren zurückgegangen.

Das Konzept des Wertebereichs liefert die exakte Lösung und ist für alle Aufgabenstellungen, die mit der Methode der kleinsten Quadrate gelöst werden sollen, anwendbar. Es ist somit von Bedeutung, dieses übergeordnete und allgemeinere Konzept auszuarbeiten. Inwieweit es der Intervallauswertung vorzuziehen ist, hängt von der Zielsetzung, Aufgabenstellung und dem damit verbundenen Aufwand ab und kann nicht allgemein beantwortet werden. Für Aufgabenstellungen, bei denen Abstandsmaße (Hausdorff-Distanz) zwischen Intervallboxen als Kriterien herangezogen werden sollen, beispielsweise als Abbruchkriterium in Algorithmen, kann die Orientierung des Koordinatenrahmens von Bedeutung sein. Im Konzept der affinen Arithmetik wird der Wertebereich zur Einschließung von Punktpositionen bei Anwendungen in der Computergraphik genutzt (DE FIGUIEREDO ET AL. 2003). Ein Kernpunkt der vorliegenden Arbeit stellt die Erarbeitung des Konzeptes des Wertebereichs dar. Hierzu werden im nächsten Kapitel einige mathematische Grundlagen zu konvexen Polytopen gelegt.

## 4.3 Zonotope

Im folgenden Abschnitt werden Zonotope vorgestellt, die mit ihren besonderen kombinatorischen und geometrischen Eigenschaften eine spezielle Klasse innerhalb der konvexen Polytope darstellen. Konvexe Polytope kann man als eine Verallgemeinerung des zweidimensionalen konvexen  $n$ -Ecks auffassen. Zusammen mit konvexen Polyedern sind sie die Grundobjekte der diskreten und kombinatorischen Geometrie. Die klassische Literatur zur Theorie der konvexen Polytope ist GRÜNBAUM (1967) und für die Darstellung regulärer Polytope vergleiche COXETER (1973). Ein modernes Buch zu diesem Thema ist ZIEGLER (1995). Einen Überblick über konvexe Polytope und verwandte Themen im Umfeld der diskreten und kombinatorischen Geometrie bieten die Handbücher GRUBER UND WILLIS (1993) und GOODMAN UND O'ROURKE (1997).

### 4.3.1 Grundlagen zu Zonotopen

#### Definitionen und geometrische Veranschaulichung

**Definition 13** Ein Polytop  $\mathcal{P}$  ist die konvexe Hülle einer endlichen Teilmenge  $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_m\} \in \mathbb{R}^{u \times m}$  in der Form:

$$\mathcal{P} = \mathcal{P}(\mathbf{V}) := \text{conv}(\mathbf{V}) = \left\{ \sum_{i=1}^m \lambda_i \mathbf{v}_i \mid \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\}. \quad (4.49)$$

Man spricht hierbei von der *V-Darstellung* (ZIEGLER 1995, S.4, S.29), da sich die Vektoren  $\mathbf{v}_i$  als Koordinaten der Eckpunkte des Polytopes auffassen lassen.

**Definition 14** Die abgeschlossene Lösungsmenge eines Systems endlich vieler linearer Ungleichungen heißt *Polytop* und es gilt:

$$\mathcal{P} = \mathcal{P}(\mathbf{H}, \mathbf{b}) := \left\{ \mathbf{x} \in \mathbb{R}^u \mid \mathbf{H}\mathbf{x} \leq \mathbf{b} \right\}. \quad (4.50)$$

Die Zeilen der  $n \times u$  Matrix  $\mathbf{H}$  lassen sich als Normalenvektoren von  $n$  Hyperebenen auffassen, deren Schnittmenge das Polytop ergibt. Man spricht deshalb auch von einer *H-Darstellung* eines Polytopes.

Beide Darstellungen sind im Sinn des Hauptsatzes für Polytope äquivalent, vgl. ZIEGLER (1995, Theorem 1.1, S.29) oder HENK ET AL. (1997, Theorem 13.1.1, S.244).

**Definition 15** Ist die Lösungsmenge (4.50) nicht abgeschlossen, so heißt  $\mathcal{P}$  *Polyeder* (ZIEGLER 1995, S.4).

Im Folgenden werden Zonotope als spezielle Polytope betrachtet. Zonotope treten in vielen technischen Anwendungen auf und können unter verschiedenen Blickwinkeln betrachtet und eingeordnet werden, wobei jeder Ansatz einen besonderen Zugang und besondere Interpretationsmöglichkeiten eröffnet. Die Darstellungen orientieren sich an ZIEGLER (1995, Kap.7.2, 7.3).

**Definition 16** Sei  $\mathcal{C}_n$  der  $n$ -dimensionale Hyperwürfel

$$\mathcal{C}_n := \{ \mathbf{w} \in \mathbb{R}^n \mid -1 \leq w_i \leq 1, \quad i = 1, \dots, n \} = [-\mathbf{e}, \mathbf{e}], \tag{4.51}$$

wobei  $\mathbf{e} = (1, \dots, 1)^T$  den  $n$ -dimensionalen Einsvektor darstellt und  $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^u$  die affine Abbildung

$$\pi : \begin{cases} \mathcal{C}_n & \rightarrow \mathcal{Z} \\ \mathbf{w} & \mapsto \mathbf{K} \mathbf{w} + \mathbf{z}_0, \end{cases} \tag{4.52}$$

so ist das Bild  $\mathcal{Z} = \pi(\mathcal{C}_n)$  des Hyperwürfels bei der affinen Abbildung ein *Zonotop*. Sei  $\mathbf{K} = (\mathbf{k}_1, \dots, \mathbf{k}_n)$  eine  $u \times n$  Matrix, so lässt sich  $\mathcal{Z}$  als  $u$ -dimensionales Polytop (*u-Zonotop*) folgendermaßen darstellen:

$$\mathcal{Z} = \mathcal{Z}(\mathbf{K}) := \mathbf{K} \mathcal{C}_n + \mathbf{z}_0 \tag{4.53}$$

$$= \{ \mathbf{K} \mathbf{w} + \mathbf{z}_0 \mid \mathbf{w} \in \mathcal{C}_n \} \tag{4.54}$$

$$= \left\{ \mathbf{z} \in \mathbb{R}^u \mid \mathbf{z} = \mathbf{z}_0 + \sum_{i=1}^n w_i \mathbf{k}_i, \quad -1 \leq w_i \leq 1 \right\}. \tag{4.55}$$

**Definition 17** Die *Minkowski-Summe* oder *Vektorsumme* zweier Mengen  $\mathcal{M}_1, \mathcal{M}_2 \subseteq \mathbb{R}^m$  ist definiert als

$$\mathcal{M}_1 \oplus \mathcal{M}_2 := \{ \mathbf{x} + \mathbf{y} \mid \mathbf{x} \in \mathcal{M}_1, \mathbf{y} \in \mathcal{M}_2 \}. \tag{4.56}$$

Damit lässt sich eine häufig verwendete, äquivalente zweite Definition von Zonotopen angeben.

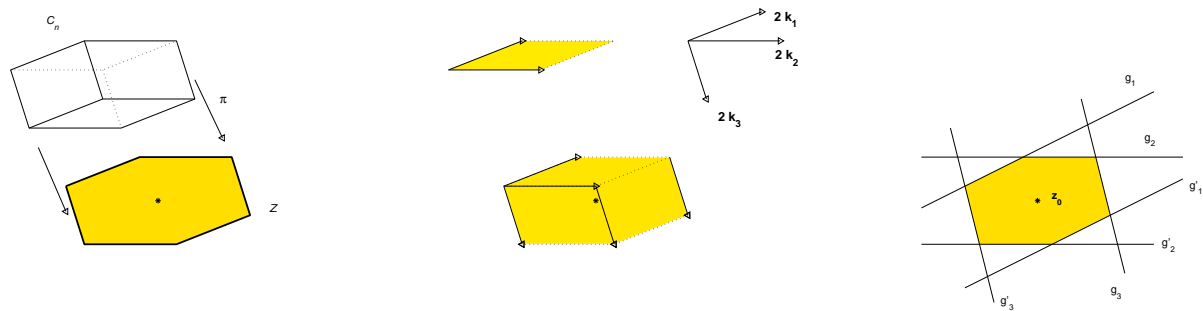
**Definition 18** Ein *Zonotop* ist die Minkowski-Summe von  $n$  Liniensegmenten

$$\mathcal{Z}(\mathbf{K}) := [-\mathbf{k}_1, \mathbf{k}_1] \oplus [-\mathbf{k}_2, \mathbf{k}_2] \oplus \dots \oplus [-\mathbf{k}_n, \mathbf{k}_n]. \tag{4.57}$$

Ein Zonotop lässt sich auch als Schnittmenge von  $n$  parallelen Hyperebenen auffassen

$$\mathcal{Z} = \mathcal{Z}(\mathbf{H}, \mathbf{b}) = \{ \mathbf{z} \in \mathbb{R}^u \mid \mathbf{H}(\mathbf{z} - \mathbf{z}_0) \leq \mathbf{b} \}. \tag{4.58}$$

Die Umkehrung gilt nur im Zweidimensionalen. In höheren Dimensionen ist nicht jede abgeschlossene Schnittmenge von  $n$  Hyperebenen ein Zonotop. Das Oktaeder ist ein solches einfaches Gegenbeispiel (EPPSTEIN 1995, S.9). Die folgenden Graphiken veranschaulichen die angedeuteten Konstruktionsprinzipien von Zonotopen für das gleiche Beispiel.



(a) Zonotop als Projektion des Hyperwürfels

(b) Zonotop als Minkowski-Summe von Liniensegmenten

(c) Zonotop als Hyperebenen-Arrangement

Abbildung 4.3: Konstruktion von Zonotopen

Abbildung 4.3(a) veranschaulicht die Projektion eines Würfels und die Entstehung eines Zonotops nach Definition 16. Der Rand des Zonotops lässt sich mit den Spaltenvektoren der Matrix  $\mathbf{K}$  (Kantenvektoren) als geschlossener Linienzug darstellen. Abbildung 4.3(b) zeigt die Konstruktion eines Zonotops als Minkowski-Summe von Vektoren. Die Minkowski-Summe der Vektoren  $[-\mathbf{k}_1, \mathbf{k}_1] = 2 \mathbf{k}_1$  und  $[-\mathbf{k}_2, \mathbf{k}_2] = 2 \mathbf{k}_2$  ergibt das Parallelogramm, das im oberen Teil der Abbildung gezeichnet ist. Wird dann der dritte Vektor  $2 \mathbf{k}_3$  zum Ergebnis dazu addiert, so entsteht das Zonotop. Es wird deutlich, dass sich 2-Zonotope in Parallelogramme und 3-Zonotope in Parallelepipede zerlegen lassen, wobei die Zerlegung nicht eindeutig ist (SHEPHARD 1974). Abbildung 4.3(c) stellt die Schnittmenge dreier paralleler Hyperebenen dar. Ihre Normalenvektoren stehen senkrecht auf den Kantenvektoren des Zonotops, die durch die Spaltenvektoren der Matrix  $\mathbf{K}$  gegeben sind.



### Einige Eigenschaften von Zonotopen

1. Jedes Zonotop ist ein konvexes Polytop.
2. Jedes Zonotop ist *zentralsymmetrisch* bezüglich des Zentrums  $\mathbf{z}_0$ , so dass gilt:

$$\mathcal{Z} - \mathbf{z}_0 = -(\mathcal{Z} - \mathbf{z}_0). \quad (4.59)$$

Diese Eigenschaft charakterisiert Zonotope innerhalb der Klasse der konvexen Polytope. Es gilt weiter, dass jedes Polytop, dessen  $k$ -dimensionale Randflächen ( $k = 3, \dots, u - 2$ ) und Kanten zentralsymmetrisch sind, ein Zonotop ist (ZIEGLER 1995, S.200). Daher sind bei 3-Zonotopen je zwei Randflächen parallel und bei 2-Zonotopen je zwei Kantenvektoren parallel.

3. Die Eigenschaft „Zonotop“ bleibt unter affinen Projektionen erhalten, d.h. jede affine Projektion eines Zonotops ist wieder ein Zonotop (ZIEGLER 1995, S.201). Insbesondere ändert sich die Form des Zonotops nicht bei Rotationen.
4. Die Anzahl der Ecken  $v$ , Kanten  $e$  und Randflächen  $f$  eines 3-Zonotops berechnen sich nach COXETER (1973, S.27):

$$v = m(m - 1) + 2, \quad e = 2m(m - 1), \quad f = m(m - 1), \quad (4.60)$$

wobei  $m = n - k$  die Anzahl der Spaltenvektoren der  $u \times n$   $\mathbf{K}$  angibt, die nicht kollinear sind und  $k$  die Anzahl der kollinearen Vektoren. Für 2-Zonotope gilt

$$v = 2m \quad \text{und} \quad e = 2m. \quad (4.61)$$

Die Gleichungen (4.60) und (4.61) verdeutlichen, dass nur ein Bruchteil der  $2^n$  Ecken des Hyperwürfels in Ecken des Zonotops abgebildet wird. Die restlichen Ecken werden in das Innere des Zonotops projiziert. Da das 3-Zonotop ein reguläres Polyeder ist, gilt für die Anzahl der Ecken, Kanten und Randflächen die *Eulersche Polyederformel* (ZIEGLER 1995, S.231):

$$v - e + f = 2. \quad (4.62)$$

5. Sind keine der  $m$  Spaltenvektoren in  $\mathbf{K}$  kollinear, so sind alle Randflächen des 3-Zonotops Parallelogramme.
6. Unter einer *Zone* versteht man alle Randflächen, die einen gemeinsamen Kantenvektor  $\mathbf{k}$  haben (ZIEGLER 1995, S.221). Im Dreidimensionalen umläuft die Zone das Zonotop wie ein Band:

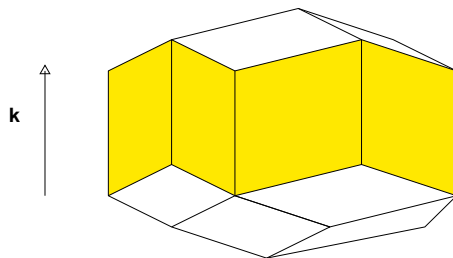


Abbildung 4.4: Zur Veranschaulichung des Begriffs Zone

7. Das Volumen von Zonotopen lässt sich nach der *Formel von McMullan* aus den Absolutbeträgen der Determinanten aller  $u \times u$  Submatrixen von  $\mathbf{K}$  berechnen, vgl. ZIEGLER (1995, S.230) oder HENK ET AL. (1997, S.264):

$$Vol(\mathcal{Z}) := 2^u \sum_{1 \leq i_1 < i_2 < \dots < i_u \leq n} \left| \det(\mathbf{k}_{i_1}, \dots, \mathbf{k}_{i_u}) \right|. \quad (4.63)$$

Diese Formel lässt sich anschaulich erklären: Im Zweidimensionalen entspricht diese Zerlegung in Indexgruppen  $\{i_1, i_2, \dots, i_u\}$  gerade der Zerlegung des Zonotops in Parallelogramme und im Dreidimensionalen in Parallelepipede, deren Volumen sich jeweils aus der Determinante der Matrix der Kantenvektoren berechnet, vgl. auch SHEPHARD (1974) oder ZIEGLER (1995, S.230). Diese Determinanten lassen sich als *Plücker-Koordinaten* interpretieren, die die homogenen Koordinaten des Zonotops unter projektiver Abbildung darstellen (BJÖRNER ET AL. 1999). Auf diese Interpretation soll in der vorliegenden Arbeit nicht weiter eingegangen werden.

### Ecken des Zonotops

Im Folgenden werden Aussagen über die Ecken des Zonotops angegeben und verschiedene Verfahren zu ihrer Konstruktion vorgestellt. Die Ecken sind die Extrempunkte des Polytops. Da  $\pi$  eine lineare Abbildung ist, ist jede Ecke des Zonotops das Bild einer Ecke des Hyperwürfels und damit in der folgenden Form darstellbar:

$$Vert_i = \sum_{j=1}^n \nu_j \mathbf{k}_j, \quad \nu_j \in \{-1, 1\}, \tag{4.64}$$

wobei  $\nu$  den  $n \times 1$  Vorzeichenvektor beschreibt. Aus Gleichung (4.60) bzw. (4.61) wird deutlich, dass nicht jedem Vorzeichenvektor eine Ecke des Zonotops entspricht, vielmehr viele Vorzeichenvektoren innere Punkte des Zonotops beschreiben. Die Vorzeichenvektoren, die die Ecken identifizieren, sind die *maximalen Kovektoren* des *orientierten Matroids*, das dem Zonotop zugeordnet ist (BJÖRNER ET AL. 1999, S.52). Orientierte Matroide beschreiben die kombinatorischen Eigenschaften von Punktfigurationen, Hyperebenenanordnungen, gerichteten Graphen oder dem Dualitätsprinzip der linearen Optimierung. In dieser Arbeit sollen diese Zusammenhänge nicht weiter vertieft werden. Der interessierte Leser sei auf ZIEGLER (1995, Kap.7.3, 7.4) bzw. BJÖRNER ET AL. (1999) verwiesen.

Die Konstruktion des Zonotops aus den Kantenvektoren kann als kombinatorische Aufgabe aufgefasst werden. Graphisch bedeutet dies, man gibt sich den Mittelpunkt  $\mathbf{z}_0$  des Zonotops vor und beginnt von dort aus sukzessive die Kantenvektoren mit positivem oder negativem Vorzeichen anzuhängen. Ist dies für alle Vektoren durchgeführt, so hat man  $2^n$  Vektorzüge gezeichnet. Die konvexe Hülle der Endpunkte stellt das Zonotop dar, wobei die in (4.60) bzw. (4.61) angegebene Eckenzahl erfüllt wird.

Abbildung 4.5 veranschaulicht die Konstruktion der Ecken des Zonotops als Kombination der Kantenvektoren  $\mathbf{k}$ , die im oberen Teil der Abbildung zu sehen sind. Für die einzelnen Eckpunkte sind jeweils die Vorzeichenvektoren  $\nu$  angegeben. In diesem Beispiel sind die Vorzeichenvektoren  $\nu = (+1, -1, +1)^T$  und  $\nu = (-1, +1, -1)^T$  inneren Punkten zugeordnet. Gegenüberliegende Eckpunkte haben wegen der Zentralsymmetrie entgegen gerichtete Vorzeichenvektoren. Dieses Verfahren eignet sich nur für kleine Konfigurationen, da der Rechenaufwand exponentiell mit  $n$  wächst.

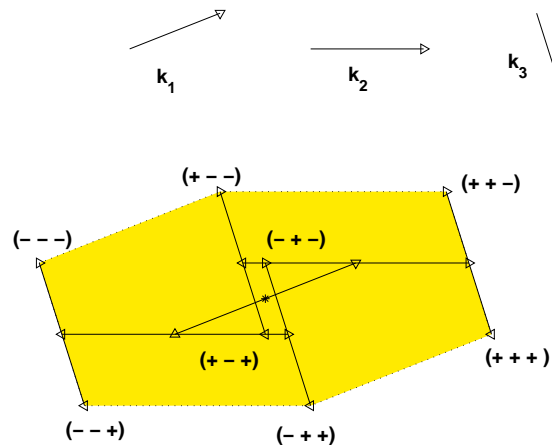


Abbildung 4.5: Darstellung der Ecken-Kombinatorik des 2-Zonotops

Als Alternative bietet sich für 2-Zonotope folgende Strategie an: Ist ein Eckpunkt bekannt, so werden alle weiteren Eckpunkte durch sukzessives Anhängen der zyklisch geordneten Kantenvektoren bestimmt. Unter der zyklischen Reihenfolge soll eine Ordnung  $\succeq$  der Kantenvektoren nach zunehmenden Richtungswinkel verstanden werden:

$$\mathbf{k}_i \succeq \mathbf{k}_j \Leftrightarrow \varphi_i \geq \varphi_j, \tag{4.65}$$

wobei  $\varphi_i, \varphi_j \in [0, 2\pi]$  die Richtungswinkel der Kantenvektoren sind, die in das Intervall  $[0, 2\pi]$  transformiert wurden. Der Startpunkt lässt sich bei dieser Reihenfolge mit Hilfe der Intervallauswertung bestimmen. Da die Intervallbox die Intervallhülle des Zonotops ist, liegt auf jeder Kante der Intervallbox mindestens ein Eckpunkt des Zonotops. Als Startpunkt ist derjenige von Interesse, für den in einem geodätischen Koordinatensystem gilt:

$y = -y_r$ . Bezeichne  $\mathbf{k}^1$  die erste Zeile der  $2 \times n$  Submatrix von  $\mathbf{K}$  und  $\text{sign}(\mathbf{k}^1)$  den Vorzeichenvektor, so lässt sich schreiben:

$$y_r = -\mathbf{k}^1 (\text{sign}(\mathbf{k}^1))^T. \quad (4.66)$$

Die Koordinaten des ersten Eckpunktes lauten dann:

$$\text{Vert}_1 := \mathbf{z}_0 - \mathbf{K} \begin{pmatrix} \text{sign}(\mathbf{k}^1) \\ \text{sign}(\mathbf{k}^1) \end{pmatrix}^T. \quad (4.67)$$

Alle anderen Eckpunkte lassen sich mit der Iterationsvorschrift berechnen:

$$\text{Vert}_i := \text{Vert}_{i-1} + 2\mathbf{k}_i, \quad i = 1, \dots, n, \quad (4.68)$$

wobei wegen der Zentralsymmetrie für die restlichen  $n$  Ecken gilt:  $\text{Vert}_{n+i} := \mathbf{z}_0 - \text{Vert}_i$ .

Eine dritte elegante Berechnungsmethode der Ecken stellt die folgende lineare Optimierungsaufgabe dar,

$$\begin{aligned} \mathbf{f}^T \mathbf{z} &\rightarrow \min, \\ \mathbf{H}(\mathbf{z} - \mathbf{z}_0) &\leq \mathbf{l}_r, \\ -\mathbf{H}(\mathbf{z} - \mathbf{z}_0) &\leq -\mathbf{l}_r, \end{aligned} \quad (4.69)$$

bei der das Zonotop die Restriktionsmenge bildet. Hierzu muss das Zonotop allerdings in der H-Darstellung vorliegen.

### 4.3.2 Wertebereich als Zonotop

Im folgenden Abschnitt wird gezeigt, dass der Wertebereich  $\mathcal{W}_{\hat{\mathbf{x}}}$  ein Zonotop ist. Dazu soll die Darstellung des Wertebereichs nach Gleichung (4.47)

$$\mathcal{W}_{\hat{\mathbf{x}}} := \left\{ \hat{\mathbf{x}} \in \mathbb{R}^u \mid \hat{\mathbf{x}} = (\bar{\mathbf{A}}^T \mathbf{P} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P} (\mathbf{l} - \mathbf{a}_0) + \overset{\circ}{\hat{\mathbf{x}}}, \mathbf{l} \in [\mathbf{l}], \mathbf{a}_0 \in [\mathbf{a}_0] \right\} \quad (4.70)$$

auf die Darstellung des Zonotops nach Definition 16 umgeformt werden. Da  $\mathbf{l} \in [\mathbf{l}]$  und  $\mathbf{a}_0 \in [\mathbf{a}_0]$  unabhängig von einander variieren können, liegt der Differenzvektor  $\mathbf{l} - \mathbf{a}_0$  im Intervallvektor  $[\mathbf{l}] - [\mathbf{a}_0]$ . Dieser lässt sich durch den Hyperwürfel  $\mathcal{C}_n$  folgendermaßen darstellen:

$$[\mathbf{l}] - [\mathbf{a}_0] = \mathbf{l}_m - \mathbf{a}_{0m} + (\mathbf{L}_r + \mathbf{A}_{0,r}) \cdot \mathcal{C}_n, \quad (4.71)$$

wobei  $\mathbf{L}_r$  und  $\mathbf{A}_{0,r}$  die Diagonalmatrizen bezeichnen, deren Diagonalelemente die (positiven) Intervallradien sind. Gleichung (4.70) lässt sich dann schreiben als:

$$\mathcal{W}_{\hat{\mathbf{x}}} := \left\{ \hat{\mathbf{x}} \in \mathbb{R}^u \mid \hat{\mathbf{x}} = (\bar{\mathbf{A}}^T \mathbf{P} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P} (\mathbf{l}_m - \mathbf{a}_{0m} + (\mathbf{L}_r + \mathbf{A}_{0,r}) \mathbf{w}) + \overset{\circ}{\hat{\mathbf{x}}}, \mathbf{w} \in \mathcal{C}_n \right\}. \quad (4.72)$$

Mit der Darstellung der Intervallmitten der geschätzten Zuschläge (4.45) und der Abkürzung

$$\mathbf{K} := (\bar{\mathbf{A}}^T \mathbf{P} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P} (\mathbf{L}_r + \mathbf{A}_{0,r}) \quad (4.73)$$

erhält man dann den Wertebereich in der Form der Definition 16:

$$\mathcal{W}_{\hat{\mathbf{x}}} := \left\{ \hat{\mathbf{x}} \in \mathbb{R}^u \mid \hat{\mathbf{x}} = \overset{\circ}{\hat{\mathbf{x}}} + \mathbf{d}\hat{\mathbf{x}}_m + \mathbf{K}\mathbf{w}, \mathbf{w} \in \mathcal{C}_n \right\}. \quad (4.74)$$

Neben dem Wertebereich des gesamten Koordinatenvektors können durch Orthogonalprojektion  $\pi_{\perp}$  die Wertebereiche der Punktposition einzelner Netzpunkte als lokale Maße abgeleitet werden:

$$\pi_{\perp} : \begin{cases} \mathbb{R}^u \rightarrow \mathbb{R}^2 \\ \hat{\mathbf{x}}_P = \mathbf{\Gamma} \hat{\mathbf{x}} \end{cases}, \quad \text{bzw.} \quad \pi_{\perp} : \begin{cases} \mathbb{R}^u \rightarrow \mathbb{R}^3 \\ \hat{\mathbf{x}}_P = \mathbf{\Gamma} \hat{\mathbf{x}} \end{cases}, \quad (4.75)$$

wobei  $\mathbf{\Gamma}$  die Selektionsmatrix ist, die die Koordinaten eines Punktes auswählt. Die Matrix  $\mathbf{K}_P = \mathbf{\Gamma} \mathbf{K}$  stellt die Matrix der Kantenvektoren des 2-Zonotops bzw. 3-Zonotops dar, das die Impräzision der Punktposition beschreibt.

Der Wertebereich lässt sich somit in der Form (4.74) als Zonotop darstellen. Der Mittelpunkt des Zonotops  $\mathbf{z}_0 := \overset{\circ}{\hat{\mathbf{x}}} + \mathbf{d}\hat{\mathbf{x}}_m = \hat{\mathbf{x}}_m$  ist die geschätzte Punktposition. Durch die Eigenschaften des Wertebereichs als Zonotop

werden die Kritikpunkte (vgl. Abschnitt 4.2.3) am Konzept der Intervallboxen als Impräzisionsmaße für geodätische Netze ausgeräumt, denn der Wertebereich ist einerseits die tatsächliche Variationsregion des Vektors der Koordinaten, so dass es per Definition zu keiner Überschätzung kommt. Die Überschätzung des Wertebereichs durch die Intervallbox lässt sich mit (4.63) folgendermaßen angeben:

$$v := \frac{Vol([\hat{\mathbf{x}}]) - Vol(\mathcal{W}(\hat{\mathbf{x}}))}{Vol(\mathcal{W}(\hat{\mathbf{x}}))}. \quad (4.76)$$

Andererseits ist das Zonotop unabhängig von Rotationen des Koordinatenrahmens, vgl. Eigenschaft (3). Zusätzlich enthält der Wertebereich noch weitere Informationen zur Charakterisierung der Impräzision, auf die in den nachfolgenden Beispielen eingegangen werden soll.

**Beispiel 23** Betrachten wir als erstes Beispiel das polare Anhängen eines Neupunkts im Dreidimensionalen. Von einem koordinierten Punkt aus werden Richtung, Strecke und Zenitdistanz zum Neupunkt gemessen sowie die Richtung zu einem bekannten Fernziel. Es soll die Modellierung der Intervallradien aus Abschnitt 3.1 verwendet werden:  $R_{r,12} = 0.47 \text{ mgon}$ ,  $R_{r,13} = 1.55 \text{ mgon}$  bzw.  $S_{r,12} = 1.7 \text{ mm}$  und  $\zeta_{r,12} = 0.81 \text{ mgon}$  für eine Strecke von  $200 \text{ m}$  bei einer Zenitdistanz von  $94.5 \text{ gon}$ .

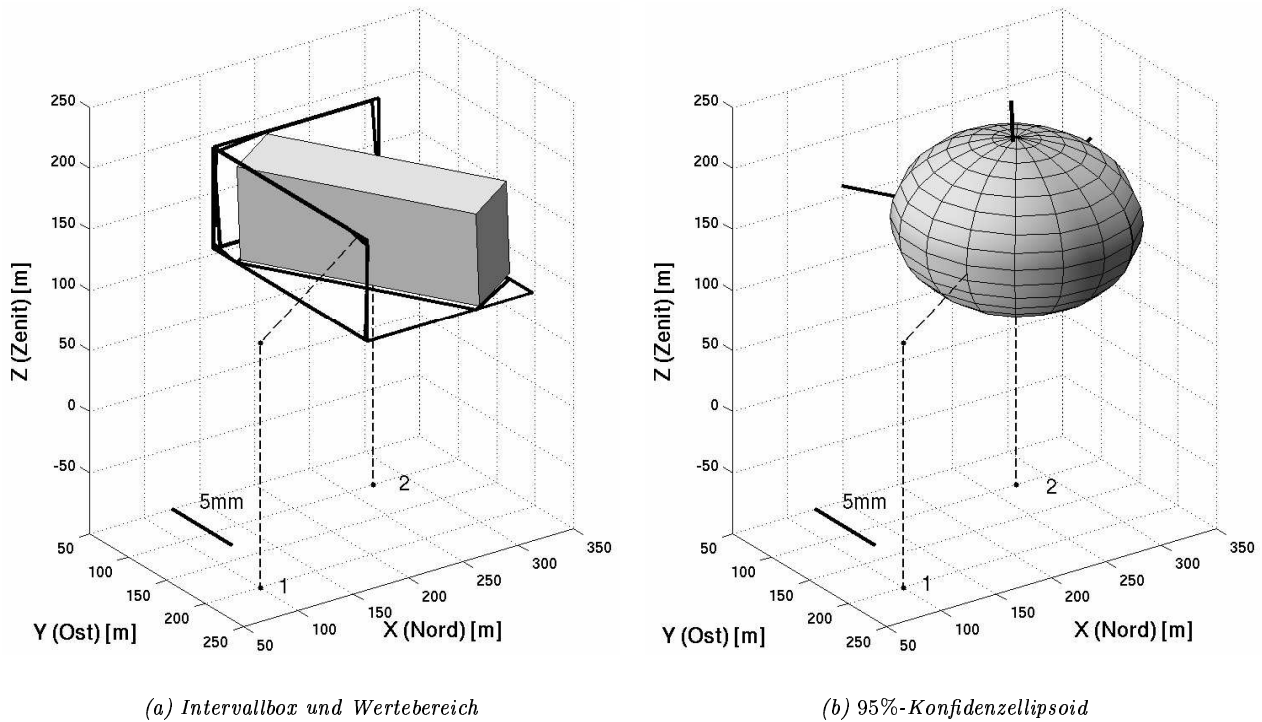


Abbildung 4.6: Punktunsicherheitsbereiche für das polare Anhängen. Die Standardabweichungen der Richtung und Zenitdistanz betragen je  $0.5 \text{ mgon}$ , die der Strecke  $2.4 \text{ mm}$ .

Abbildung 4.6(a) stellt die Intervallbox und den Wertebereich für das dreidimensionale polare Anhängen dar. Die Intervallbox ist parallel zu den Koordinatenachsen ausgerichtet, wohingegen der Wertebereich hier parallel zur Messungslinie ausgerichtet ist. Abbildung 4.6(b) zeigt das 95%-Konfidenzellipsoid. Es gibt den stochastischen Anteil der Gesamtunsicherheit an. Die große Halbachse steht in diesem Beispiel ebenfalls senkrecht zur Messungslinie.

Da dieses Beispiel noch relativ einfach analytisch darstellbar ist, soll die Matrix  $\mathbf{K}$  der Kantenvektoren angegeben werden. Mit  $\varrho = \frac{200 \text{ gon}}{\pi}$  ergibt sich:

$$\mathbf{K} = \begin{pmatrix} \cos \varphi \sin \zeta S_r & -\sin \varphi \sin \zeta S \frac{R_{r,1}}{\varrho} & \sin \varphi \sin \zeta S \frac{R_{r,2}}{\varrho} & \cos \varphi \cos \zeta S \frac{\zeta_r}{\varrho} \\ \sin \varphi \sin \zeta S_r & \cos \varphi \sin \zeta S \frac{R_{r,1}}{\varrho} & \cos \varphi \sin \zeta S \frac{R_{r,2}}{\varrho} & \sin \varphi \cos \zeta S \frac{\zeta_r}{\varrho} \\ \cos \zeta S_r & 0 & 0 & \sin \zeta S \frac{\zeta_r}{\varrho} \end{pmatrix}. \quad (4.77)$$

Die erste Spalte ist der Strecke zugeordnet, die zweite und dritte Spalte der Richtungsmessung und die vierte Spalte der Zenitdistanz. Der zweite und dritte Spaltenvektor sind kollinear, alle anderen Spaltenvektoren stehen

paarweise senkrecht aufeinander. Daher ist das Zonotop in diesem Beispiel ein Quader mit folgenden Seitenlängen und Volumen:

$$\|\mathbf{k}_1\| = S_r, \quad (4.78)$$

$$\|\mathbf{k}_2\| = \sqrt{(\Delta x^*)^2 + (\Delta y^*)^2} \frac{R_{r1}}{\varrho} = \sin \zeta S \frac{R_{r1}}{\varrho}, \quad (4.79)$$

$$\|\mathbf{k}_3\| = \sqrt{(\Delta x^*)^2 + (\Delta y^*)^2} \frac{R_{r2}}{\varrho} = \sin \zeta S \frac{R_{r2}}{\varrho}, \quad (4.80)$$

$$\|\mathbf{k}_4\| = S \frac{\zeta_r}{\varrho}, \quad (4.81)$$

$$Vol(\mathcal{W}_{\mathbf{x}}) = S_r \frac{S \zeta_r}{\varrho} \frac{\sqrt{(\Delta x^*)^2 + (\Delta y^*)^2} (R_{r,1} + R_{r,2})}{\varrho}. \quad (4.82)$$

Der Kantenvektor  $\mathbf{k}_1$  gibt die Impräzision der Streckenmessung, die Kantenvektoren  $\mathbf{k}_2$  und  $\mathbf{k}_3$  geben die Impräzision der Richtungsbeobachtungen (umgerechnet als Quermaß) und  $\mathbf{k}_4$  die Impräzision der Zenitdistanzbeobachtung (umgerechnet als Quermaß) an. Die Kantenvektoren des Wertebereichs verlaufen senkrecht bzw. parallel zur Verbindungslinie zwischen Standpunkt und Neupunkt. Das Restvolumen zwischen dem Quader des Wertebereichs und der Intervallbox stellt die Überschätzung des Wertebereichs dar, vgl. (4.76). Sie beträgt in diesem Beispiel  $v \approx 68\%$ .

Es wird deutlich, dass der Wertebereich mehr Informationen über die Impräzision der Punktlage enthält als die Intervallbox. Zum Einen kann bei eindeutig bestimmten Punktsituationen eine einfache Zuordnung der Kantenvektoren zu einzelnen Messungen wie im obigen Beispiel direkt erfolgen. Zum Anderen lässt der Wertebereich Aussagen zu, in welcher Raumrichtung die Punktposition beispielsweise am „besten“ im Sinne einer geringsten Impräzision bestimmt werden kann.

**Beispiel 24** Betrachten wir den überbestimmten Rückwärtsschnitt. Vom Neupunkt 4 aus werden Richtungen zu vier Fernzielen beobachtet. Es werden die in Abschnitt 3.1.2 bestimmten Intervallradien verwendet:  $R_{r,41} = 0.38 \text{ mgon}$ ,  $R_{r,42} = 0.47 \text{ mgon}$ ,  $R_{r,43} = R_{r,45} = 0.49 \text{ mgon}$ . Die Standardabweichung der Richtungsbeobachtungen beträgt  $\sigma_R = 0.5 \text{ mgon}$ .

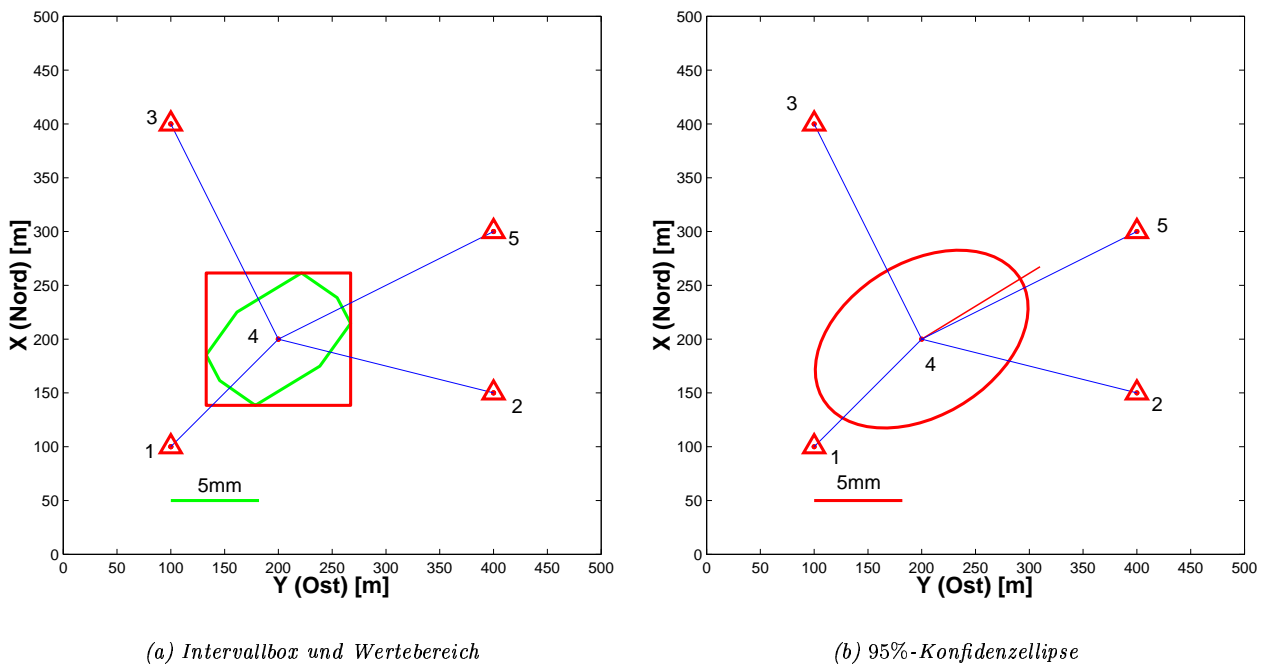


Abbildung 4.7: Punktunsicherheitsbereiche für den überbestimmten Rückwärtsschnitt

Abbildung 4.7(a) stellt die Intervallbox und den Wertebereich dar. Der Wertebereich ist ein Achteck: jede der vier Messungen liefert eine Begrenzung in Form eines Paares paralleler Geraden. Im Gegensatz zur eindeutig

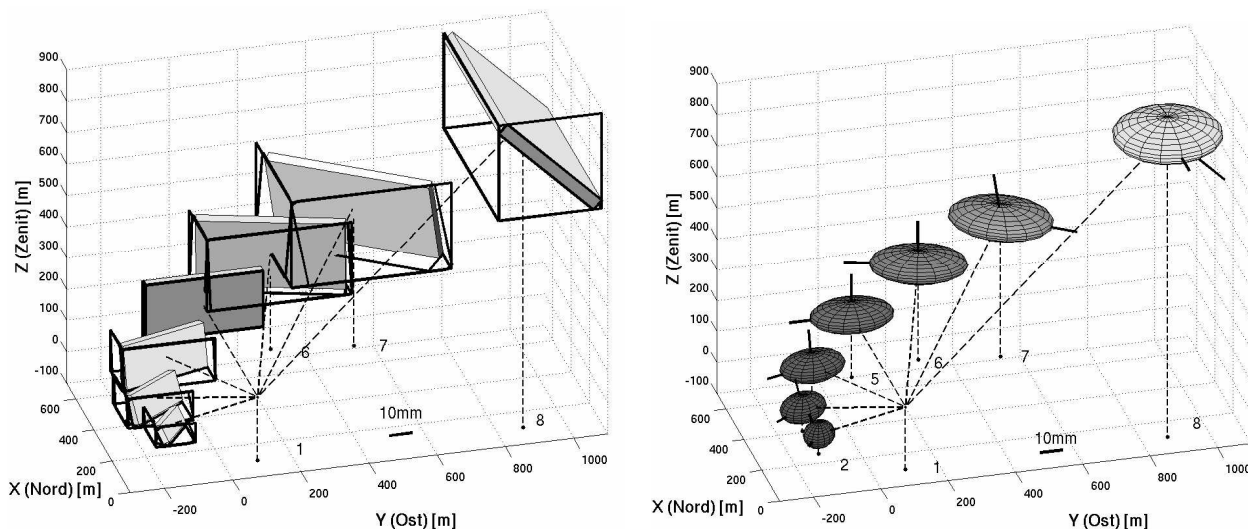
bestimmten Situation im vorherigen Beispiel verlaufen die Messungslinien nicht mehr parallel zu diesen Geradenpaaren, die die Seiten des Wertebereichs bilden. Dies ist Ausdruck der vermittelnden Eigenschaft der Ausgleichung nach der Methode der kleinsten Quadrate im überbestimmten Fall. In Richtung der Verbindungslinie der Punkte 1 und 4 ist der Wertebereich am schlechtesten bestimmt. Diese Information kann aus der fast quadratischen Intervallbox nicht abgeleitet werden. Die Überschätzung des Wertebereichs durch die Intervallbox beträgt in diesem Beispiel  $v = 74\%$ . Abbildung 4.7(b) zeigt die 95%-Konfidenzellipse für den Neupunkt. Die Richtung der großen Halbachse zeigt ebenfalls in Richtung der Verbindung der Punkte 1 und 4.

**Beispiel 25** Bisher hat sich die Überschätzung des Wertebereichs „in Grenzen gehalten“. Das folgende Beispiel soll deshalb dieses Thema kritisch beleuchten. Es sollen mehrere Punkte von einem Standpunkt aus polar angehängt werden, wobei sich die Richtung der Messungslinie in Bezug auf das Koordinatensystem ändert. Die Streckenlänge variiert von  $S = 300\text{ m}$  bis  $S = 1100\text{ m}$ , die Richtungen von  $R = -60\text{ gon}$  bis  $R = 100\text{ gon}$  und die Zenitdistanzen von  $\zeta = 130\text{ gon}$  bis  $\zeta = 50\text{ gon}$ . Die Intervallradien wurden wie in Abschnitt 3.1 beschrieben bestimmt. Sie variieren entsprechend zwischen  $1.8\text{ mm}$  und  $2.6\text{ mm}$  für die Streckenmessungen, zwischen  $0.7\text{ mgon}$  und  $2\text{ mgon}$  für die Richtungsbeobachtungen und zwischen  $0.9\text{ mgon}$  und  $1.7\text{ mgon}$  für die Zenitdistanzbeobachtungen. Die Standardabweichung der Streckenmessung wurde zu  $\sigma_S = \sqrt{4 + 4 \cdot 10^{-6} S^2}\text{ mm}$  angesetzt. Für die Richtungs- und Zenitdistanzbeobachtung wurde  $\sigma_r = \sigma_\zeta = 0.5\text{ mgon}$  angenommen.

Pkt.-Nr.	$x_r$ [mm]	$y_r$ [mm]	$z_r$ [mm]	$Vol(\mathcal{W})$ [cm] <sup>3</sup>	$Vol([\mathbf{x}])$ [cm] <sup>3</sup>	$v$ [%]
2	10.2	8.7	4.6	0.6	3.3	427.8
3	12.0	14.4	7.0	1.5	9.8	517.4
4	9.7	20.5	10.0	3.3	16.1	385.0
5	2.1	26.1	13.3	6.0	6.0	0.0
6	29.1	35.4	20.7	15.0	171.8	1043.3
7	14.6	32.0	17.3	10.0	64.9	547.7
8	45.2	22.8	22.8	28.9	189.3	553.7

Tabelle 4.1: Kenngrößen der Intervallbox und des Wertebereichs der bestimmten Punktkoordinaten

In der ersten Spalte sind die Punktnummern, in der zweiten bis vierten Spalte die Intervallradien der einzelnen Koordinaten angegeben. Die fünfte Spalte listet das Volumen der Wertebereiche auf und erlaubt so einen schnellen Vergleich mit dem Volumen der Intervallbox, das in der nächsten Spalte angegeben wird. Die siebte Spalte zeigt die daraus resultierende Überschätzung des Wertebereichs durch die Intervallbox: Der Wertebereich wird hier bis zum 10-fachen überschätzt (Punkt 6).



(a) Intervallboxen und Wertebereiche

(b) 95%-Konfidenzellipsoide

Abbildung 4.8: Punkunsicherheitsbereiche für das dreidimensionale polare Anhängen

Abbildung 4.8(a) zeigt die Intervallboxen und Wertebereiche für alle Punkte. Die Größe des Wertebereichs variiert nach (4.78) - (4.82). Die Ausrichtung in Bezug auf die Messungslinie ist für alle Punkte identisch, da diese nicht von einer Variation der Messungslinie abhängt. Die Intervallboxen sind jedoch an die Orientierung des Koordinatenrahmens gebunden. Bei ungünstiger Lage der Messungslinie überschätzt die Intervallauswertung den Wertebereich erheblich, was in Tabelle 4.1 dargestellt wird. Abbildung 4.8(b) zeigt die 95%-Konfidenzellipse. Diese sind als quadratische Formen unabhängig von der Rotation des Koordinatenrahmens.

### 4.3.3 Kenngrößen des Wertebereichs

Dieser Abschnitt soll verschiedene Kenngrößen zusammenstellen, die die Zonotope beschreiben und somit als Maßzahlen für die Impräzision der Punktposition herangezogen werden können.

#### Volumen von Zonotopen

Eine erste Kenngröße ist das Volumen, das sich nach der Formel von McMullan (HENK ET AL. 1997, S.264) berechnet:

$$\begin{aligned} Vol(\mathcal{W}_{\mathbf{x}_P}) &= 4 \sum_{1 \leq i_1 < i_2 \leq n} |\det(\mathbf{k}_{P,i_1}, \mathbf{k}_{P,i_2})|, & \text{im Zweidimensionalen,} \\ &= 8 \sum_{1 \leq i_1 < i_2 < i_3 \leq n} |\det(\mathbf{k}_{P,i_1}, \mathbf{k}_{P,i_2}, \mathbf{k}_{P,i_3})|, & \text{im Dreidimensionalen.} \end{aligned} \quad (4.83)$$

#### Breite von Zonotopen

Unter der Breite konvexer Körper versteht man den Abstand zweier paralleler Hyperebenen (Stützebenen), vgl. BONNESEN UND FENCHEL (1934, S.51f) bzw. HADWIGER (1957, S.137f). Die Stützebenen lassen sich nach der Art der Berührung mit dem Körper einteilen in beispielsweise Eckstützebenen oder Kantenstützebenen, vgl. HADWIGER (1957, S.2f) und auch SCHMIDBAUER (1989, S.16f). Die Breite nimmt einen Minimal- und Maximalwert an, der als *Dicke* bzw. *Durchmesser* bezeichnet wird (HADWIGER 1957, S.138).

In der vorliegenden Arbeit soll eine verfeinerte Unterteilung der Breite vorgenommen werden. Hierfür wird der *Abstand paralleler Randflächen des Zonotops*  $d_i$  im Dreidimensionalen bzw. paralleler Kanten im Zweidimensionalen eingeführt. Als zweites Maß wird der *Abstand diametral gegenüber liegender Punkte*  $w_i$  angesetzt.

Mit der H-Darstellung der Zonotope aus (4.50) lassen sich die Abstände paralleler Kanten bzw. Randflächen für Zonotope angeben:

$$\mathbf{h}_i^T (\mathbf{z} - \mathbf{z}_0) = b_i, \quad (4.84)$$

wobei

$$d_i := \frac{2 b_i}{\|\mathbf{h}_i\|}. \quad (4.85)$$

Als Kenngrößen des Zonotops eignet sich der maximale und minimale Abstand:

$$d_{max} := \max\{d_i\}, \quad d_{min} := \min\{d_i\}. \quad (4.86)$$

Die Abstände  $d_i$  werden hauptsächlich durch die Impräzision der Messungen bestimmt.

Der Abstand gegenüberliegender Ecken kann als Norm des Vektorzuges der Kantenvektoren berechnet werden, der diese beiden Eckpunkte miteinander verbindet.

$$\begin{aligned} w_i &= \left\| \sum_{j=i}^{n+i-1} \mathbf{k}_j \right\|, \quad \mathbf{k}_{l+n} = -\mathbf{k}_l, \\ &= \sqrt{\mathbf{\Gamma}^T (\mathbf{K}, -\mathbf{K})^T (\mathbf{K}, -\mathbf{K}) \mathbf{\Gamma}}, \quad \text{mit } \mathbf{\Gamma} = (0, \dots, 0, 1, \dots, 1, 0, \dots, 0)^T, \end{aligned} \quad (4.87)$$

wobei der  $2n \times 1$  Selektionsvektor  $\mathbf{\Gamma}$  genau  $n$  Spalten der erweiterten Matrix  $(\mathbf{K}, -\mathbf{K})$  auswählt. Als Maße eignen sich wieder die minimalen bzw. maximalen Abstände

$$w_{min} := \min\{w_i\}, \quad w_{max} := \max\{w_i\}. \quad (4.88)$$

Für  $w_{max} = w_{min}$  haben alle Ecken den gleichen Abstand vom Zentralpunkt des Zonotops. Allgemein gilt: Die Dicke des Zonotop ist identisch mit  $d_{min}$  und der Durchmesser des Zonotops ist identisch mit  $w_{max}$ . Weitere Klassifizierungen von Zonotopen, die sich mit der Breite ableiten lassen, sind beispielsweise in LEICHTWEISS (1959) oder LINHART (1986,1988) angegeben.

### Singulärwerte der Matrix $\mathbf{K}$ der Kantenvektoren

Analog zu den spektralen Kenngrößen für die Varianz-Kovarianz-Matrix, die in Abschnitt 4.1.3 vorgestellt wurden, soll auf die Singulärwerte der Matrix  $\mathbf{K}$  eingegangen werden. Die Singulärwertzerlegung berechnet sich zu:

$$\mathbf{K} = \mathbf{U} \mathbf{S} \mathbf{V}^T, \quad (4.89)$$

mit  $\mathbf{U}^T \mathbf{U} = \mathbf{I}_u$  der  $u \times u$  Matrix der orthonormierten Linkseigenvektoren von  $\mathbf{K}$ ,  $\mathbf{V}^T \mathbf{V} = \mathbf{I}_n$  der  $n \times n$  Matrix der orthonormierten Rechtseigenvektoren von  $\mathbf{K}$  und  $\mathbf{S} = (\mathbf{S}_1, \mathbf{0})$  der  $u \times n$  Matrix der Singulärwerte. Die  $u \times u$  Diagonalmatrix  $\mathbf{S}_1$  enthält als Diagonalelemente die Singulärwerte. Folgende Eigenschaften der Singulärwertzerlegung sind für die weitere Betrachtung wichtig:

1. Die Singulärwertzerlegung ist eindeutig.
2. Die Singulärwerte sind die Wurzeln der Eigenwerte der Matrix  $\mathbf{K}\mathbf{K}^T$  bzw.  $\mathbf{K}^T\mathbf{K}$ . Sie seien der Größe nach geordnet, so dass gilt:  $s_1 \geq s_2 \geq \dots \geq s_u \geq 0$ .
3. Die Matrix  $\mathbf{U}$  ist die Matrix der orthonormierten Eigenvektoren von  $\mathbf{K}\mathbf{K}^T = \mathbf{U}\mathbf{S}_1^2\mathbf{U}^T$ ,  $\mathbf{V}$  ist die Matrix der Eigenvektoren von  $\mathbf{K}^T\mathbf{K} = \mathbf{V} \begin{pmatrix} \mathbf{S}_1^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T$ .
4. Die der euklidischen Vektornorm zugeordnete Matrixnorm ist durch den größten Singulärwert gegeben:

$$\|\mathbf{K}\|_2 := s_1. \quad (4.90)$$

5. Die Frobeniusnorm der Matrix  $\mathbf{K}$  entspricht der Summe der Quadrate der Singulärwerte

$$\|\mathbf{K}\|_F := \sqrt{\sum_{i=1}^u \sum_{j=1}^n k_{ij}^2} = \sqrt{\sum_{i=1}^u s_i^2} = \sqrt{\text{spur}(\mathbf{K}^T\mathbf{K})}. \quad (4.91)$$

6. Die Singulärwerte geben die Hauptachsllängen und die Linkseigenvektoren die Hauptachsrichtungen einer Ellipse bzw. eines Ellipsoids an, das das Bild der Einheitskugel unter der Projektion  $\pi$  aus (4.52) ist:

$$\mathcal{E} = \pi(\mathcal{S}_n), \quad \mathcal{S}_n := \{ \mathbf{w} \in \mathbb{R}^n \mid \mathbf{w}^T \mathbf{w} = 1 \}. \quad (4.92)$$

### Umfang von 2-Zonotopen

Der Umfangs des 2-Zonotops lässt sich direkt aus der Kantenmatrix ableiten:

$$\text{Peri} := \text{spur}(\sqrt{\mathbf{K}^T\mathbf{K}}). \quad (4.93)$$

### Oberfläche von 3-Zonotopen

Alle  $(n-k)(n-k-1)$  Randflächen des 3-Zonotops lassen sich in Parallelogramme zerlegen bzw. sind Parallelogramme, falls  $k=0$ , d.h. falls keine kollinearen bzw. komplanaren Kantenvektoren auftreten. Die Parallelogramme setzen sich jeweils aus zwei unterschiedlichen Kantenvektoren zusammen (SHEPHARD 1974):

$$O = 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \mathbf{k}_i \times \mathbf{k}_j. \quad (4.94)$$



### 4.3.4 Geometrische Deutung

Im Folgenden soll die Projektion, die den Hyperwürfel auf das Zonotop abbildet, geometrisch gedeutet werden. Die Projektion  $\pi : \mathbb{R}^n \mapsto \mathbb{R}^2(\mathbb{R}^3)$  (4.52) setzt sich insgesamt aus vier Einzelabbildungsschritten zusammen. Diese lassen sich mit der Singulärwertzerlegung der Matrix der Kantenvektoren erklären.

In einem ersten Schritt wird der Hyperwürfel  $\mathcal{C}_n$  im  $\mathbb{R}^n$  gedreht. Diese Drehung wird mit der orthogonalen Matrix  $\mathbf{V}^T$  beschrieben. Danach wird der gedrehte Würfel orthogonal in den Unterraum  $\mathbb{R}^u$  projiziert. COXETER (1973, S.240ff) unterscheidet hierbei drei verschiedene Positionen, die der Hyperwürfel vor der Orthogonalprojektion hat und die sein Bild beeinflussen. Für die Klassifikation stelle man sich vor, dass der Hyperwürfel durch die Projektionsebene hindurch gleite. Hierbei kann eine Seite des Hyperwürfels die Projektionsebene zuerst berühren (*face first*) oder eine Kante (*edge first*) oder ein Eckpunkt (*vertice first*). Die Schnitte, die beim Durchgleiten des Hyperwürfels durch die Projektionsebene entstehen, eignen sich zur Darstellung höherer Dimensionen. Je nach Position des Hyperwürfels variiert die Anzahl seiner  $2^n$  Ecken, die ins Innere des Zonotops abgebildet werden. Bei der „face-first-Position“ bleibt eine minimale Anzahl an Ecken, bei der „vertice-first-Position“ eine maximale Anzahl an Ecken erhalten. Für  $u = 3$  sind dies maximal  $n(n - 1) + 2$  bzw. minimal 8 Ecken des Würfels. Letzteres wird erreicht, wenn  $\mathbf{V} = \mathbf{I}$  gilt.

Im  $\mathbb{R}^u$  wird das Zonotop skaliert mit der Matrix  $\mathbf{S}_1$ . Das Ergebnis wird abschließend erneut gedreht, wobei die Matrix  $\mathbf{U}$  die verwendete Rotationsmatrix angibt. So wird das Zonotop  $\mathcal{Z}_P$  erhalten, das die Impräzision der Punktposition  $P$  beschreibt. Formal lässt sich die Abbildung als Verknüpfung der vier Einzelabbildungen beispielsweise für  $u = 2$  darstellen:

$$\begin{aligned} \pi : \quad & \mathbb{R}^n \rightarrow \mathbb{R}^2 \\ \mathcal{Z}_P = & \mathbf{U} \mathbf{S}_1 \mathbf{\Gamma}_1 \mathbf{V}^T \mathcal{C}_n. \end{aligned} \quad (4.95)$$

### 4.3.5 Abschließendes Beispiel: Kenngrößen für ein Überwachungsnetz

Als abschließendes Beispiel soll das Überwachungsnetz der Linachtalsperre im Schwarzwald herangezogen werden (KAHMEN 1993, S.692). Das Überwachungsnetz besteht aus fünf Kontrollpfeilern (Punkte 1-5) und Objektpunkten an der Staumauer, die hier durch drei Punkte exemplarisch repräsentiert werden (Punkte 6-8). Die Intervallradien der Messungen werden nach Abschnitt 3.1 bestimmt, die Standardabweichungen der Messungen ergeben sich zu  $\sigma_R = \sigma_\zeta = 0.5 \text{ mgon}$  bzw.  $\sigma_S = \sqrt{4 + 4 \cdot 10^{-6} S^2} \text{ mm}$ .

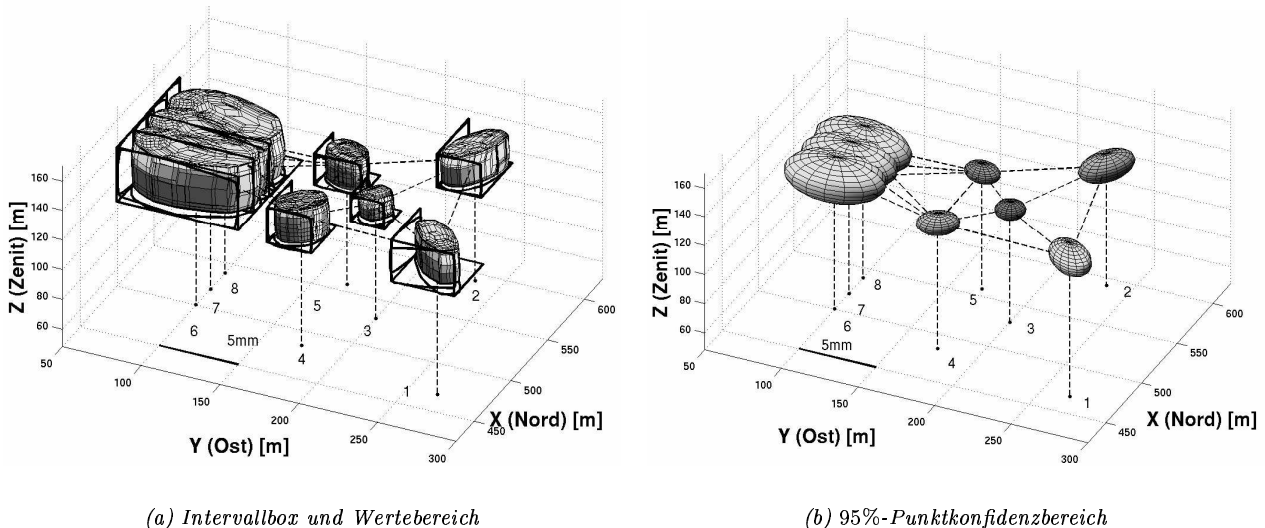


Abbildung 4.9: Punktunsicherheitsbereiche für ein Überwachungsnetz

Abbildung 4.9 zeigt die Punktunsicherheitsbereiche, die unabhängig von der Orientierung des Koordinatensystems sind, das bei Anwendungen in der Ingenieurgeodäsie frei gewählt werden kann. Die Ausgleichung wurde als freies Netz mit Teilspurminimierung (vgl. Abschnitt 4.1.3) über die Kontrollpunkte berechnet, was ein typischer Ansatz in der Deformationsanalyse ist, vgl. WELSCH ET AL. (2000, S.202). Da jede der 14 Strecken-

bzw. Zenitdistanzbeobachtungen und jede der 18 Richtungen zur Form des Wertebereichs beiträgt, hat jedes Zonotop bis zu 2070 Randflächen, vgl. (4.60).

Teilabbildung 4.9(a) zeigt die Wertebereiche der Intervallauswertung und die Intervallboxen, die durch drei Randflächen angedeutet sind. In diesen Randflächen sind die entsprechenden Parallelprojektionen des Zonotops eingezeichnet. Da die horizontale Netzausdehnung in diesem Beispiel größer ist als die vertikale Netzausdehnung, sind die Zonotope eher zylindrisch als kugelförmig. Teilabbildung 4.9(b) zeigt die 95%-Konfidenzellipse für das dreidimensionale Netz. Es fällt auf, dass Zonotope und Konfidenzellipsen ähnlich ausgerichtet sind.

Im Folgenden sollen die Kenngrößen für Zonotope, die im vorherigen Abschnitt eingeführt wurden, diskutiert werden, um festzustellen, inwieweit sie sich für eine Charakterisierung der Impräzision der Punktposition eignen.

Pkt.-Nr.	$x_r$ [mm]	$y_r$ [mm]	$z_r$ [mm]	$Vol(\mathcal{W})$ [mm] <sup>3</sup>	$Vol([\mathbf{x}])$ [mm] <sup>3</sup>	$v$ [%]	$O$ [mm] <sup>2</sup>
1	1.99	1.98	1.39	17.81	44.18	148.0	9.91
2	2.15	1.38	1.12	17.12	26.93	57.3	9.41
3	1.17	1.05	0.87	5.26	8.66	64.5	3.87
4	1.70	1.40	1.17	14.62	22.51	53.9	8.04
5	1.31	1.48	1.05	11.18	16.49	47.4	6.69
6	2.49	3.82	1.83	90.38	139.69	54.5	28.19
7	2.28	3.80	1.82	84.49	126.84	50.1	26.85
8	2.25	3.87	1.79	84.51	125.19	48.1	26.85

Tabelle 4.2: Vergleich der Intervallradien der Koordinaten und der Volumina und Oberflächen der Zonotope

Tabelle 4.2 gibt eine Übersicht über die Intervallradien der geschätzten Punktkoordinaten (Spalte 2-4). In der fünften Spalte ist das Volumen des Wertebereichs angegeben, das nach (4.83) berechnet wurde. Als Vergleich dazu sind das Volumen der Intervallbox in der sechsten und die Überschätzung nach (4.76) in der siebten Spalte aufgelistet. Mit Ausnahme des Punktes 1 liegen die Wertebereiche günstig bezüglich des Koordinatensystem, so dass die Überschätzung durchschnittlich ca. 50% beträgt. In der letzten Spalte ist die Oberfläche des Zonotops nach (4.94) angegeben. Volumina und Oberfläche eignen sich, die Impräzision der Punktposition zu beurteilen. Je größer das Volumen ist, desto größer ist die Impräzision der Punktposition. Allerdings weisen diese Größen integralen Charakter auf. Daher werden keine Aussagen über die Form des Wertebereichs geliefert, beispielsweise können keine Fragen nach Richtungen, in denen die Punktlage am besten oder am schlechtesten bestimmt ist, beantwortet werden.

Hierfür eignen sich die Abstände paralleler Randflächen  $d_i$  (vgl. 4.85) bzw. Abstände gegenüberliegender Ecken  $w_i$  (vgl. 4.87). Je kleiner die Differenz  $w_{max} - w_{min}$  desto kugelförmiger ist das Zonotop. Aufgrund der großen Anzahl der Randflächen besteht in diesem Beispiel kein großer Unterschied zwischen den Größen  $w_i$  und  $d_i$ , so dass in der folgenden Tabelle nur die Abstände paralleler Randflächen angegeben wird.

Pkt.-Nr.	$d_{max}$ [mm]	$Az_{max}$ [gon]	$El_{max}$ [gon]	$d_{min}$ [mm]	$Az_{min}$ [gon]	$El_{min}$ [gon]
1	2.7	146.6	29.6	1.0	-142.6	-2.1
2	2.3	35.4	23.8	1.0	-65.0	94.2
3	1.2	11.9	-32.7	0.8	-1.0	97.5
4	2.0	-162.8	38.5	1.2	-37.3	93.5
5	1.7	137.5	31.2	1.0	-47.8	95.1
6	3.9	-113.4	34.8	1.6	-139.5	-94.2
7	3.8	-121.2	38.3	1.6	76.5	92.7
8	3.7	-93.7	10.6	1.6	91.5	94.0

Tabelle 4.3: Maximaler und minimaler Abstand paralleler Randflächen der Zonotope

Der Abstand gegenüberliegender paralleler Randflächen liefert lokale Information. In den dritten und vierten bzw. sechsten und letzten Spalte sind die Richtungen (Azimut und Elevation) angegeben, in denen die Punktlage am schlechtesten bzw. am besten bestimmt ist. Für integrale Aussagen zur Beurteilung der Impräzision der Punktposition ist dieses Maß weniger geeignet als die oben vorgestellten Volumina bzw. Oberflächen.

Ergänzend sollen noch spektrale Kenngrößen angegeben werden, die mit der Singulärwertzerlegung aus der Matrix  $\mathbf{K}$  abgeleitet werden (4.89). Diese werden im nachfolgenden Abschnitt geometrisch gedeutet. Die zweite

bis sechste Spalte liefert die Singulärwerte und die Hauptrichtung des größten Singulärwertes der Matrix  $\mathbf{K}_P$ , die für jeden Netzpunkt nach (4.75) erhalten wird. In den letzten drei Spalten sind verschiedene Matrixnormen der Matrix  $\mathbf{K}$  angegeben. Die der euklidischen Vektornorm zugeordnete Matrixnorm  $\|\mathbf{K}\|_2$  ist mit dem größten Singulärwert identisch, vgl. (4.90). Die Zeilennorm  $\|\mathbf{K}\|_\infty$  liefert den größten Intervallradius der drei Punktkoordinaten, vgl. Tabelle 4.2, 2.-4. Spalte. Die Frobeniusnorm ist proportional zum quadratischen Mittel der Singulärwerte.

Pkt.-Nr.	$s_1$ [mm]	$s_2$ [mm]	$s_3$ [mm]	$Az_{max}$ [gon]	$El_{max}$ [gon]	$\ \mathbf{K}\ _2$ [mm]	$\ \mathbf{K}\ _\infty$ [mm]	$\ \mathbf{K}\ _F$ [mm]
1	0.60	0.52	0.26	146.8	17.7	0.60	1.99	0.84
2	0.53	0.42	0.29	-168.8	-4.1	0.53	2.15	0.74
3	0.27	0.25	0.24	192.8	86.5	0.27	1.17	0.44
4	0.36	0.32	0.27	57.5	-9.3	0.36	1.70	0.56
5	0.34	0.29	0.23	-76.4	-7.7	0.34	1.48	0.51
6	0.86	0.67	0.53	-108.8	7.4	0.86	3.82	1.21
7	0.83	0.66	0.52	-102.9	11.6	0.83	3.80	1.18
8	0.85	0.65	0.51	-96.0	10.2	0.85	3.87	1.19

Tabelle 4.4: Spektrale Kenngrößen der Zonotope und Matrix-Normen der Matrix  $\mathbf{K}$

Zusammenfassend lässt sich festhalten, dass mit Hilfe der Wertebereiche detailliertere Aussagen über die Impräzision der Punktlage abgeleitet werden können als mit der Intervallbox. Neben integralen Kenngrößen wie Volumina oder Flächeninhalt der Zonotopoberflächen lassen sich minimale Abstände paralleler Randflächen bzw. gegenüberliegender Ecken berechnen und damit Aussagen über Richtungen ableiten, in denen die Impräzision der Punktlage am größten bzw. am kleinsten ist. Spektrale Kenngrößen liefern sowohl integrale Informationen (Matrixnormen) als auch Detailinformationen (einzelne Singulärwerte), deren geometrische Bedeutung im folgenden Abschnitt erklärt wird.

### 4.3.6 Ersatzformen für den Wertebereich

#### Einführung

Aus dem umfangreichen Beispiel im vorherigen Abschnitt 4.3.5 wurde deutlich, dass die Form des Zonotops mit steigender Anzahl der Beobachtungen  $n$  immer komplexer wird. Für  $n = 46$  hat das 3-Zonotop beispielsweise 2070 Randflächen. Es ist deshalb besonders für eine übersichtliche Darstellung von Bedeutung, nach einfachen Ersatzformen zu suchen, die die Impräzision beschreiben. Alle Qualitätsmaße, die in Abschnitt 4.3.2 vorgestellt wurden, können weiter aus der Matrix der Kantenvektoren abgeleitet werden. Die Ersatzform muss den Wertebereich, das Zonotop, gut approximieren und von der Orientierung des Koordinatensystems unabhängig sein. Hierzu bieten sich beispielsweise Ellipsoide an, die als quadratische Formen diese Anforderungen erfüllen. Ellipsoide werden in anderen Fachdisziplinen (Verfahrenstechnik, Elektrotechnik oder Wirtschaftswissenschaften) zur Beschreibung und Vorhersage von dynamischen Prozessen genutzt. Eine mengentheoretische Beschreibung und Einschließung der Lösung wird mit den Begriffen *Ellipsoidal Bounding* oder *Ellipsoidal Calculus* belegt, vergleiche zur Übersicht MILANESE ET AL. (1996) oder KURZHANSKI UND VÁLYI (1997). Ein Vorteil von Ellipsoiden als Ersatzformen ist beispielsweise die Abgeschlossenheit bezüglich linearer Abbildungen. Innerhalb der Klasse der Ellipsoide können verschiedene Anforderungen an diese gestellt werden, um eine eindeutige Lösung zu erhalten. Beispielsweise kann ein umschließendes Ellipsoid mit minimaler Spur oder minimalem Volumen gefordert werden, vgl. KURZHANSKI UND VÁLYI (1997, S.132-142) oder GRÖTSCHEL ET AL. (1988, Kap.3).

#### Projektion der Umkugel

Eine einfache und direkte Möglichkeit, ein Ellipsoid anzugeben, das das Zonotop enthält, ist die Projektion der Umkugel  $\mathcal{S}_{\mathcal{C}_n}$  des Hyperwürfels  $\mathcal{C}_n$ . Für die Projektion ins Zweidimensionale nach (4.52) ergibt sich beispielsweise:

$$\begin{aligned}
 \mathcal{E}_P &= \pi(\mathcal{S}_{\mathcal{C}_n}), \text{ mit } \mathcal{S}_{\mathcal{C}_n} := \{ \mathbf{w} \in \mathbb{R}^n \mid \mathbf{w}^T \mathbf{w} = n \} \\
 &= \{ \mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x}^T (n \mathbf{K} \mathbf{K}^T)^{-1} \mathbf{x} \leq 1 \}.
 \end{aligned} \tag{4.96}$$

Da die Umkugel und der Hyperquader durch dieselbe lineare Abbildungsvorschrift auf das Ellipsoid bzw. das Zonotop abgebildet werden, stellt das Ellipsoid eine Einschließung des Zonotops dar.

Abbildung 4.10(a) zeigt die umschließende Ellipse und den Wertebereich für die Konfiguration des Vorwärtsschnittes aus Beispiel 23. Hier liegen alle vier Ecken des Zonotops auf der umschließenden Ellipse. Abbildung 4.10(b) stellt den Wertebereich und die ihn umschließende Ellipse für die Konfiguration des überbestimmten Rückwärtsschnittes aus Beispiel 24 dar. Je nach Lage des Hyperwürfels vor der Orthogonalprojektion, berührt die Ellipse das Zonotop in unterschiedlich vielen Eckpunkten. Ist der Hyperwürfel in der „face-first-Position“ so liegt keine Ecke des Zonotops auf der umschließenden Ellipse.

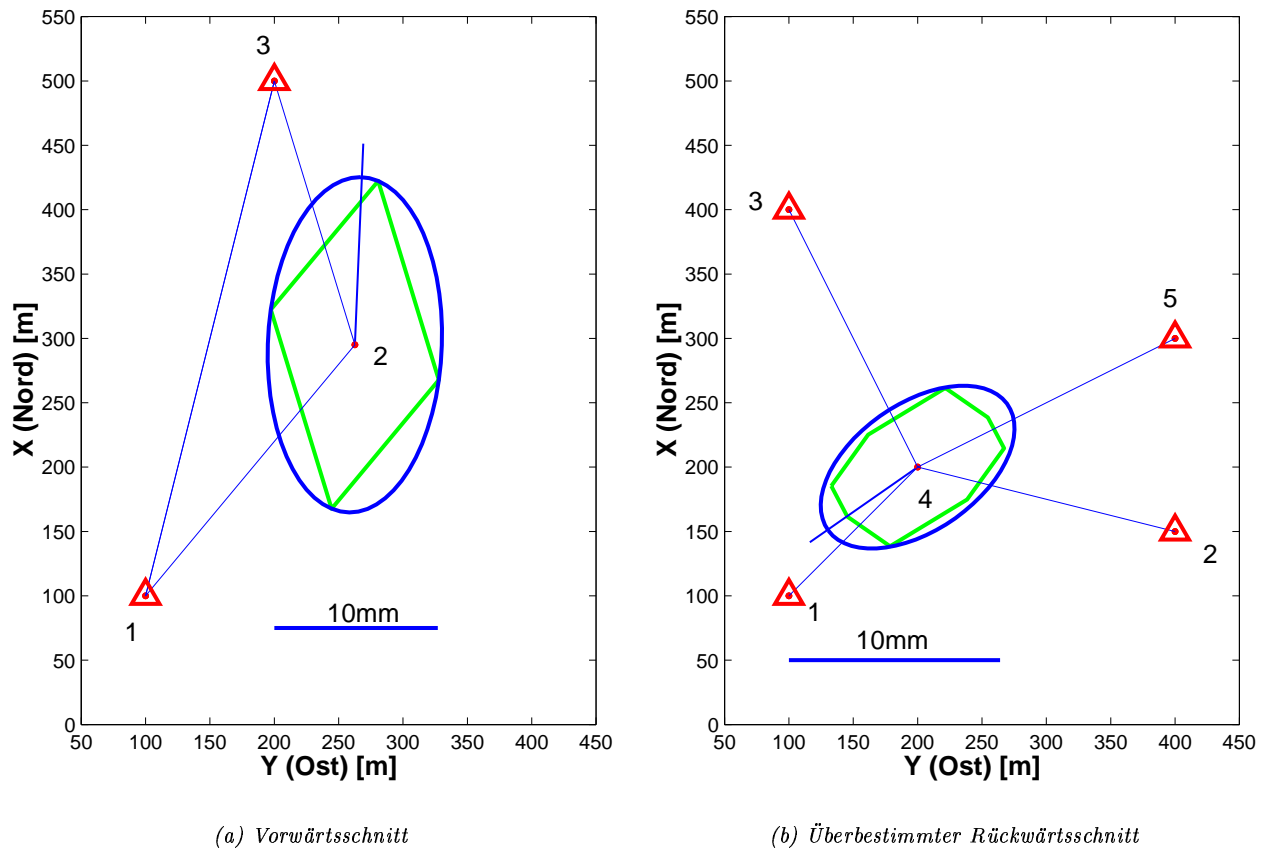


Abbildung 4.10: Einschließung des Zonotops durch umschließende Ellipsen

### Löwner-Ellipsoid

Um eine bessere Einschließung zu garantieren, kann minimales Volumen für das Ellipsoid gefordert werden, das das Zonotop einschließt. Dieses Ellipsoid, das für konvexe Mengen diese Bedingung erfüllt, wird als *Löwner-Ellipsoid* bezeichnet (JUHNE 1990):

$$\mathcal{E}_L := \left\{ \begin{array}{l} \min_{\mathbf{x}} (\det \mathbf{B})^{-1} \\ \mathbf{x}^T \mathbf{B} \mathbf{x} \leq 1, \quad \mathbf{x} \in \mathcal{Z} \subset \mathbb{R}^u \end{array} \right. \quad (4.97)$$

Das Löwner-Ellipsoid zeichnet sich durch folgende Eigenschaften aus:

1. Das Löwner-Ellipsoid ist eindeutig, vgl. DANZER ET AL. (1957) oder JUHNE (1990).
2. Minimales Volumen für umschließende Ellipsoide ist äquivalent zu größter Anzahl an Berührungspunkten (BEHREND 1938).
3. Für zentralsymmetrische Körper fallen der Mittelpunkt des Ellipsoids und das Zentrum des Körpers zusammen.

4. Es existiert ein zu  $\mathcal{E}_L$  konzentrisch und maßstäblich um den Faktor  $\frac{1}{u}$  verkleinertes Ellipsoid, das in der konvexen Hülle des konvexen Körpers liegt. Für zentralsymmetrische Körper ist der Faktor  $\frac{1}{\sqrt{u}}$  ausreichend, vgl. JOHN (1948) oder JUHNKE (1990).

Abbildung 4.11 stellt für das Netz aus Beispiel 22 einen Vergleich der umschließenden Ellipsen und der Löwner-Ellipsen für die Netzpunkte dar. Die Wertebereiche sind jeweils 12-Ecke. In diesem Beispiel verändert sich die Richtung der großen Halbachse zwischen der umschließenden Ellipse und der Löwner-Ellipse um maximal 38 gon. Die Einschließung des Wertebereichs verbessert sich um ca. 30%.

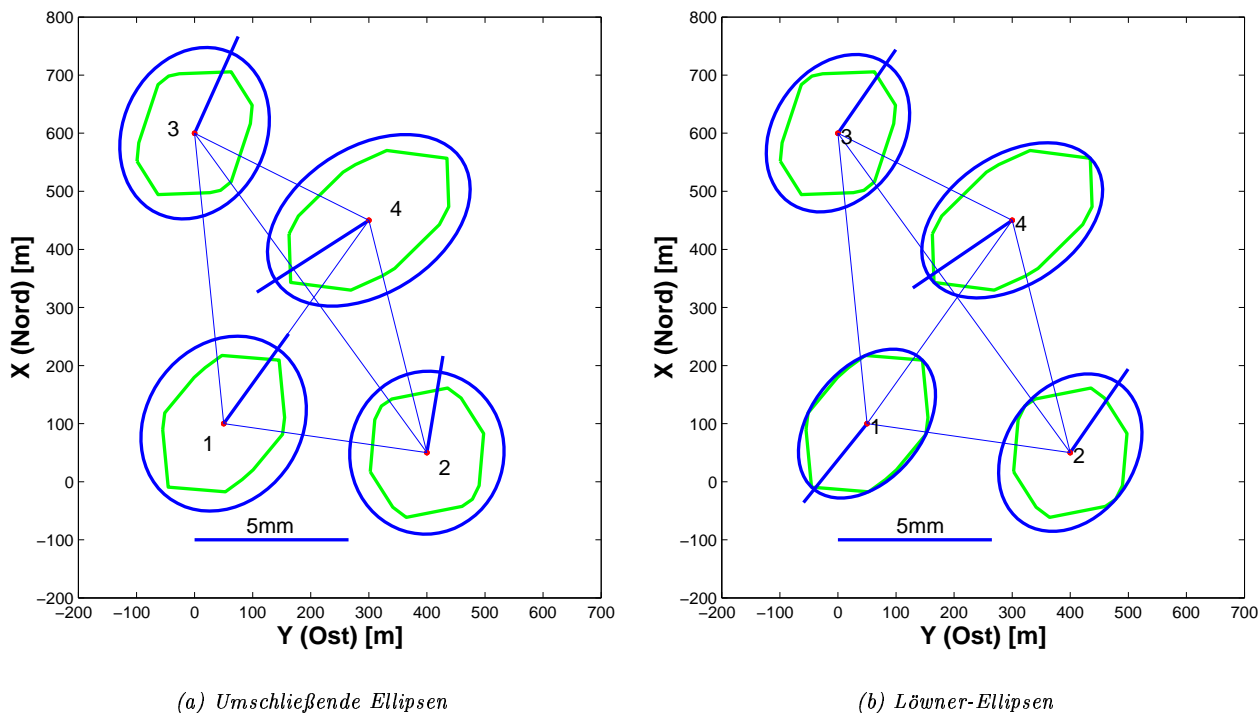


Abbildung 4.11: Vergleich der umschließenden Ellipsen mit den Löwner-Ellipsen

Die Berechnung umschließender Ellipsoide als eine „beste“ Ersatzform für konvexe Mengen ist sehr aufwändig, wenn keine weiteren Informationen vorliegen (GRÖTSCHEL ET AL. 1988, S.70f, S.80ff). Im Gegensatz zu den meisten Anwendungen, bei denen die konvexen Körper durch Punkte diskretisiert sind, ist im Kontext der vorliegenden Arbeit der konvexe Körper (Zonotop) durch seine Kantenvektoren angegeben. Die Koordinaten der Eckpunkte sind a priori unbekannt. Die Berechnung des Löwner-Ellipsoid nach (4.97) basiert aber auf der Kenntnis der Koordinaten der Eckpunkte. Diese können nur aufwändig aus der Darstellung gewonnen werden (EPPSTEIN 1995). Die explizite Berechnung des Löwner-Ellipsoids ist also nur sinnvoll, wenn die Punkte bekannt sind. Aus diesem Grund wird das umschließende Ellipsoid  $\mathcal{E}_P$  zur Approximation des Wertebereichs vorgeschlagen.

#### 4.4 Verknüpfung von Impräzision und Stochastizität

Eine wesentliche Fragestellung ist, wie die Stochastizität und Impräzision des Ergebnisses miteinander verknüpft werden können. Ziel soll hierbei sein, ein erweitertes Maß zu erhalten, das vollständig stochastisch interpretierbar bleibt, für das also beispielsweise Hypothesentests und Konfidenzintervalle weiterhin existieren. Im Folgenden werden zwei Ansätze vorgestellt, mit denen *erweiterte Konfidenzbereiche* und *erweiterte Testgrößen* abgeleitet werden können.

### 4.4.1 Konzept der Superposition

Eine erste einfache Möglichkeit, erweiterte Konfidenzbereiche  $\mathcal{K}_{1-\gamma}^*$  abzuleiten, ist die Superposition von Zonotopen als Maß der Impräzision mit Konfidenzellipsen bzw. -ellipsoiden, den Maßen für Stochastizität. Fasst man beide Bereiche als Mengen auf, so kann die Superposition als Minkowski-Summe geschrieben werden, vgl. Definition 14:

$$\mathcal{K}_{1-\gamma}^* := \mathcal{Z} \oplus \mathcal{E}_K = \mathcal{Z} \oplus \mathcal{K}_{1-\gamma}. \quad (4.98)$$

Durch diesen Ansatz wird der Wertebereichsgedanke konsequent fortgesetzt. Der resultierende erweiterte Konfidenzbereich ist dann folgendermaßen zu interpretieren: Unter Berücksichtigung der Impräzision überdeckt mit einer Konfidenzwahrscheinlichkeit von  $1 - \gamma$  der *erweiterte Konfidenzbereich* die geschätzte Punktposition.

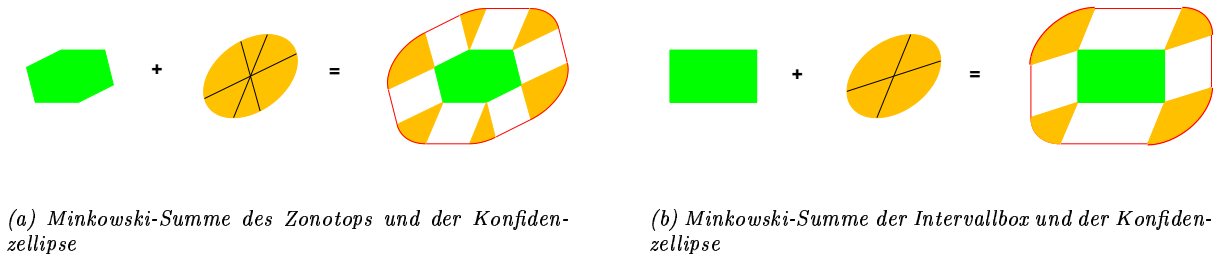


Abbildung 4.12: Erweiterte Konfidenzbereiche

Die Konstruktion lässt sich graphisch motivieren, vgl. Abbildung 4.12(a): Das Zonotop stellt den Bereich aller Punktpositionen dar, die sich unter Berücksichtigung der Impräzision der Messungen ergeben. Für jede dieser Punktpositionen existieren darüber hinaus Informationen über ihre stochastischen Eigenschaften in Form von Punktconfidenzellipsen oder -ellipsoiden. Maximale Bereiche ergeben sich durch Verschieben der Konfidenzellipse auf dem Rand des Zonotops. Die resultierende Form des erweiterten Konfidenzbereichs ist keine Ellipse und kein Zonotop mehr, der Rand ist jedoch eine glatte Kurve.

Die Darstellung als Minkowski-Summe enthält allerdings keine konkrete Anleitung, wie der Rand des erweiterten Konfidenzbereichs funktional beschreibbar ist. Ein mögliches Verfahren soll deshalb kurz skizziert werden. In einem ersten Schritt sind die Punkte  $\mathbf{x}^*$  der Ellipse zu bestimmen, in denen die Tangentenrichtung parallel zur Kantenrichtung des Zonotops ist. An den Eckpunkten des Zonotops werden die entsprechenden Ellipsensektoren der Ellipsenzerlegung angebracht, vgl. Abbildung 4.12. Die Endpunkte der Ellipsensektoren werden durch die passenden Kantenvektoren des Zonotops verbunden.

Um einen erweiterten Konfidenzbereich auch ohne Kenntnis der Ecken des Zonotops berechnen zu können, kann die Minkowski-Summe der Konfidenzellipse bzw. des Konfidenzellipsoids und der Intervallbox gebildet werden, vgl. Abbildung 4.12(b). Mit (4.98) und der Einschließungseigenschaft gilt:

$$\mathcal{Z} \subseteq [\hat{\mathbf{x}}] \Rightarrow \mathcal{K}_{1-\gamma}^* \subseteq [\hat{\mathbf{x}}] \oplus \mathcal{E}_K =: \mathcal{K}_{1-\gamma}^{**}. \quad (4.99)$$

Eine weitere Vereinfachung kann erzielt werden, wenn auf die umschließende Ellipse zurückgegriffen wird. Der somit erhaltene erweiterte Konfidenzbereich  $\mathcal{K}_{1-\gamma}^{***}$  ist wie  $\mathcal{K}_{1-\gamma}^*$  unabhängig von der Orientierung des Koordinatensystems. Der Rand des erweiterten Konfidenzbereichs  $\mathcal{K}_{1-\gamma}^{***}$  lässt sich geschlossen angeben. Für den zweidimensionalen Fall erhält man beispielsweise für eine Darstellung in Polarkoordinaten:

$$r(\tau) = \sqrt{a'^2 - b'^2 \cos^2 \tau}, \quad (4.100)$$

wobei  $a'^2 = a_{\mathcal{K}_{1-\gamma}}^2 + a_{\mathcal{E}_P}^2 + 2a_{\mathcal{K}_{1-\gamma}} a_{\mathcal{E}_P} \cos(\phi_{\mathcal{K}_{1-\gamma}} - \phi_{\mathcal{E}_P})$  und  $b'^2 = b_{\mathcal{K}_{1-\gamma}}^2 + b_{\mathcal{E}_P}^2 + 2b_{\mathcal{K}_{1-\gamma}} b_{\mathcal{E}_P} \cos(\phi_{\mathcal{K}_{1-\gamma}} - \phi_{\mathcal{E}_P})$  gilt. Die Halbachsen der Konfidenzellipse bzw. der umschließenden Ellipse werden mit  $a_{\mathcal{K}_{1-\gamma}}$ ,  $b_{\mathcal{K}_{1-\gamma}}$  bzw.  $a_{\mathcal{E}_P}$ ,  $b_{\mathcal{E}_P}$  bezeichnet, die Richtungswinkel der großen Halbachsen mit  $\phi_{\mathcal{K}_{1-\gamma}}$  bzw.  $\phi_{\mathcal{E}_P}$ .

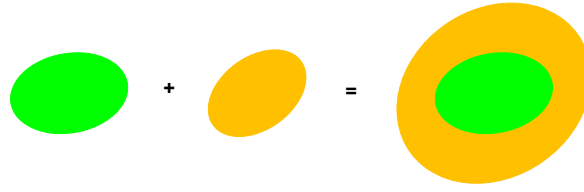


Abbildung 4.13: Minkowski-Summe der umschließenden Ellipse und der Konfidenzellipse

Abbildung 4.13 zeigt die Minkowski-Summe der umschließenden Ellipse und der Konfidenzellipse. Diese Summe ist im Allgemeinen keine Ellipse mehr, sondern nur im Spezialfall, falls beide Richtungen der großen Halbachsen zusammenfallen.

#### 4.4.2 Konzept des Erweiterungsprinzips in der Fuzzy-Theorie

Die eher mengentheoretische Ableitung der erweiterten Kenngröße stellt einen Spezialfall des Erweiterungsprinzips der Fuzzy-Theorie dar und kann in dieses allgemeinere Konzept eingebettet werden. Die Fuzzy-Erweiterung des Kleinsten-Quadrate-Schätzers für geodätische Anwendungen wurde von KUTTERER (1994, 2002b) entwickelt und vorgestellt. Im Folgenden sollen die dort präsentierten Ergebnisse auf Intervallgrößen als Spezialfälle von Fuzzy-Größen angewandt und kurz skizziert werden. Für die Überlegungen ist es wesentlich, Vektoren von Fuzzy-Größen (*unscharfe Vektoren*) zu definieren. Dies wird durch Angabe einer Zugehörigkeitsfunktion  $m : \mathbb{R}^n \rightarrow [0, 1]$  erreicht.

##### Unscharfer Vektor der ausgeglichenen Parameter

Ziel ist es, einen unscharfen Vektor für die geschätzten Parameter anzugeben. In einem ersten Schritt wird hierzu der unscharfe Vektor für die Beobachtungen bestimmt. Mit dessen Hilfe wird in einem zweiten Schritt der unscharfe Vektor der geschätzten Parameter ermittelt. Wir greifen dafür auf  $\alpha$ -Schnitte zurück, die in Definition 11 eingeführt wurden. Die gesuchte Zugehörigkeitsfunktion darf genau einen Maximalwert annehmen (*Gipfelpunkt*) und alle  $\alpha$ -Schnitte müssen einfache und zusammenhängende Teilmengen des  $\mathbb{R}^n$  sein, vgl. KUTTERER (2002b, S.47).

Es lassen sich verschiedene Ansätze zur Ableitung von Zugehörigkeitsfunktionen für Vektoren unterscheiden. Einerseits können diese aus den Zugehörigkeitsfunktionen der einzelnen Vektorkomponenten beispielsweise durch die Minimumregel (BANDEMER UND GOTTWALD 1993, S.59ff) oder die „Kombination der Komponenten“ (VIERTL 1996, S.75ff) erhalten werden. Da die Fuzzy-Größen in dieser Arbeit Intervalle sind, ist die Diskussion für jeden  $\alpha$ -Schnitt möglich und lässt sich mit der Minimumregel direkt übertragen. Die Zugehörigkeitsfunktion  $m_{\tilde{\Delta}\mathbf{l}}(\Delta\mathbf{l}) : \mathbb{R}^n \rightarrow [0, 1]$  lautet für den Vektor der reduzierten Beobachtungen  $\Delta\mathbf{l} = \mathbf{l} - \mathbf{a}_0$ :

$$m_{\tilde{\Delta}\mathbf{l}}(\Delta\mathbf{l}) = \begin{cases} 1, & \Delta\mathbf{l} \in [\Delta\mathbf{l}] \\ 0, & \text{sonst.} \end{cases} \quad (4.101)$$

Bei dieser Verknüpfung entspricht der  $\alpha$ -Schnitt des kartesischen Produktes der Vektorkomponenten dem kartesischen Produkt der  $\alpha$ -Schnitte. Darüber hinaus können direkt Zugehörigkeitsfunktionen angegeben werden, ohne auf die einzelnen Komponenten zurückzugehen. KUTTERER (2002b, S.47) verwendet beispielsweise eine Zugehörigkeitsfunktion vom elliptischen Typ, deren  $\alpha$ -Schnitte alle Ellipsen sind. Mit der positiv definiten  $n \times n$  Matrix  $\mathbf{U}$  und einer streng monoton fallenden Funktion  $k$  lässt sich die Zugehörigkeitsfunktion folgendermaßen angeben:

$$m_{\tilde{\Delta}\mathbf{l}}(\tilde{\Delta}\mathbf{l}) = k\sqrt{(\Delta\mathbf{l} - \Delta\mathbf{l}_m)^T \mathbf{U}^{-1} (\Delta\mathbf{l} - \Delta\mathbf{l}_m)}, \quad k : \mathbb{R}_0^+ \rightarrow [0, 1], \quad k(0) = 1. \quad (4.102)$$

Im zweiten Schritt wird das *Erweiterungsprinzip nach Zadeh* angewendet, um den unscharfen Vektor der geschätzten Parameter zu bestimmen, vgl. BANDEMER UND GOTTWALD (1993, S.46f) oder KUTTERER (2002b, S.67f):

$$m_{\tilde{\hat{\mathbf{x}}}}(\hat{\mathbf{x}}) := \sup_{\tilde{\mathbf{x}} = (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} \Delta\mathbf{l} + \hat{\mathbf{x}}} m_{\tilde{\Delta}\mathbf{l}}(\Delta\mathbf{l}). \quad (4.103)$$

Diese Vorschrift ist folgendermaßen zu lesen: Verschiedene Linearkombinationen der reduzierten Beobachtungen können den gleichen Vektor  $\hat{\mathbf{x}}^*$  liefern. Als Wert der Zugehörigkeitsfunktion von  $\hat{\mathbf{x}}^*$  soll das Maximum der Werte von  $m_{\hat{\Delta}\mathbf{1}}$  für alle diese Linearkombinationen verwendet werden. Mit der Zugehörigkeitsfunktion aus (4.101) ergibt sich dann:

$$m_{\hat{\mathbf{x}}}(\hat{\mathbf{x}}) = \begin{cases} 1, & \hat{\mathbf{x}} + \mathbf{d}\hat{\mathbf{x}} \in \mathcal{W}_{\hat{\mathbf{x}}} \\ 0, & \text{sonst.} \end{cases} \quad (4.104)$$

Da die Zugehörigkeitsfunktion die Indikatorfunktion (vgl. Definition 9) ist, haben alle  $\alpha$ -Schnitt die identische Form: die Form des Wertebereichs  $\mathcal{W}_{\hat{\mathbf{x}}}$ . Zugehörigkeitsfunktionen vom Typ „Intervallbox“ (4.101) sind somit nicht abgeschlossen bezüglich linearer Abbildungen, (KUTTERER 2002b, S.48). Um die Abgeschlossenheit zu garantieren, können entweder Zugehörigkeitsfunktionen vom elliptischen Typ für die Beobachtungen verwendet (KUTTERER 2002b, S.47f) oder (4.104) über der Intervallbox  $[\hat{\mathbf{x}}]$  anstelle des Wertebereichs  $\mathcal{W}_{\hat{\mathbf{x}}}$  definiert werden.

### Konfidenzbereiche für unscharfe Vektoren

Der Konfidenzbereich für die geschätzten Parameter ist eine Funktion derselben. Daher ist mit Hilfe des Erweiterungsprinzip eine Zugehörigkeitsfunktion für den Konfidenzbereich zu bestimmen (VIERTL 1996). KUTTERER (2002b, S.67ff) diskutiert dafür verschiedene Möglichkeiten und Verfahren und führt die Berechnung auf eine Optimierungsaufgabe zurück. Da in der vorliegenden Arbeit mit sehr einfachen Zugehörigkeitsfunktionen, nämlich Indikatorfunktionen, vgl. Definition 9, gearbeitet wird, kann die Optimierungsaufgabe analytisch gelöst werden. Die Zugehörigkeitsfunktion  $m_{\hat{\mathbf{x}}}(\hat{\mathbf{x}})$  ist konstant und somit für alle Parameter identisch, so dass gilt:

$$m_{\mathcal{K}_{1-\gamma}^*} = \begin{cases} 1, & \mathbf{x} \in \mathcal{K}_{1-\gamma}^* \\ 0, & \text{sonst,} \end{cases} \quad (4.105)$$

wobei  $\mathcal{K}_{1-\gamma}^*$  den erweiterten Konfidenzbereich darstellt. Es besteht das Problem, den Konfidenzbereich einfach zu parametrisieren, worauf schon in Abschnitt 4.4.1 eingegangen wurde. KUTTERER (2002b, S.70) hat dieses Problem elegant umgangen, indem er Zugehörigkeitsfunktionen vom elliptischen Typ verwendet.

Wenn es zu aufwändig ist, den Rand des Wertebereichs bzw. der erweiterten Konfidenzregion  $\mathcal{K}_{1-\gamma}^*$  zu berechnen, wird vorgeschlagen, alternativ den erweiterten Konfidenzbereich  $\mathcal{K}_{1-\gamma}^{***}$  heranzuziehen. Es gilt dann beispielsweise für den zweidimensionalen Fall:

$$\mathcal{K}_{1-\gamma}^{***} = \{ \hat{\mathbf{x}}_P \in \mathbb{R}^2 \mid \hat{\mathbf{x}}_P^T \hat{\mathbf{x}}_P \leq r^2(\tau) \}, \quad (4.106)$$

mit  $r(\tau)$  aus (4.100). Eine graphische Darstellung des erweiterten Konfidenzbereichs für Intervalle als Fuzzy-Größen liefert Abbildung 4.13. Für elliptische Zugehörigkeitsfunktionen werden entsprechende Bereiche bei KUTTERER (2002a) gezeigt.

### Statistische Tests mit impräzisen Daten

Statistische Tests für impräzise Daten wurden von KUTTERER (2002b) in seiner Habilitationsschrift weiterentwickelt und auf die Problemstellungen in der Geodäsie angepasst. Auf die Reformulierung der Testgrößen für Intervalle als spezielle Fuzzy-Größen wird verzichtet, da die in KUTTERER (2002b) genannten Konzepte direkt mit den hier abgeleiteten Fuzzy-Größen angewandt werden können. Der interessierte Leser sei auf KUTTERER (2003) verwiesen, der das Vorgehen anhand eines Beispiels für GPS-Beobachtungen anschaulich erklärt.

#### 4.4.3 Vergleich einer rein stochastischen mit einer erweiterten Vorgehensweise

Im Folgenden soll gezeigt werden, welche Aussagen und Schlussfolgerungen man ziehen kann bei einer rein stochastischen Betrachtung der Gesamtunsicherheit (kombinierte Unsicherheit) im Vergleich zur Betrachtung der erweiterten Gesamtunsicherheit, die durch eine Überlagerung von Impräzision und Stochastizität erhalten wird. In diesem Abschnitt soll auf die Überlegungen zurückgegriffen werden, die bereits in Abschnitt 2.3.6 vorgestellt wurden. Es wird die Lagekomponente des dreidimensionalen Linach-Netzes (Abschnitt 4.3.4) untersucht.

Abbildung 4.14(a) zeigt die erweiterten 95%-Punktkonfidenzbereiche  $\mathcal{K}_{0,95}^*$ . Diese werden, wie in den vorangegangenen Abschnitten 4.4.1 und 4.4.2 beschrieben, entweder aus der Minkowski-Summe des Zonotops und der



95%-Konfidenzellipse erhalten oder aus der Fuzzy-Erweiterung nach (4.103) und (4.104). An den 64 Ecken des Zonotops, das hellgrau dargestellt ist, werden die entsprechenden Ellipsoidsektoren addiert. Die Bereiche  $\mathcal{K}_{0,95}^*$  überdecken die geschätzte Punktposition mit einer 95% Konfidenzwahrscheinlichkeit, wenn die Impräzision der Beobachtungen durch Intervalle und die zufällige Variabilität der Beobachtungen durch Stochastizität adäquat ausgedrückt sind.

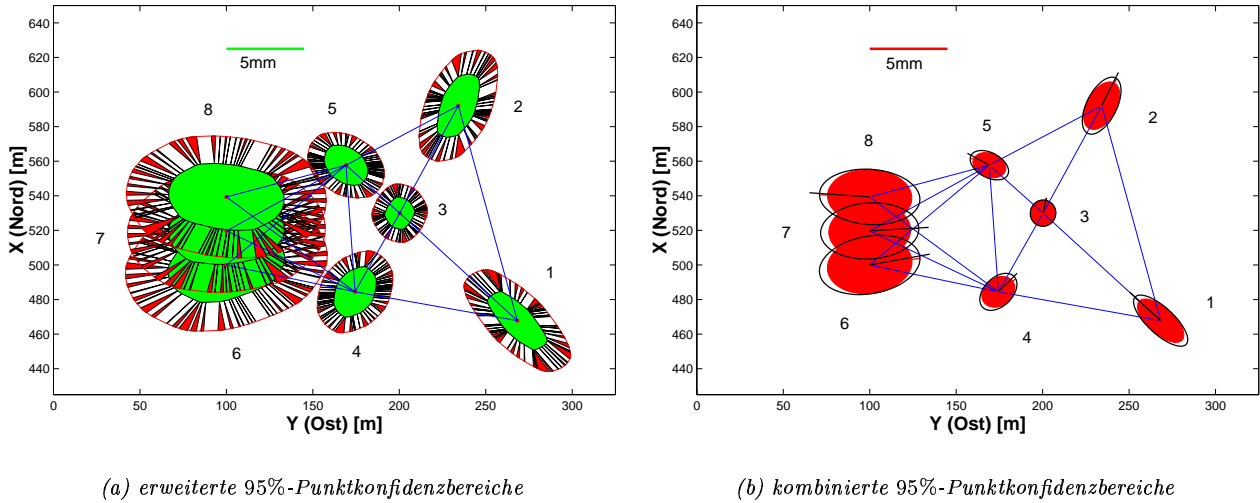


Abbildung 4.14: Erweiterte und kombinierte Punktkonfidenzbereiche

Vergleichend dazu sollen die Punktkonfidenzbereiche betrachtet werden, die man nach den Empfehlungen des Leitfadens (DEUTSCHES INSTITUT FÜR NORMUNG 1995) erhält. Hierzu werden beide Anteile der Gesamtunsicherheit (Impräzision und zufällige Variabilität) durch Varianzen beschrieben und mit dem Varianzfortpflanzungsgesetz kombiniert (KREYSZIG 1979, S.154f). Unterscheiden sich die Verteilungen der einzelnen kombinierten Größen, so ist das Ergebnis nicht mehr statistisch interpretierbar (DEUTSCHES INSTITUT FÜR NORMUNG 1995, S.19ff). Um dieses Problem zu umgehen und um Kenngrößen abzuleiten, die mit der Vorgehensweise der erweiterten Unsicherheitsmaße vergleichbar sind, wird die Normalverteilung für alle beteiligten Einflussparameter angenommen. Es gelte:  $dp_i = \sigma_i$  und damit

$$\mathbf{C}_{pp} = \text{diag}(\mathbf{p}_r \mathbf{p}_r^T). \quad (4.107)$$

Dann berechnet sich die kombinierte Unsicherheit mit  $\mathbf{F}$  und  $\mathbf{p}_r$  aus (2.71):

$$\mathbf{C}_{ll, komb} := \mathbf{C}_{ll} + \mathbf{F} \mathbf{C}_{pp} \mathbf{F}^T. \quad (4.108)$$

Die Betrachtung der Unsicherheit der Einflussparameter der Aufbereitungsschritte der Messungen als Varianzen liefert die vollbesetzte Varianz-Kovarianz-Matrizen  $\mathbf{C}_{ll, komb}$  für die korrigierten Messungen.

Abbildung 4.14(b) zeigt als grau ausgefüllte Ellipsen die 95%-Punktkonfidenzellipsen, die man unter Verwendung von  $\mathbf{C}_{ll}$  erhält, d.h. unter Vernachlässigung des Unsicherheitsbeitrags der Aufbereitungsschritte der Messung. Wird dieser Unsicherheitsbeitrag stochastisch modelliert (Verwendung von  $\mathbf{C}_{ll, komb}$ ), so erhält man die dunkel gezeichneten Ellipsen. Ein Vergleich beider Ellipsen zeigt, dass sich die Hauptrichtungen nur wenig oder gar nicht ändern. Die Form der Ellipse vergrößert sich in Richtung der Streckenmessungen. Dies kann folgendermaßen erklärt werden: Durch die Messung der Richtungen in zwei Lagen sind die verbleibenden Unsicherheitsbeiträge gering, vgl. Abbildung 3.3 und 3.5, im Vergleich zu denen der Strecken, für die keine Reduktionsmechanismen der Unsicherheit des Aufbereitungsprozesses verwendet werden, bzw. den Zenitdistanzen, die nur in einer Lage gemessen wurden, vgl. Abbildung 3.1 und 3.4 sowie die Diskussion im Abschnitt 3.2.

Ein Vergleich der Abbildungen 4.14(a) und (b) zeigt den Unterschied einer quadratischen Fortpflanzung und Mittelung der Unsicherheit im Kontext einer stochastischen Beschreibung (Teilabbildung (b)) und der linearen Fortpflanzung der Impräzision (Teilabbildung (a)).

## 4.5 Alternative Ansätze zur Fortpflanzung der Impräzision

In diesem Abschnitt werden zwei alternative Ansätze zur Fortpflanzung von Impräzision oder allgemeiner zur Behandlung endlicher und beschränkter Fehlerintervalle beispielhaft vorgestellt. Im Gegensatz zum Ansatz der Intervallauswertung, der in der vorliegenden Arbeit verwendet wird und der auf eine statistische Interpretierbarkeit der Ergebnisse zielt, verfolgen die beiden alternativen Ansätze eine rein mengentheoretische Interpretation.

### 4.5.1 Konzept der Minimalen Maximalfehler

Ein früher Ansatz in der Geodäsie zur Betrachtung von Toleranzen durch Intervalle geht auf REINHART (1975) zurück und kann als *Konzept minimaler Maximalfehler* charakterisiert werden. An der Technischen Universität München entstanden in der Folgezeit weitere Arbeiten, die dieses Konzept verwenden: von allgemeiner Natur sind beispielsweise die Aufsätze HEINDL UND REINHART (1976a-c, 1978), Anwendungen für die Netzausgleichung zweidimensionaler Netze stellen GRIMHARDT (1984) und für die Ausgleichung dreidimensionaler Netze SCHMIDBAUER (1989) vor.

Ausgangspunkt ist das funktionale Modell (4.2), wobei keine Stochastizität der Beobachtungen angenommen wird, sondern *Toleranzen* in Form von Intervallen angesetzt werden. Die Intervallgrenzen bilden dabei Schranken, die durch etwaige Beobachtungsfehler nicht überschritten werden können. Es wird nach einer Lösungsmenge, dem *Toleranzpolyeder*, gefragt, für die gilt:

$$\mathcal{T} := \{ \mathbf{x} \in \mathbb{R}^u \mid \mathbf{A}\mathbf{x} = \mathbf{l}, \mathbf{l} \in [\mathbf{l}] \}. \quad (4.109)$$

Die Berechnung der Ecken oder Extrempunkte der Menge  $\mathcal{T}$  wird als lineare Optimierungsaufgabe für den Vektor  $\mathbf{x}$  aufgefasst, vgl. REINHART (1975, S.16ff) und Anmerkungen in Abschnitt 4.3.1. Das Toleranzpolyeder wird dabei in Form von  $2n$  Restriktionen eingeführt. Mit dem Simplexalgorithmus können dann elegant die Ecken der Toleranzpolyeders berechnet werden.

Allerdings liefert das Toleranzpolyeder nur die Menge aller möglichen Lösungsvektoren, die alle gleichberechtigt sind. Es muss nun noch eine Vorschrift angegeben werden, welcher dieser Vektoren als eigentliche Lösung verstanden werden soll, beispielsweise welche Koordinaten konkret einem Punkt zugeordnet werden sollen. SCHMIDBAUER (1989, S.68) weist ausdrücklich darauf hin, dass der Schwerpunkt des Toleranzpolyeders keine Schätzung oder Approximation des Parametervektors darstellt, wie durch die Abbildungen in den Arbeiten REINHART (1975), GRIMHARDT (1984) und SCHMIDBAUER (1989) zu vermuten wäre.

Für den eindimensionalen Fall wird von HEINDL UND REINHART (1976b) eine eindeutige Lösung angegeben, die kurz motiviert werden soll. Unter dem *maximalen Fehler* soll der größte Abstand eines Wertes aus dem Toleranzintervall von den Intervallgrenzen verstanden werden. Folglich wird der minimale maximale Fehler für die Intervallmitte erzielt (HEINDL UND REINHART 1976b). Sie ist deshalb als repräsentativer Wert aufzufassen. Dies entspricht gerade der Vorgehensweise im Konzept der affinen Arithmetik, die in Abschnitt 2.3.5 vorgestellt wurde. Für Parametervektoren kann diese Fragestellung nur komponentenweise gelöst werden, vgl. REINHART (1975, S.5) oder mit Methoden der Vektoroptimierung. HEINDL UND REINHART (1976a) schlagen hierzu die lexikographische Methode vor. Vergleiche für Details der Vektoroptimierung die Ausführungen in Abschnitt 5.1.2.

KUTTERER (1994, S.62ff) hat sich in seiner Arbeit ausführlich mit diesem Ansatz auseinandergesetzt und eine Einordnung in das Intervallkonzept vorgenommen. Im Folgenden soll kurz betrachtet werden, welche Beziehungen zwischen dem Wertebereich der Intervallauswertung und dem Toleranzpolyeder für das selbe Ausgleichungsmodell bestehen. Hierbei muss in den Planungsfall und den Fall des Arbeitens mit realen Daten unterschieden werden.

Betrachten wir zunächst den Planungsfall. Sind die Messdaten mit dem Modell konsistent, wie dies beispielsweise bei Simulationen oder Planungsrechnungen erreicht wird, so gilt:

$$\mathbf{A}\hat{\mathbf{x}} = \mathbf{l}_m. \quad (4.110)$$

Das resultierende Toleranzpolyeder ist dann zentralsymmetrisch (SCHMIDBAUER 1989, S.88ff). In diesem Fall lässt sich aussagen, je kleiner das Toleranzpolyeder, desto besser ist die Punktlage bestimmt. Das Toleranzpolyeder kann maximal  $2n$  Kanten im Zweidimensionalen bzw.  $n(n-1)$  Randflächen im Dreidimensionalen haben. Weniger Kanten bzw. Randflächen treten auf, wenn die Toleranz einer Beobachtung so groß ist, dass sie nicht zur Bestimmung des Toleranzpolyeders beiträgt. Die entsprechende Restriktion in (4.109) heißt dann

redundant, vgl. SCHMIDBAUER (1989, S.16f). Da Zonotope die einzigen zentralsymmetrischen Polytope im Zweidimensionalen darstellen, sind Toleranzpolyeder im Zweidimensionalen Zonotope. Im Dreidimensionalen reicht die Zentralsymmetrie als Kriterium für Zonotope nicht aus, daher sind Toleranzpolyeder im Dreidimensionalen im Allgemeinen keine Zonotope.

Beim Arbeiten mit Realdaten wird (4.110) in den seltensten Fällen erfüllt sein. Daher sind im Allgemeinen keine Aussagen über die Form des Toleranzpolyeders möglich. Je größer die Diskrepanzen zwischen Modell und Daten sind, desto stärker verschieben sich die entsprechenden Toleranzbänder gegeneinander. Liegen grobe Fehler in den Messwerten vor, so ist das Toleranzpolyeder die leere Menge, d.h. Messungen und Modellierung sind nicht kompatibel. Toleranzpolyeder können somit zum Test des Datenmaterials auf grobe Fehler verwendet werden. Entsprechend darf ein kleines Toleranzpolyeder nicht als Hinweis auf eine gut bestimmte Punktlage fehlinterpretiert werden, sondern ist im Gegenteil eher als Indikator von Diskrepanzen zwischen Modell und Daten zu verstehen. KUTTERER (1994, S.65f) weist in diesem Zusammenhang darauf hin, dass Toleranzpolyeder eher als Verträglichkeitsmaße und nicht als Genauigkeitsmaße aufgefasst werden sollten.

**Beispiel 26** Betrachten wir zur graphischen Veranschaulichung der Diskussion in diesem Abschnitt das Beispiel des überbestimmten Bogenschnitts. Abbildung 4.15 zeigt einen Vergleich der Toleranzpolyeder, die schraffiert dargestellt sind, und der Zonotope für den überbestimmten Bodenschnitt. Die Intervallradien der drei Streckenmessungen von den Festpunkten 1, 2, 4 zum Neupunkt 3 wurden nach der Faustformel (3.6) berechnet:  $\mathbf{l}_r = (1.6, 1.8, 1.6)^T [mm]$ .

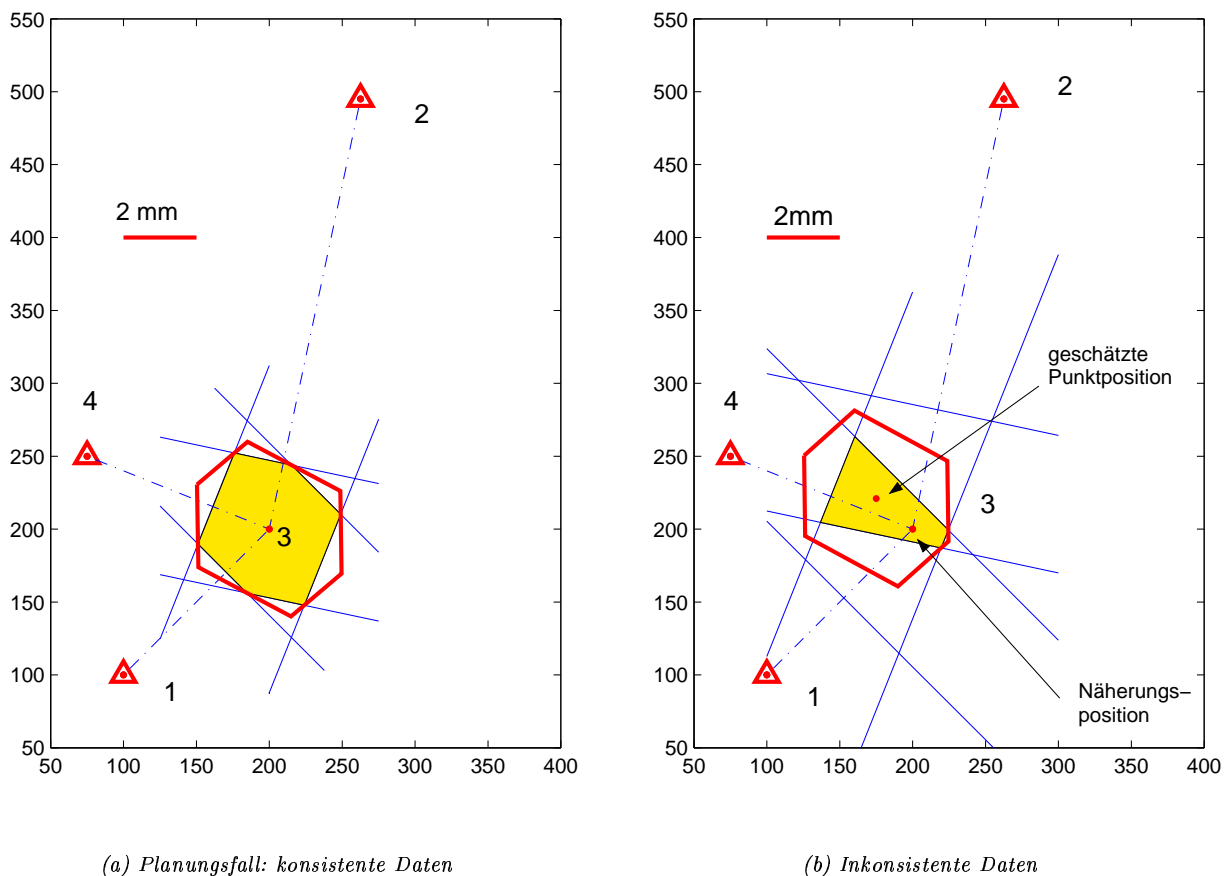


Abbildung 4.15: Vergleich der Toleranzpolyeder (schraffiert) und Zonotope für den überbestimmten Bogenschnitt

Teilabbildung 4.15(a) zeigt den Planungsfall, d.h. Gleichung (4.110) ist erfüllt. Die Toleranzbänder, deren Grenzen eingezeichnet sind, verlaufen senkrecht zu den gestrichelt angedeuteten Messungslinien. Die Kanten des Zonotops weisen gegenüber diesen Richtungen leichte Drehungen auf. Der Schwerpunkt des Toleranzpolyeders fällt mit dem Schwerpunkt des Zonotops, d.h. mit der Näherungsposition zusammen. Das Toleranzpolyeder stellt eine Teilmenge des Wertebereichs dar.

Teilabbildung 4.15(b) zeigt das Toleranzpolyeder (schraffiert) und das Zonotop im Fall des Arbeitens mit (inkonsistenten) Realdaten, die die Gleichung (4.110) nicht erfüllen. Die Diskrepanz betrage  $\mathbf{A}\hat{\mathbf{x}} - \mathbf{1} = -(1, 1.5, 0.75)^T [mm]$ . Dadurch werden im Gegensatz zum Planungsfall die Toleranzbänder gegeneinander verschoben, so dass das resultierende Toleranzpolyeder nicht mehr zentralsymmetrisch ist. Im Extremfall (große Diskrepanzen) ist die Schnittmenge sogar leer. Das Zonotop ist bezüglich der geschätzten Punktposition  $\mathbf{d}\hat{\mathbf{x}} = (-1, 0.8)^T [mm]$  dargestellt, wobei dafür der überhöhte Maßstab verwendet wurde. Es bildet wie im Planungsfall eine Obermenge des Toleranzpolyeders.

Das Beispiel des überbestimmten Bogenschnitts hat die Unterschiede des Toleranzansatzes und der Intervallauswertung deutlich gemacht: Bei der Intervallauswertung liegt eine wohldefinierte Punktposition vor: der geschätzte Koordinatenvektor. Der Wertebereich ist immer ein Zonotop und alle Messungen, die im gesamten Netz vorkommen, tragen zur Form und Größe des Wertebereichs bei. Für eindeutig bestimmte Punktfigurationen fallen Wertebereich und Toleranzpolyeder zusammen, da hier der Wertebereich parallel zu den Messungslinien ausgerichtet ist, vgl. Beispiel 23. Für überbestimmte Punktfigurationen ist dies nicht mehr der Fall, vergleiche Beispiel 26.

Abschließend soll eine Verbindung zwischen der Größe des Toleranzpolyeders im Planungsfall und bei Verwendung von inkonsistenten Daten hergestellt werden. Wir gehen für die folgende Betrachtung davon aus, dass keine kollinearen Zeilenvektoren  $\mathbf{a}^i$  in der Matrix  $\mathbf{A}$  vorliegen. Dies ist gleichbedeutend damit, dass es keine Doppelmessungen im Netz ausgeführt wurden. Zusätzlich sollen keine redundanten Restriktionen vorliegen. Dann ist durch

$$b_i = \frac{2 l_{r,i}}{\|\mathbf{a}^i\|} \quad (4.111)$$

die Ausdehnung des Toleranzbandes für die  $i$ -te Messung gegeben, wobei  $l_{r,i}$  den Intervallradius dieser Messung bezeichnet. Diese maximale Ausdehnung tritt nur auf, wenn die Ebenen mit Normalenvektor  $\mathbf{a}^i$  auch Stützebenen des Polyeders sind, beispielsweise wenn (4.110) erfüllt ist. Daraus folgt, dass das Toleranzpolyeder aus der Planungsphase eine Obermenge des Toleranzpolyeders ist, das bei realen Anwendungen erhalten wird, wobei der Schwerpunkt des Toleranzpolyeders mit den Näherungskoodinaten gleichzusetzen ist.

## 4.5.2 Konzept der Set-Inversion

Der Ansatz der Set-Inversion wurde im Kontext der Kontrolltheorie vorgestellt (JAULIN ET AL. 2001). Ausgangspunkt ist die Annahme, dass bandbegrenzte, endliche Fehler für die Messungen bzw. Eingangsgrößen vorliegen, die durch Intervalle beschreiben werden. JAULIN ET AL. (2001, S. 54ff) unterscheiden zwei prinzipielle Aufgabenstellung: die direkte und die inverse Fragestellung.

**Definition 19** Gegeben seien eine nichtlineare Abbildung:  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , eine abgeschlossene Ausgangsmenge  $\mathcal{X} \subset \mathbb{R}^n$  und eine abgeschlossene Bildmenge  $\mathcal{Y} \subset \mathbb{R}^m$ . Dann heißt

$$\begin{aligned} \mathcal{Y} &= \{ \mathbf{y} \in \mathbb{R}^m \mid \mathbf{y} = \mathbf{f}(\mathbf{x}), \mathbf{x} \in \mathcal{X} \} && \text{direkte Fragestellung,} \\ \mathcal{X} &= \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{f}(\mathbf{x}) \in \mathcal{Y} \} =: \mathbf{f}^{-1}(\mathcal{Y}) && \text{indirekte Fragestellung,} \end{aligned} \quad (4.112)$$

wobei  $\mathbf{f}^{-1}$  nicht als mathematisch strenge Inverse zu betrachten ist, sondern als eine Umkehrabbildung.

Die indirekte Aufgabenstellung (*Set-Inversion*) ist in der Regel einfacher zu lösen als die direkte Aufgabenstellung, vgl. (JAULIN ET AL. 2001, S.59). Für jede beliebige Menge  $\mathcal{Y} \subset \mathbb{R}^m$  und für jede Funktion  $\mathbf{f}$ , für die eine konvergente Einschließungsfunktion existiert, können eine äußere und eine innere Einschließung der Menge  $\mathcal{X}$  als Vereinigung von Hyperboxen berechnet werden, vgl. (JAULIN ET AL. 2001, S.54ff.). Es wird somit garantiert, dass der Lösungsvektor sicher innerhalb der Einschließungen liegt. KIEFFER ET AL. (2000) zeigen die garantierte und robuste Positionsbestimmung von Robotern aus Trilateration unter Verwendung der Set-Inversion. Die tatsächliche Position ist garantiert in der Lösungsmenge enthalten, die durch das Einschließungsverfahren bestimmt wurde. Eine interessante Eigenschaft dieser Vorgehensweise ist, dass beim Vorliegen von Inkonsistenzen der Schnitt der Intervallboxen bei der Abbildung  $\mathbf{f}^{-1}(\mathbf{y})$  leer ist. Inkonsistenzen können durch unzureichende Modellierung, grobe Fehler in den Messungen oder zu optimistisch angenommene Fehlerbänder entstehen. Da solche Situationen aufgedeckt werden können, bezeichnen KIEFFER ET AL. (2000) das Verfahren als robust. BRAEMS ET AL. (2003) verwenden diese Ansätze zum Design von Experimenten bzw. der Auswahl von geeigneten Sensoren anhand von Kalibrierungskurven. Der Artikel veranschaulicht das Vorgehen für den Vergleich zweier Thermometer.

Die Berechnung der Intervallauswertung des Schätzwertes nach der Methode der kleinsten Quadrate stellt einen einfachen Fall der Set-Inversion dar. Die Funktion  $\mathbf{f}$  ist linear und die Menge  $\mathcal{Y}$  ist durch eine Menge in Form der Intervallbox gegeben. Die Set-Inversion könnte zur direkten Lösung nichtlinearer geodätischer Problemstellungen ohne vorherige Linearisierung angewendet werden. Voraussetzung ist hierfür das Vorliegen konvergenter Einschließungsfunktionen für die nichtlinearen Zusammenhänge zwischen Daten und Modellparametern. Für bestimmte Aufgabenstellungen wie die Netzoptimierung, bei der ggf. die Orientierung des Koordinatensystems frei wählbar ist, kann das Verfahren nur mit Einschränkungen verwendet werden, da die Lösungsmenge als Vereinigung von achsparallelen Intervallboxen von der Orientierung abhängt. Ist der Koordinatenrahmen a priori festgelegt, wie bei Arbeiten in konventionellen Referenzsystemen, beispielsweise dem ITRF oder im Gauß-Krüger-Koordinatensystem der Landesvermessung, oder sind die zu analysierenden Parameter keine Koordinaten, sondern beispielsweise Koeffizienten der Kugelfunktionsentwicklung des Schwerefeldes, so kann das Verfahren ohne Probleme angewandt werden.

Die Toleranzaufgabe kann als einfache direkte Fragestellung aufgefasst werden, wobei durch das vorgeschlagene Verfahren nur eine äußere Einschließung erhalten werden kann (JAULIN ET AL. 2001, S.59). Ein ähnlicher Ansatz wurde bereits von KUTTERER (1994, S.73f.) in seiner Dissertation entwickelt.

## 5. Optimierung geodätischer Messoperationen

In den vorangegangenen Kapiteln wurden Intervallradien als Maße für Impräzision motiviert und diese linear auf die geschätzten Parameter (Punktkoordinaten geodätischer Netze) fortgepflanzt. Durch eine Kombination von Stochastizität und Impräzision sind erweiterte Unsicherheitsmaße abgeleitet worden. An verschiedenen Beispielen wurde aufgezeigt, dass diese erweiterten Maße eine realistischere Beurteilung der Unsicherheit geodätischer Netze erlauben, da beide Aspekte der Unsicherheit (Impräzision und Stochastizität) mit ihren unterschiedlichen Charakteristika behandelt werden.

Die Untersuchungen und Ergebnisse des vorliegenden Kapitels sollen den Themenkreis zur Planung und Messung geodätischer Netze schließen. Dabei wird der Begriff „Optimierung“ nicht primär im Sinne einer konkreten optimalen Gestaltung eines vorliegenden Netzes verstanden. Netzoptimierung wird hingegen als Analysetool aufgefasst, das dem Ingenieur erlaubt, grundsätzliche Eigenschaften geodätischer Netze zu verstehen und dieses Wissen bei der Konzeption neuer Netze umzusetzen.

In einem ersten Abschnitt werden Grundbegriffe der skalaren und der Vektoroptimierung zusammengestellt, sofern sie für das Verständnis wesentlich sind. Als Einführung zur geodätischen Netzoptimierung werden die Verfahren der klassischen Vorgehensweise zusammenfassend vorgestellt. Darauf aufbauend werden in einem zweiten Schritt Zielfunktionen für die Impräzision entwickelt. Es wird gezeigt, wie eine Variation der „Grundelemente“ geodätischer Netze (Datum, Geometrie, Topologie und Gewichtung der Messungen) sich auf die Kenngrößen auswirkt.

Neben der Betrachtung der klassischen „Designs“, die Ähnlichkeiten im Verhalten der Kenngrößen der Stochastizität und Impräzision aufweisen, sind neue Gruppierungen notwendig. Diese erlauben erst die adäquate Beschreibung des Einflusses des Mess- und Aufbereitungsprozess der Messung und der Performance der Instrumente. In einem weiteren Abschnitt wird gefragt, wie die Mess- und Auswertemethoden und welche Sensoren konkret gewählt werden sollten, damit die Impräzision der Messung oder der geschätzten Parameter gewisse Maximalwerte nicht übersteigt. Abschließend wird die Optimierung der erweiterten Unsicherheit betrachtet. Durch Verfahren der Vektoroptimierung können verschiedene Zielfunktionen für die Impräzision, die Stochastizität und die Zuverlässigkeit kombiniert werden.

### 5.1 Grundlagen der Optimierung

Im folgenden Abschnitt werden Begriffe und Grundlagen der mathematischen Optimierung mit skalaren Zielfunktionen zusammengefasst. Es genügt dabei, Minimumprobleme für die Zielfunktion  $\Phi$  zu betrachten. Jedes Maximumproblem kann durch Übergang von  $\Phi$  auf  $-\Phi$  in ein äquivalentes Minimumproblem umgewandelt werden. Der Optimierung skalarer Zielfunktionen wird die Vorgehensweise der Vektoroptimierung (*Multi Objective Optimization*) gegenübergestellt. Diese wird angewandt, wenn mehrere Zielforderungen gleichzeitig erfüllt werden sollen.

#### 5.1.1 Skalare Optimierung

**Definition 20** Sei  $\Phi$  eine stetig differenzierbare Funktion  $\Phi : \mathbb{R}^p \rightarrow \mathbb{R}$  von  $p$  freien Parametern  $\mathbf{x}$ , so soll unter einer *skalaren Minimierungsaufgabe* folgendes Problem verstanden werden:

$$\begin{cases} \min_{\mathbf{x}} \Phi(\mathbf{x}) \\ \mathbf{x} \in \Omega \subset \mathbb{R}^p, \end{cases} \quad (5.1)$$

wobei  $\Phi$  *skalare Zielfunktion* heißt und  $\Omega$  *zulässiger Bereich*.

Der zulässige Bereich wird durch  $c$  lineare oder nichtlineare Restriktionen für die Parameter beschrieben. Diese stellen als Nebenbedingungen der Optimierung die maximalen Variationsbereiche für die Parameter dar:

$$\Omega := \{\mathbf{x} \in \mathbb{R}^p \mid \mathbf{c}(\mathbf{x}) \leq \mathbf{0}\}. \quad (5.2)$$

Oftmals werden die Optima auf dem Rand des zulässigen Bereichs angenommen, so dass ihm eine besondere Bedeutung zukommt. Die Wahl der freien Parameter und der zu optimierenden Zielfunktionen legt die Art

der Optimierungsaufgabe fest. Damit ist in der Regel auch über eine geeignete mathematische Lösungsstrategie des Problems entschieden: Sind Zielfunktion und Nebenbedingungen beispielsweise lineare Funktionen der Parameter, so kann (5.1) durch Lineare Programmierung gelöst werden. Andere Problemstellungen lassen sich so formulieren, dass sie beispielsweise geeignet mit der Methode der Kleinsten Quadrate oder dem Minimax-Algorithmus behandelt werden können. Die Rückführung konkreter Problemstellungen auf „Standardprobleme“ hat den Vorteil, dass Standardlösungsverfahren verwendet werden können und wichtige Eigenschaften wie Eindeutigkeit oder Effizienz für die „Standardproblemstellung bzw. -lösungsstrategie“ bekannt sind. Für die Umsetzung der einzelnen Ansätze stehen entsprechende Programmbibliotheken zur Verfügung. Die Beispiele in diesem Kapitel wurden mit der MATLAB-Optimization-Toolbox (Vers.2.2) berechnet. Sollen mehrere Zielfunktionen gleichzeitig optimiert werden, so spricht man von einer Vektoroptimierung. Im Folgenden werden einige Begriffe zur Vektoroptimierung vorgestellt.

### 5.1.2 Vektoroptimierung

#### Definitionen und Ausgangsfragestellungen

**Definition 21** Sei  $\Phi$  eine stetig differenzierbare vektorwertige Funktion  $\Phi : \mathbb{R}^p \rightarrow \mathbb{R}^z$ ,  $z \geq 2$  von  $p$  freien Parametern  $\mathbf{x}$ , so soll unter einer *Vektoroptimierungsaufgabe* folgendes Problem verstanden werden:

$$\begin{cases} \min_{\mathbf{x}} \Phi(\mathbf{x}) \\ \mathbf{x} \in \Omega : \{\mathbf{x} \in \mathbb{R}^p \mid \mathbf{c}(\mathbf{x}) \leq \mathbf{0}\}, \end{cases} \quad (5.3)$$

wobei  $\Phi$  *Vektor der Zielfunktion* heißt und  $\Omega$  *zulässiger Bereich*.

Die grundlegende Problematik der Vektoroptimierung besteht darin, dass keine totale Ordnung der Vektoren im  $\mathbb{R}^n$  existiert, vgl. GÖPFERT UND NEHSE (1990, S.17). Eine Aussage des Typs

$$\mathbf{a} < \mathbf{b}, \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}^n \quad (5.4)$$

ist daher nicht definiert. Im Gegensatz dazu ist für  $a, b \in \mathbb{R}$  die Relation  $a < b$  implizit durch die totale Ordnung der reellen Zahlen gegeben, vgl. GÖPFERT UND NEHSE (1990, S.17f). Die Existenz einer Ordnungsrelation  $\prec$  ist jedoch beispielsweise für die Definition eines Minimums notwendig. Um die Vektoroptimierung zu lösen, muss daher eine Ordnungsrelation  $\prec$  im  $\mathbb{R}^n$  und ein Optimalitätsbegriff definiert werden. Mögliche Ansätze sind beispielsweise:

$$\mathbf{a} \prec \mathbf{b} \Leftrightarrow a_i < b_i, \text{ für alle } i, \text{ oder } \mathbf{a} \prec \mathbf{b} \Leftrightarrow a_1 < b_1 \quad \text{oder} \quad \mathbf{a} \prec \mathbf{b} \Leftrightarrow \|\mathbf{a}\| < \|\mathbf{b}\|. \quad (5.5)$$

Die Ordnungsrelation  $\prec$  wird also auf die Ordnungsrelation  $<$  des  $\mathbb{R}$  zurückgeführt. Ist die Ordnungsrelation  $\prec$  gewählt, so kann der Begriff der Optimalität definiert werden.

**Definition 22** Bezeichne  $\Omega \in \mathbb{R}^u$  den zulässigen Bereich für den Parametervektor  $\mathbf{x}$ , so stellt

$$\Omega' := \Phi(\Omega) \quad (5.6)$$

die Abbildung des zulässigen Bereichs dar.  $\Omega'$  heißt *zulässiger Bereich der Zielfunktion*.

**Definition 23** Ein Element  $\Phi^* \in \Omega'$  heißt *effizient*, falls im Sinne der Ordnungsrelation  $\prec$  in  $\Omega'$  kein Vektor  $\Phi$  existiert, für den gilt:

$$\Phi \prec \Phi^*. \quad (5.7)$$

Die Menge aller effizienten Elemente heißt *Effizienzmenge*  $\Theta$ .

**Definition 24** Sei  $\Omega'$  das Bild des zulässigen Bereichs und  $\Phi^* \in \Omega'$  effizient, so heißt

$$\mathbf{x}^* \in \Omega \text{ mit: } \Phi(\mathbf{x}^*) = \Phi^* \quad (5.8)$$

*paretomaximal* oder *funktionaleffizient*, vgl. GÖPFERT UND NEHSE (1990, S.18) oder KUANG (1996, S.251ff).

## Lösungsverfahren der Vektoroptimierung

Grundidee aller Verfahren ist die Rückführung der vektorwertigen Zielfunktion auf skalare Zielfunktionen (*Skalarisierung*) unter Erhaltung der Ordnungsrelation. Für skalare Zielfunktionen ist die Theorie der Optimierung bekannt. Für die Skalarisierung sind verschiedene Ansätze möglich, vgl. GÖPFERT UND NEHSE (1990, S.41ff). Neben einer komponentenweise Betrachtung aller oder ausgewählter Zielfunktionen kann eine Ordnungen (Präferenz) für einzelne Vektorkomponenten eingeführt werden. Der interessierte Leser sei für weitere Details auf die Literatur zur Vektoroptimierung wie beispielsweise SAWARAGI ET AL. (1985) oder GÖPFERT UND NEHSE (1990) verwiesen. Im Folgenden werden einige Lösungsverfahren der Vektoroptimierung kurz skizziert, die sich aus den verschiedenen Ordnungsrelation (5.5) ergeben.

- Eine einfache Möglichkeit der Definition der Ordnungsrelation  $\prec$  ist die *komponentenweise Betrachtung*:

$$\Phi^* \prec \Phi \Leftrightarrow \Phi_i^* < \Phi_i \quad \text{für alle } i = 1, \dots, n. \quad (5.9)$$

Diese Forderung ist jedoch recht stark. Sie verlangt, dass Punkte im zulässigen Bereich existieren, für die alle Komponenten des Vektors der Zielfunktionen eine Verbesserung ihrer Werte im Vergleich zur Startposition erfahren.

- Ein weiteres Verfahren stellt die *lexikographische Methode* dar, vgl. GÖPFERT UND NEHSE (1990, S.100f). Hierbei wird von einer „natürlichen“ Ordnung der Vektorkomponenten ausgegangen, die beispielsweise durch ihre Indizes festgelegt wird. Die Zielfunktion, die an der ersten Komponente im Vektor steht, ist am wichtigsten. Der Vektor ist somit nach den Vorstellungen des Anwenders total geordnet. Es ist dann iterativ über alle Vektorkomponenten das folgende Optimierungsproblem zu lösen:

$$\begin{cases} \Phi_j \rightarrow \min \\ \mathbf{x} \in \Omega_j, \end{cases} \quad (5.10)$$

wobei  $\Omega_j$  die Elemente von  $\Omega$  enthält, für die die Zielfunktion im  $j - 1$  Iterationsschritt minimal war. Die Iteration wird abgebrochen, falls  $\Omega_j$  noch genau ein Element enthält. Dieses ist dann der Lösungsvektor  $\mathbf{x}^*$ , vgl. GÖPFERT UND NEHSE (1990, S.100).

- Alternativ zu einer Ordnung der Vektorkomponenten durch ihre Indizes, kann jeder Zielfunktion gemäß ihrer Bedeutung für die betrachtete Problemstellung ein Gewicht  $\mu > 0$  zugewiesen werden (*Wichtungsmethode*).

$$\begin{cases} \min_{\mathbf{x}} \sum_{i=1}^z \mu_i \Phi_i, & \mu_i > 0 \\ \mathbf{x} \in \Omega. \end{cases} \quad (5.11)$$

Man erhält eine skalare Ersatzzielfunktion in Form der gewichteten Summe der einzelnen Ausgangszielfunktionen. Das Optimierungsproblem (5.11) hat die besondere Eigenschaft, dass jede Lösung pareto-minimal ist (GÖPFERT UND NEHSE 1990, S.100).

- Eine einfache Möglichkeit, die Vektoroptimierung zu lösen, ist die *Beschränkung auf eine Zielfunktion*, vgl. GÖPFERT UND NEHSE (1990, S.99f):

$$\begin{cases} \Phi_i \rightarrow \min \\ \Omega'' := \{ \mathbf{x} \in \Omega \mid \Phi_i(\mathbf{x}) \leq b_i, \quad \text{für } i \neq j \}. \end{cases} \quad (5.12)$$

Für die anderen Komponenten des Vektors der Zielfunktionen des Ausgangsproblems der Vektoroptimierung, die nicht als Zielfunktion des skalaren Problems gewählt wurden, werden Ungleichungsrestriktionen als Nebenbedingungen eingeführt. Dadurch wird garantiert, dass diese Zielfunktionen bestimmte Kriterien oder Schranken  $b_i$  nicht überschreiten. Für diesen Ansatz gilt beispielsweise (GÖPFERT UND NEHSE 1990, S.99f): Ist  $\mathbf{x}^* \in \Omega''$  die einzige Lösung der Optimierungsaufgabe (5.12), so gehört  $\mathbf{x}^*$  zur Effizienzmenge des Ausgangsproblems der Vektoroptimierung.

Dieser Ansatz wird für die Lösung von Vektoroptimierungsaufgaben in der Geodäsie beispielsweise von SCHAFFRIN (1985) oder KUANG (1996) angewandt. SCHAFFRIN (1985, S.548ff) stellt insgesamt drei Modelle vor, wobei jeweils eines der drei Kriterien Genauigkeit, Zuverlässigkeit und Wirtschaftlichkeit als Zielfunktion angesetzt wird. Für die verbleibenden Forderungen werden durch Restriktionen Werte garantiert, die kleiner als gewisse Vorgaben sind (*control*). Die Auswahl geeigneter Schranken muss sorgfältig getroffen werden, damit der zulässig Bereich nicht zu stark eingeschränkt wird.



- Schließlich ermöglicht die Verwendung einer *gewichteten Norm des Vektors der Zielfunktionen* die Transformation des Ausgangsproblems der Vektoroptimierung auf ein skalares Ersatzproblem. Hierbei wird von der Monotonieeigenschaft der Vektornorm Gebrauch gemacht, vgl. GÖPFERT UND NEHSE (1990, S.103ff).

$$\begin{cases} \min_{\mathbf{x}} \sum_{i=1}^z \mu_i (|\Phi_i|^\alpha)^{\frac{1}{\alpha}}, & \mu_i > 0, 1 \leq \alpha < \infty \\ \mathbf{x} \in \Omega. \end{cases} \quad (5.13)$$

Dieser Ansatz wird auch in der Geodäsie verwendet. Erste Umsetzungen wurden in GRAFAREND ET AL. (1979) besprochen, wobei keine spezielle Wahl der Zielfunktionen vorliegt. KUANG (1996, S.253f) entwickelt ein allgemeines Optimierungsproblem für geodätische Netze auf diesem Ansatz, der auch in weiteren Einzelveröffentlichungen von KUANG besprochen wird. XU (1993) weist darauf hin, dass a priori die Einheiten der Zielfunktionen vereinheitlicht werden müssen, da diese sonst implizit in die Gewichte  $\mu_i$  einfließen. Anwendungen für die dreidimensionale Netzoptimierung werden bei XU UND GRAFAREND (1995) vorgestellt.

## 5.2 Klassische Geodätische Netzoptimierung - ein Überblick

### 5.2.1 Motivation und Einordnung

Nachdem in Abschnitt 5.1 die Grundlagen der Optimierung zusammengestellt worden sind, wird im folgenden Abschnitt ein Überblick über die klassische geodätische Netzoptimierung gegeben. Da die Kenngrößen geodätischer Netze für die Stochastizität und die Zuverlässigkeit der Ergebnisse, vgl. (4.26)-(4.41), wie auch für die Impräzision nur von den Näherungskordinaten und nicht von den Messwerten abhängen, können diese Kenngrößen a priori ohne die Durchführung von Messungen bestimmt werden. Insbesondere kann somit gefragt werden, wie ein geodätisches Netz angelegt werden soll oder welche Messinstrumente zum Einsatz kommen sollen, um optimale Qualitätsmaße zu erhalten oder Anforderungen an die Netzqualität mit geringstmöglichem Aufwand zu erfüllen. Die wesentlichen Kriterien, die dabei an ein geodätisches Netz gestellt werden, sind Anforderungen an die *Genauigkeit*, die *Zuverlässigkeit* und die *Wirtschaftlichkeit*. Dabei wird die Genauigkeit eines geodätischen Netzes durch die Varianz-Kovarianz-Matrix der geschätzten Punktkordinaten oder aus ihr abgeleiteten Größen beschrieben, vgl. Abschnitt 4.1.3. Unter der Zuverlässigkeit eines Netzes wird verstanden, inwieweit grobe Fehler in den Messungen aufgedeckt werden können (innere Zuverlässigkeit) und welche Einflüsse nicht aufgedeckte grobe Fehler in den Messungen auf die geschätzten Koordinaten haben (äußere Zuverlässigkeit). Im Rahmen der Wirtschaftlichkeit sollen der Aufwand und die Kosten, die die Planung, konkrete Messung und Berechnung eines geodätischen Netzes verursachen, quantifiziert werden.

Diese Zielvorgaben werden für reale Netze in der Regel nicht alle gemeinsam erfüllt, da sie gegenläufig sind. So erhöht beispielsweise das Ausführen zusätzlicher (redundanter) Messungen die Zuverlässigkeit des Netzes, gleichzeitig vergrößern sich jedoch die Kosten für die Messung. Ziel ist es, diese sich widersprechenden Zielvorgaben zu einem Ausgleich zu führen. Die geodätische Netzoptimierung ist hierbei als ein mathematisches Hilfsmittel zu verstehen.

Der Zweck der geodätischen Netzoptimierung, so wie sie in der vorliegenden Arbeit verstanden werden soll, ist zweierlei: Primär ist die Optimierung als Baustein der Netzanalyse mit *didaktischen Charakter* zu sehen, d.h. durch die Netzoptimierung soll das grundlegende Verhalten der Kenngrößen geodätischer Netze untersucht und dargestellt werden. Dadurch können komplexe Zusammenhänge wie die Reaktion des Netzes und der Qualitätsmaße auf eine Variationen der Netzgeometrie, des Netzdatums oder der Gewichtung besser verstanden werden. Dieses Wissen und Verständnis bietet den ausführenden Ingenieuren und Wissenschaftlern Leitlinien zur Planung oder Verbesserung geodätischer Netze, die bestmöglich ihre jeweiligen Anforderungen erfüllen. Dies galt im 19.Jh. für die Planung und Ausmessung großer Triangulationsnetze der Landesvermessung wie heute bei der satellitengestützten Vermessung der Erde oder der Überwachung große Ingenieurbauprojekte wie dem Alptransittunnel.

Als zweite Aufgabe können mit den mathematischen Methoden und den Algorithmen, die zur Verfügung gestellt werden, konkret geodätische Netze optimiert werden. Allerdings stößt diese zweite Aufgabenstellung schnell an Grenzen, da bei konkreten Projekten in der Regel nur kleine Variationsbereiche für die Parameter bestehen. So können bei terrestrischen Netzen oftmals die Lage der Netzpunkte nicht oder nur wenig variiert werden, weil sonst beispielsweise Sichtverbindungen verloren gehen, weil eine dauerhafte Vermarkung der Punkte nicht realisiert werden kann oder weil Objektpunkte an Bauwerken spezifische Bauwerkspunkte charakterisieren, so

dass keine Variationsmöglichkeit gegeben ist. Ebenso liefert das verwendete Instrumentarium feste Werte für die Gewichte der Messungen, die nicht beliebig und individuell für jede Messung variiert werden können. Auf diese Aspekte wird in den nachfolgenden Kapiteln an den jeweiligen Stellen eingegangen.

Erste Ansätze der Netzoptimierung findet man im 19. Jh. bei Helmert und Schreiber, vgl. GRAFAREND (1974) oder SCHMITT (1982) für einen historischen Überblick. Seither haben sich die Fragestellungen und Lösungsansätze der klassischen geodätischen Netzoptimierung den Entwicklungen der Optimierungsmethoden einerseits und den Rechnerkapazitäten andererseits angepasst und ständig weiterentwickelt. Standen anfangs skalare Zielfunktionen und eindeutig bestimmbare Netzkonfigurationen im Vordergrund, so können mit der Steigerung der Leistungsfähigkeit der Rechner umfangreiche Netze und Fragestellungen gelöst werden.

Eine Kategorisierung und Klassifizierung der geodätischen Netzoptimierung geht auf GRAFAREND (1974) zurück. Die Literatur zur Netzoptimierung kann angesichts ihrer Fülle im Rahmen dieses Kapitels nicht umfassend zitiert werden. Die folgenden Literaturstellen sollen dem interessierten Leser als Orientierungshilfe und Ausgangspunkt dienen, für weitere Arbeiten soll auf die Fachzeitschriften und Dissertation, die u.a. bei der Deutschen Geodätischen Kommission veröffentlicht sind, verwiesen werden. Eine Übersicht über die Arbeiten bis in die 80er Jahre liefern GRAFAREND ET AL. (1979) oder SCHMITT (1982) und der Sammelband GRAFAREND UND SANSÒ (1985). Weitere Entwicklungen und Anwendungen für die Deformationsanalyse sind bei KUANG (1996) zu finden und in den Proceedings der alle zwei Jahre stattfindenden FIG-Symposien zur Deformationsanalyse. Die jeweiligen Arbeiten sind dabei immer in ihrem historischen Kontext (Entwicklungsstand der Optimierung und Rechnerkapazitäten) zu würdigen und einzuordnen.

## 5.2.2 Zielfunktionen der klassischen Netzoptimierung

Im Folgenden werden Zielfunktionen der geodätischen Netzoptimierung für die Forderungen nach Genauigkeit, Zuverlässigkeit und Wirtschaftlichkeit vorgestellt. Es lässt sich feststellen, dass sich die meisten Autoren mit der Optimierung der Genauigkeit befasst haben. Erst neuere Arbeiten (ab den 1980er Jahren) beschäftigen sich konkret mit der Optimierung der Zuverlässigkeit.

### Zielfunktionen und Optimalitätskriterien für die Netzgenauigkeit

Träger der Genauigkeit der geschätzten oder projizierten Punktkoordinaten des geodätischen Netzes ist die Varianz-Kovarianz-Matrix der Punktkoordinaten. Aus ihr sind die Optimalitätskriterien abzuleiten, wobei gefordert wird, dass diese Kriterien bezüglich einer Translation oder Rotation des Netzes unabhängig sind (GRAFAREND 1974). Es wird zwischen *lokalen* und *globalen Zielfunktionen* unterschieden. Lokale Optimalitätskriterien eignen sich für eine detaillierte Beschreibung der Genauigkeit jedes Netzpunktes. Sie können beispielsweise in Form von absoluten oder relativen Konfidenzellipsen bzw. -ellipsoiden für jeden Netzpunkt beschrieben werden. Globale Optimalitätskriterien weisen dagegen integralen Charakter auf. Mit einer skalaren Kenngröße wird die Varianz-Kovarianz-Situation des gesamten Netzes beschrieben.

Betrachten wir zunächst globale Zielfunktionen. Diese lassen sich als Invarianten der Varianz-Kovarianz-Matrix der Punktkoordinaten bezüglich Rotation und Translation des Netzes darstellen. Diese Invarianten können mit den Eigenwerten  $\lambda_i$  der Kofaktormatrix  $\mathbf{Q}_{\hat{x}\hat{x}}$  dargestellt werden. Mit den in der Optimierung üblichen Bezeichnungen, die von GRAFAREND (1974) bzw. GRAFAREND ET AL. (1979, S.46) auf geodätische Netze übertragen wurden, gilt: Ein geodätisches Netz entspricht den folgenden Optimalitätskriterien, wenn der Wert der Zielfunktion minimal ist.

**A-optimal:** wenn die durchschnittliche Varianz oder der sogenannte mittlere Helmertsche Punktfehler minimal ist:

$$\Phi_A = \sigma_0^2 \operatorname{spur} \mathbf{Q}_{\hat{x}\hat{x}} = \sigma_0^2 \sum_{i=1}^u \lambda_i. \quad (5.14)$$

**D-optimal:** wenn die verallgemeinerte Varianz oder der sogenannte Werkmeistersche Punktfehler minimal ist:

$$\Phi_D = \sigma_0^2 \det \mathbf{Q}_{\hat{x}\hat{x}} = \sigma_0^2 \prod_{i=1}^u \lambda_i. \quad (5.15)$$

**E-optimal:** wenn der maximale Eigenwert der Varianz-Kovarianz-Matrix minimal wird:

$$\Phi_E = \sigma_0^2 \lambda_{max}. \quad (5.16)$$

**N-optimal:** wenn die Frobenius-Norm der Varianz-Kovarianz-Matrix minimal ist:

$$\Phi_N = \|\mathbf{C}_{\hat{x}\hat{x}}\|_F = \sigma_0^2 \sqrt{\sum_{i=1}^u \lambda_i^2}. \quad (5.17)$$

**S-optimal:** wenn die Spannweite  $\lambda_{max} - \lambda_{min}$  oder die Länge des Spektrums der Varianz-Kovarianz-Matrix minimal ist:

$$\Phi_S = \sigma_0^2 (\lambda_{max} - \lambda_{min}). \quad (5.18)$$

Die oben angegebenen Optimalitätsforderungen beziehen sich auf das gesamte Netz. Sie können abgeschwächt werden, wenn nur die Varianz-Kovarianz-Situation einer gewissen Anzahl von Punkten von besonderem Interesse ist wie beispielsweise die Menge der Objektpunkte bei Deformationsnetzen. Anstelle der gesamten Varianz-Kovarianz-Matrix ist dann die entsprechende Submatrix  $\mathbf{\Gamma} \mathbf{Q}_{\hat{x}\hat{x}} \mathbf{\Gamma}^T$  mit der Selektionsmatrix  $\mathbf{\Gamma}$  auszuwählen.

Neben diesen globalen Optimalitätskriterien stehen *lokale Kriterien* zur Verfügung, die die relative und absolute Punktfehlersituation aller Netzpunkte oder einer Auswahl einzeln und detailliert beschreiben. Sie können für jeden Netzpunkt in Form der großen und kleinen Halbachse der Punktfehlerellipse oder durch *Kriteriummatrizen* angegeben werden. Letztere spiegeln eine idealisierte synthetische Form der Varianz-Kovarianz-Situation wider, der sich die Varianz-Kovarianz-Matrix „bestmöglich“ anpassen soll. Das theoretische Konzept geht auf GRAFAREND (1972) und BAARDA (1973) zurück. Als „ideale“ Varianz-Kovarianz-Situation wird angesehen, dass alle Punktfehlerellipsen die gleiche Form aufweisen (*Homogenität*) und dass die Unsicherheit der Punktposition in alle Raumrichtungen gleich groß ist (*Isotropie*).

BILL (1985, S.8ff) stellt verschiedene Vorgehensweisen zur Bestimmung von Kriteriummatrizen zusammen. Die einfachste Kriteriummatrix ist die Einheitsmatrix bzw. eine Diagonalmatrix. Diagonalmatrizen tragen allerdings nicht der Tatsache Rechnung, dass die ausgeglichenen Parameter korreliert sind. Verfeinerte Kriteriummatrizen können aus Korrelationsfunktionen aufgebaut werden. Ausgangspunkt ist die Interpretation geodätischer Netze als Ergebnis eines stochastischen Prozesses (zufällige Abweichung der geschätzten oder projizierten von der wahren Punktlage bei der Realisierung eines Netzes). In diesem Konzept haben sich insbesondere Kriteriummatrizen mit Taylor-Karman oder vollständig isotroper Struktur durchgesetzt, da sie die Eigenschaften der Homogenität und Isotropie aufweisen (BILL 1985, S.9). BILL (1985, Kap.3) diskutiert verschiedene Kovarianz- bzw. Korrelationsfunktionen und weist darauf hin, dass nur bei Verwendung der sogenannten Bessel-, Exponential- oder Gaußfunktion positiv semidefinite Kriteriummatrizen abgeleitet werden können. Das Abklingverhalten der Korrelationsfunktionen hängt von der *charakteristischen Distanz* ab. SCHMITT (1982) bzw. KALTENBACH UND SCHMITT (1988) empfehlen, dafür einen Wert kleiner dem minimalen Abstand zwischen zwei Punkten in einem Netz zu wählen.

Um die Kriteriummatrix als synthetische, ideale Varianz-Kovarianz-Matrix mit der tatsächlichen Varianz-Kovarianz-Matrix vergleichen zu können, müssen beide Matrizen im gleichen Netzdatum vorliegen. Wurde beispielsweise eine reguläre Kriteriummatrix generiert, so muss sie mittels S-Transformation (4.31) in eine Matrix transformiert werden, die den gleichen Datumsdefekt aufweist wie die Varianz-Kovarianz-Matrix. Hierbei gehen allerdings im Allgemeinen die Eigenschaften der Homogenität und Isotropie verloren. Eine Übersicht über die Eigenschaften der S-transformierten Kriteriummatrizen liefert BILL (1985, Tab.3, S.27).

Abschließend soll kurz skizziert werden, wie die Kriteriummatrix  $\mathbf{K}_{xx}$  durch die Varianz-Kovarianz-Matrix angenähert werden kann.

- V. MIERLO (1982) schlägt zur Beurteilung der Approximation der durch die Kriteriummatrix bestimmten idealisierten Varianz-Kovarianz-Situation ein allgemeines Eigenwertproblem vor

$$\mathbf{Q}_{\hat{x}\hat{x}} \mathbf{m} - \mu \mathbf{K}_{xx} \mathbf{m} = \mathbf{0}, \quad (5.19)$$

das bei regulärer Kriteriummatrix in ein spezielles Eigenwertproblem überführt werden kann, so dass gilt:

$$(\mathbf{K}_{xx}^{-1} \mathbf{Q}_{\hat{x}\hat{x}} - \lambda \mathbf{I}) \mathbf{m} = \mathbf{0}. \quad (5.20)$$

Ist der größte Eigenwert  $\mu_{max} \leq 1$ , so werden die Genauigkeitsanforderungen erfüllt (BILL 1985, S.8).

- Ausgangspunkt für eine Optimierung ist das inkonsistente Gleichungssystem

$$\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}} \doteq \mathbf{K}_{xx}^-. \quad (5.21)$$

SCHMITT (1985a, S.78ff) vergleicht drei Ansätze, die zur Lösung von (5.21) entwickelt wurden, worauf der interessierte Leser verwiesen sei.

### Zielfunktionen und Kriterien für die Zuverlässigkeit

Die Zuverlässigkeit stellt ein wesentliches Kriterium für geodätische Netze dar. Umso erstaunlicher scheint, dass das Interesse lange hauptsächlich der Optimierung der Varianz-Kovarianz-Situation gegolten hat, worauf SCHAFFRIN (1985, S.595) hinweist. Ein Grund hierfür mag sein, dass die analytische Formulierung von Zielfunktionen der Zuverlässigkeit aufwändiger ist als für Zielfunktionen, die auf der Varianz-Kovarianz-Matrix basieren, (SCHMITT 1982). Im Folgenden werden einige mögliche Zielfunktionen für die Zuverlässigkeit vorgeschlagen und diskutiert. Sie basieren auf den Kenngrößen, die in Abschnitt 4.1.3 vorgestellt wurden. Die Formulierung von Zielfunktionen findet man beispielsweise bei v. MIERLO (1981), SCHAFFRIN (1985, S.560ff) aber auch bei XU (1993) und XU UND GRAFAREND (1995) oder KUANG (1996, S.228ff, S.269ff).

Setzt man unkorrelierte Messungen voraus, so beschreiben die Redundanzanteile  $r_i$ , vgl. Gleichung (4.35), die Kontrollierbarkeit einer Beobachtung: Je größer der Redundanzanteil, desto kleiner ist ein nicht aufdeckbarer grober Fehler und desto besser ist somit die Beobachtung kontrolliert. Eine erste Forderung, um optimal zuverlässige Netze zu erhalten, ist, die Redundanzanteile zu maximieren. Werden die Redundanzanteile in einem Vektor zusammengefasst, so eignen sich Vektornormen als einfache Zielfunktionen:

$$\Phi_{r,p} = \|(\mathbf{Q}_{\hat{v}} \mathbf{P} l)_{ii}\|_p, \quad (5.22)$$

wobei beispielsweise  $p = 2$  oder  $p = \infty$  gewählt werden kann. Bei der Formulierung von Zielfunktionen für die Redundanzanteile muss bedacht werden, dass bei unkorrelierten Beobachtungen implizit folgende Restriktionen für die Redundanzanteile bestehen:

$$\mathcal{R} : \sum_{i=1}^n r_i = r \quad \text{und} \quad 0 \leq r_i \leq 1. \quad (5.23)$$

Die Gleichungen (5.23) lassen sich geometrisch als  $n + 1$  Hyperebenen für die „Koordinaten“  $r_i$  deuten. Als Schnittmenge ergibt sich ein konvexes Polygon, das in der Ebene  $\mathcal{R}$  liegt und den zulässigen Bereich der Optimierungsaufgabe beschreibt. Die Minimierung der Zielfunktion (5.22) entspricht der Berechnung des (kürzesten) Abstandes der Ebene  $\mathcal{R}$  vom Ursprung. Mit diesen Informationen lassen sich das Minimum und die Minimalstelle der Zielfunktion nach (5.22) direkt angeben:

$$\begin{aligned} \Phi_{r,2,min} &= \frac{r}{\sqrt{n}}, & \text{falls } r_i &= \bar{r} = \frac{r}{n}, \\ \Phi_{r,\infty,min} &= \bar{r} = \frac{r}{n}, & \text{falls } r_i &= \bar{r} = \frac{r}{n}. \end{aligned} \quad (5.24)$$

Ein geodätisches Netz ist in Bezug auf diese Zielfunktionen optimal, wenn alle Redundanzanteile gleich groß sind, vergleiche hierzu BAARDA (1968, S.66f) und die Anmerkungen in KUANG (1996, S.215).

Der *durchschnittliche Redundanzanteil*  $\bar{r}$  eignet sich nur beim Verfahren der *Netzreduktion* oder *Netzerweiterung* als Zielgröße. Bei der Netzreduktion startet man mit einer Netzkonfiguration mit allen möglichen Messelementen zwischen den Punkten. Solange gewisse Grenzwerte für die Genauigkeit oder Zuverlässigkeit nicht über- bzw. unterschritten sind, werden iterativ Beobachtungen gestrichen. Dadurch wird der Aufwand der Realisierung des Netzes minimiert, wobei gewisse Schranken garantiert werden. Bei der Netzerweiterung startet man umgekehrt mit einer Minimalconfiguration des Netzes und fügt iterativ Beobachtungen hinzu, um die Anforderungen an Genauigkeit oder Zuverlässigkeit einzuhalten. Durch Hinzufügen oder Weglassen von Beobachtungen oder Veränderungen in der Netztopologie ändert sich die Gesamtredundanz bzw. der mittlere Redundanzanteil, so dass dieser als Kenngröße in diesen Fällen eingeführt werden kann.

Als Maß der *inneren Zuverlässigkeit* eines geodätischen Netzes wird der Grenzwert für einen nicht aufdeckbaren groben Fehler  $\nabla_0 l_i$  (4.37) angesehen (BAARDA 1968; HECK 1981; JÄGER UND BILL 1985). In einem zuverlässigen Netz sollten für alle Beobachtungen die Grenzwerte möglichst klein sein. Fasst man die Grenzwerte aller Beobachtungen in einem Vektor zusammen:

$$\nabla_0 \mathbf{l} := (\nabla_0 l_1, \dots, \nabla_0 l_n)^T, \quad (5.25)$$

so können diese Forderungen mit folgenden Zielfunktionen umgesetzt werden:

$$\begin{aligned} \Phi_{\nabla_0 \mathbf{l}} &= \nabla_0 \mathbf{l}^T \mathbf{M}^T \mathbf{M} \nabla_0 \mathbf{l}, \\ &= \sigma_0^2 \delta_0^2 \sum_{i=1}^n \frac{m_i^2}{r_i p_i}, \end{aligned} \quad (5.26)$$

wobei die  $n \times n$  Diagonalmatrix  $\mathbf{M}$  mit Diagonalelementen  $m_i$  zur Vereinheitlichung der Einheiten dient und  $\delta_0 = \delta_0(\gamma_0, \beta_0)$  den Nichtzentralitätsparameter der Verteilung bezeichnet, vgl. (4.37). Wählt man speziell  $\mathbf{M}^T \mathbf{M} =$

$\mathbf{P}_{ll}$ , so ergibt sich:

$$\Phi_{\nabla_0 \mathbf{l}, \mathbf{P}_{ll}} = \sigma_0^2 \delta_0^2 \sum_{i=1}^n \frac{1}{r_i}. \quad (5.27)$$

Für diese Wahl lassen sich mit den Restriktionen (5.23) das Minimum und die Minimalstelle angeben. Dies kann geometrisch folgendermaßen motiviert werden: Es wird nach optimalen Kantenlängen  $r_i$  eines Hyperquaders gesucht, die bei konstanter Summe der Kantenlängen das Verhältnis Oberfläche zu Volumen des Hyperquaders  $\prod r_i$  minimieren. Für den zweidimensionalen Fall erfüllt ein Quadrat diese Anforderungen, wobei dabei das Verhältnis der Summe der Kantenlängen zum Flächeninhalt minimiert wird. Analog gilt im Mehrdimensionalen:

$$\Phi_{\nabla_0 \mathbf{l}, \mathbf{P}_{ll}, \min} = \sigma_0^2 \delta_0^2 \frac{n^2}{r}, \quad \text{für } r_i = \frac{r}{n}, i = 1, \dots, n. \quad (5.28)$$

Ein weiterer Ansatz ist die Forderung nach der Reduktion des maximalen Grenzwertes, vgl. XU (1993):

$$\Phi_{\nabla_0 \mathbf{l}, \infty} = \|\mathbf{M} \nabla_0 \mathbf{l}\|_{\infty} \quad (5.29)$$

Diese Forderung zielt auf eine Homogenisierung der Grenzwerte. XU (1993) schlägt weiter vor, das Spektrum der Grenzwerte zu minimieren:

$$\Phi_{\nabla_0 \mathbf{l}, S} = \min_i \left\{ \max_i \{p_i \sqrt{r_i}\} - \min_i \{p_i \sqrt{r_i}\} \right\} \quad (5.30)$$

Für mehrdimensionale Beobachtungen wie Koordinatendifferenzen aus GPS-Beobachtungen kann das Grenzwertellipsoid als Qualitätsmaß herangezogen werden, vgl. (4.38). KUANG (1993) führt dazu eine Minimierung der Eigenwerte der Matrix  $\mathbf{\Gamma}^T \mathbf{P}_{ll} \mathbf{Q}_{\hat{\mathbf{x}} \hat{\mathbf{x}}} \mathbf{P}_{ll} \mathbf{\Gamma}$  durch, was einer Minimierung und Homogenisierung der Halbachsenlängen der Grenzwertellipse für den Vektor der Koordinatendifferenz zwischen zwei Punkten entspricht.

Als Maß der *äußeren Zuverlässigkeit* soll die Netzverzerrung  $\|\nabla_0 \hat{\mathbf{x}}\|_{\mathbf{C}_{\hat{\mathbf{x}}}^{-1}}$ , vgl. Gleichung (4.40), den Abschluss für die Diskussion der Zuverlässigkeitsmaße bilden. Für ein zuverlässiges Netz wird gefordert, dass sich nicht aufdeckbare grobe Fehler nur gering auf die geschätzten Parameter auswirken. Ein quadratisches Maß hierfür ist die Netzverzerrung. Schreibt man die Werte der Netzverzerrung in einen Vektor:

$$\delta^2 := (\delta_{0,1}^2, \dots, \delta_{0,n}^2)^T, \quad (5.31)$$

so kann die Summe der Netzverzerrungen  $\delta_{0,i}^2$  für alle Beobachtungen als Zielfunktion dienen:

$$\begin{aligned} \Phi_{\delta^2, 2} &= \sum_{i=1}^n (\mathbf{P}_{ll} \bar{\mathbf{A}} \mathbf{Q}_{\hat{\mathbf{x}} \hat{\mathbf{x}}} \bar{\mathbf{A}}^T \mathbf{P}_{ll})_{ii} \nabla_0 \mathbf{l}_i \\ &= \delta_0^2 \sum_{i=1}^n \frac{u_{x_i}}{r_i}. \end{aligned} \quad (5.32)$$

Treten keine Zusatzparameter auf, d.h.  $u_{z,i} = 0$ , so ergibt sich mit Gleichung (4.36),  $u_{x_i} + u_{z_i} + r_i = 1$ :

$$\Phi_{\delta^2, 2} = \delta_0^2 \left( \sum_{i=1}^n \frac{1}{r_i} - n \right). \quad (5.33)$$

Ein Vergleich der Struktur der Gleichungen (5.27) und (5.33) liefert, falls keine Zusatzparameter geschätzt wurden, das Optimum für homogene Redundanzanteile  $r_i = \bar{r}$ ,  $i = 1, \dots, n$ :

$$\Phi_{\delta^2, 2, \min} = \delta_0^2 \left( \frac{n^2}{r} - n \right) = \delta_0^2 \left( \frac{u^2}{r} + u \right). \quad (5.34)$$

Eine Anwendung dieser Maße für die Deformationsanalyse zeigt KUANG (1996, S.270ff). Alternativ kann gefordert werden, die maximale Netzverzerrung zu minimieren, vgl. KUANG (1993):

$$\Phi_{\delta^2, \infty} = \|\delta_i\|_{\infty}. \quad (5.35)$$

Zusammenfassend lässt sich festhalten, dass die Zuverlässigkeit weitgehend durch die Redundanzanteile determiniert ist. Darauf weist schon BAARDA (1968, S.79) hin. In Bezug auf die angesprochenen Zielfunktionen ist ein Netz dann optimal zuverlässig, wenn die Redundanzanteile gleich groß sind:  $r_i = \bar{r}$ ,  $i = 1, \dots, n$ .

## Zielfunktionen und Optimalitätskriterien für die Wirtschaftlichkeit

Die Wirtschaftlichkeit eines Netzentwurfes hängt von zwei Aspekten ab. Zum Einen soll mit einem gegebenen Aufwand ein maximaler Ertrag erzielt werden, zum Anderen soll ein vorgegebenes Ertragsziel mit minimalem Aufwand erreicht werden, vgl. GRAFAREND ET AL. (1979, S.79f). GRAFAREND ET AL. (1979) verstehen unter dem Nutzen oder Ertrag eines geodätischen Netzes die „Netzgenauigkeit“ und „Netzzuverlässigkeit“. Als Aufwand sind alle Phasen zu verstehen, die zur Umsetzung des Netzes nötig sind wie Planung, Messung und Berechnung.

Jedes Netz, das die Anforderungen an die Netzgenauigkeit und die Zuverlässigkeitskriterien optimal erfüllt, optimiert somit in der Argumentation von GRAFAREND ET AL. (1979) auch den Ertrag. Ein gegenläufiges Ziel stellt die Kostenminimierung dar. Hierzu sind die Vermessungskosten sinnvoll abzuschätzen. GRAFAREND ET AL. (1979, S.81) stellen für terrestrische Netze eine Kostenfunktion vor, die beispielsweise Fixkosten, Beobachtungskosten und Fahrtkosten beinhaltet. Die Aufstellung dieser Funktion hängt vom Vermessungsprojekt ab und kann nicht in eine einfache Form gebracht werden. Von Seiten der Anwendung, beispielsweise für ein Ingenieurbüro sind diese Fragestellungen aber wesentlich.

SCHAFFRIN (1985) beschreibt die Kosten in Funktion der Gewichtsmatrix der Messungen. Diese Vorgehensweise erlaubt eine sehr einfache direkte Umsetzung. Ob der Ansatz die gesamten Vermessungskosten adäquat beschreibt, hängt davon ab, wie die Gewichte bestimmt werden. Geht man von einem üblichen Ansatz aus, die Gewichte direkt aus den Angaben der Instrumentenhersteller zu verwenden, so sind sie mit der Performance der Instrumente und damit an die Anschaffungskosten oder Leihgebühren gekoppelt. Eine Reduktion der Kosten ist somit gleichbedeutend mit der Wahl eines Instrumentariums aus einer geeigneteren Instrumentenklasse. Auf dieser Ebene greift der Ansatz viel zu kurz, da nur ein sehr geringer Anteil der Vermessungskosten wirklich berücksichtigt wird. Die Kosten für die Planung und Berechnung eines Netzes oder den konkreten Messaufwand bleiben außen vor.

Werden abweichend zum obigen Ansatz die Gewichte von den angewandten Messmethoden oder dem Aufbereitungsprozess der Messung beeinflusst, so spiegeln sie eher den gesamten Aufwand wider, der zur Messung eines Netzes nötig ist. Mit einer gewissen Approximation ließe sich dann ein geringes Gewicht so interpretieren, dass geringere Ansprüche an die Qualität der Messung gestellt werden und somit ein geringerer Aufwand notwendig ist. Als Konsequenz könnte dann beispielsweise ausgesagt werden, dass die Messung in einer Lage ausreiche oder keine Temperaturprofile oder Feuchte erfasst werden müssten. Weitere Ansätze zur Optimierung der Wirtschaftlichkeit werden in SCHMITT (1985a) diskutiert.

Es muss festgehalten werden, dass die Optimierung der Wirtschaftlichkeit, die auf eine Minimierung der Vermessungskosten zielt, der Genauigkeits- oder Zuverlässigkeitsmaximierung entgegensteht. In dieser Polarität liegt der Schlüssel, um von einer reinen Analyse der Reaktion geodätischer Netze auf die Variation der Netzgeometrie, des Datums oder der Gewichtung (erste Aufgabe der geodätischen Netzoptimierung) hin zur konkreten mathematischen Optimierung realer Netze für Ingenieurbüros zu gelangen. Diese Problemstellung kann durch Vektoroptimierung gelöst werden.

### 5.2.3 Feste und freie Parameter in der Optimierung

Eine weitere Einteilung der klassischen geodätischen Netzoptimierung kann durch eine Unterscheidung in feste und freie Parameter bei der Optimierung vorgenommen werden. GRAFAREND (1974) hat dazu vier sogenannte *Designs* vorgeschlagen. Er unterteilt dafür die „Grundbausteine“ eines geodätischen Netzes (Topologie, Geometrie und Datum des Netzes sowie Gewichtung der Beobachtungen) in freie Parameter, die bei der Optimierung verändert werden, um ein Optimum der Zielfunktion zu erreichen und in feste Größen, die nicht verändert werden. Als „Design“ ist somit die Analyse der Reaktion eines geodätischen Netzes auf die Variation gewisser Parametersätze zur Optimierung einer gewählten Zielfunktion zu verstehen. Sollen die Aussagen der einzelnen Designs unabhängig von einander sein, so muss dies auch für die einzelnen Parametergruppen gelten. Die Gewichte der Beobachtungen sind, wenn keine Zusatzparameter im Netz auftreten, am einfachsten zugänglich. Netzgeometrie und Topologie sind in der Designmatrix enthalten und ihr Nullraum beschreibt die Freiheitsgrade der Netzlagerung, die durch eine geeignete Datumsgebung ausgefüllt werden müssen.

**Design 0. Ordnung:** Das Design Nullter Ordnung (Zero Order Design, ZOD) stellt die Frage nach der besten Datumsgebung für ein geodätisches Netz bezüglich der gewählten Zielfunktion. Da die oben vorgestellten Kenngrößen der inneren und äußeren Zuverlässigkeit sowie die Redundanzanteile von der Datumsgebung unabhängig sind, eignen sich nur Optimalitätskriterien der Genauigkeit. Wichtigstes Hilfsmittel für diese

Untersuchungen ist die S-Transformation, mit der die Varianz-Kovarianz-Matrix ohne erneute Ausgleichung in ein beliebiges Datum transformiert werden kann. Mit der Selektionsmatrix  $\mathbf{\Gamma}$  können Teilmengen der Netzknoten zur Datumsgebung verwendet werden, vgl. Abschnitt 4.1.3.

Mit den Eigenschaften der Pseudoinversen lässt sich das ZOD-Problem direkt lösen. Für jede generalisierte Inverse  $\mathbf{N}^-$  der Normalgleichungsmatrix, die die Bedingung  $\mathbf{N}^- \mathbf{\Gamma} \mathbf{G} = \mathbf{0}$  erfüllt, vgl. (4.13), gilt (JÄGER 1988, S.53ff):

$$\text{spur}(\mathbf{N}^+) < \text{spur}(\mathbf{N}^-), \quad \|\mathbf{N}^+\|_F < \|\mathbf{N}^-\|_F, \quad \text{und} \quad \lambda_{\max}(\mathbf{N}^+) < \lambda_{\max}(\mathbf{N}^-). \quad (5.36)$$

Als optimale Datumswahl wird daher das sogenannte *innere Datum* angesehen, das durch die Pseudoinverse  $\mathbf{N}^+$  als Varianz-Kovarianz-Matrix der ausgeglichenen Koordinaten gekennzeichnet ist.

**Design 1. Ordnung:** Im Design erster Ordnung (First Order Design, FOD) werden die Netzgeometrie und Netztopologie variiert, um die Zielfunktionen zu minimieren. Dabei bleibt die Datumsgebung fest. Werden streckenlängenabhängige Gewichte eingeführt, so hängt die Gewichtung ebenfalls von der Netzgeometrie ab. In diesen Fällen ist eine strikte Trennung in ein FOD und das nachfolgend besprochene Design zweiter Ordnung nicht möglich.

Die Lösung des FOD kann unter verschiedenen Zielsetzungen erfolgen. Zum Einen kann der Beobachtungsplan (Topologie) konstant gehalten werden und nur die Lage der Netzknoten variiert werden. KUANG (1996, S.288ff) zeigt verschiedene Netzoptimierungen im FOD, wobei die Punktverschiebung 10 – 20% der Netzausdehnung betragen, um deutliche Verbesserungen der Zielfunktionen zu liefern. Durch eingeschränkte Sichtverbindungen und bewegte Topographie ist die Variationsregion für Punktpositionen in Realnetzen im Allgemeinen jedoch stark beschränkt.

Zum Anderen kann die Netztopologie verändert werden, d.h. Beobachtungen werden zum Beobachtungsplan hinzugenommen oder weggelassen. Diese Problemstellung wird bevorzugt durch Simulation gelöst, indem man entweder von der maximal möglichen Anzahl an Messungen zwischen allen Punkten ausgeht und dann sukzessive Beobachtungen aus dem Beobachtungsplan streicht (*Netzreduktion*) oder indem man mit dem minimalen Design startet, das nötig ist, um das Netz überhaupt zu bestimmen und dann Messungen hinzufügt (*Netzerweiterung*).

**Design 2. Ordnung:** Das Design zweiter Ordnung (Second Order Design, SOD) zielt auf eine Variation der Gewichtung bei festem Netzdatum sowie fester Netztopologie und -geometrie. Mathematisch gesehen ist es am einfachsten handhabbar. Da durch eine Skalierung der Gewichte mit einem Faktor  $\mu > 0$ :  $\mathbf{P}'_{ll} := \mu \mathbf{P}_{ll}$  jede Zielforderung erfüllt werden kann, muss dieser Fall durch eine geeignete Restriktion ausgeschlossen werden, beispielsweise  $\sum_{i=1}^n p_i = \text{const.}$  (XU 1993). Für das SOD lassen sich zwei Anwendungsaspekte unterscheiden:

Viele Autoren führen im SOD alle Gewichte als Parameter ein (Individualgewichte). Als „Nebeneffekt“ dieser Parameterwahl werden Ergebnissgewichte erzielt, die sehr klein werden. Dies wird in der Regel so interpretiert, dass die entsprechende Beobachtung keinen Einfluss auf die Zielfunktion hat und als Konsequenz aus dem Beobachtungsplan gestrichen werden kann. Das SOD wird somit zur Optimierung der Netztopologie verwendet. Sekundär ist dabei, wie die Ergebnissgewichte konkret umgesetzt werden sollen.

Als Lösung bietet sich an, nach der Optimierung solche Instrumente zu verwenden, die Gewichte liefern, die sich den optimalen Vorgaben am „besten“ anpassen. Alternativ dazu werden Wiederholzahlen eingeführt, vgl. SCHMITT (1985b), mit deren Hilfe die optimierten Gewichte so transformiert werden, dass sie durch Instrumentarien realisiert werden können. Dieser Ansatz ist kritisch zu bewerten. Einerseits wurde das angezeigte Messergebnis bei der Messung mit modernen Tachymetern bereits aus einer Vielzahl von Einzelmessungen gewonnen, so dass gefragt werden muss, was Wiederholzahlen dann aussagen. Andererseits wurde in Abschnitt 2.3.6 gezeigt, dass zu optimistische Endergebnisse entstehen können, wenn nur die Stochastizität als Unsicherheitsmaß verwendet wird. Erst bei einer Betrachtung der erweiterten Unsicherheit (Stochastizität und Impräzision) können Wiederholzahlen als Mittel angesehen werden, die Stochastizität zu reduzieren. Die Impräzision wird durch Wiederholungsmessungen nicht beeinflusst, vgl. Abschnitt 2.3.6.

Im Hinblick auf die konkrete Umsetzung des optimierten Netzentwurfs wäre es vorzuziehen, nur die Performance-Parameter der Instrumente als freie Parameter bei der Optimierung einzuführen. Aus diesen werden mit einem Modell die Gewichte berechnet. Wegen der darin bestehenden Koppelung können die Gewichte nur bestimmte Werte annehmen und nicht individuell variiert werden. Das SOD liefert dann

Aussagen zur Gewichtsrelation der einzelnen Messungstypen (Richtungen, Strecken und Zenitdistanzen im terrestrischen Fall).

**Design 3. Ordnung:** Im Design dritter Ordnung (Third Order Design, THOD) wird von einem bestehenden Netz mit bekanntem festen Datum, Topologie und Geometrie ausgegangen, vgl. SCHMITT (1985c). Bezüglich der Zielfunktion wird dann nach der optimalen Bestimmung von ergänzenden Neupunkten im bestehenden Netz gefragt. Hierbei steht neben der optimalen Punktlage (FOD-Anteil) auch die optimale Wahl der Gewichtung der neu einzuführenden Messungen (SOD-Anteil) im Interesse. Das THOD ist somit eine Koppelung aus FOD und SOD Anteilen. Für die konkrete Lösung vgl. beispielsweise ILLNER (1986).

## 5.3 Optimierung der Impräzision: direkte Fragestellung

In diesem Abschnitt wird ein erster neuer Aspekt der Netzoptimierung betrachtet, der sich aus der Berücksichtigung der Impräzision ergibt. Unter der *direkten Fragestellung* soll die Sensitivität oder Abhängigkeit der Kenngrößen der Impräzision von den „Grundbausteinen“ des geodätischen Netzes: der Gewichtung, des Datums, der Topologie und der Geometrie verstanden werden. In einem ersten Abschnitt werden Zielfunktionen für die Impräzision formuliert, die auf den Kenngrößen der Impräzision aufbauen. In einem weiteren Abschnitt wird die Reaktion der Kenngrößen der Impräzision auf die Variation der Grundbausteine des Netze gezeigt, wobei die Einteilung den klassischen Designs der Netzoptimierung folgt. Eine abschließende Wertung arbeitet einerseits Ähnlichkeiten im Reaktionsverhalten für Zielfunktionen der Netzgenauigkeit und -impräzision heraus und verdeutlicht andererseits die wesentlichen Unterschiede dieser beiden Komponenten der erweiterten Unsicherheit.

### 5.3.1 Zielfunktionen für die Impräzision

Erste Ansätze zur Optimierung der Impräzision wurden von SCHÖN UND KUTTERER (2001b,c) vorgestellt, wobei die Diskussion auf Basis der Intervallradien geführt wurde. Als Zielfunktionen wurden Matrixnormen der Matrix  $\mathbf{K}$  vorgeschlagen:

$$\Phi_2(\mathbf{K}) = \|\mathbf{K}\|_2 = s_{max}(\mathbf{K}), \quad (5.37)$$

$$\Phi_F(\mathbf{K}) = \|\mathbf{K}\|_F = \sqrt{\sum_{i=1}^u \sum_{j=1}^n k_{ij}^2}, \quad (5.38)$$

$$\Phi_\infty(\mathbf{K}) = \|\mathbf{K}\|_\infty = \max\{\hat{x}_{r,i}\}. \quad (5.39)$$

Die angegebenen Zielfunktionen sind als *globale Zielfunktionen* für die Impräzision zu verstehen. Auf eine detaillierte Interpretation wird im Abschnitt 5.3.3 eingegangen. Es muss festgehalten werden, dass die 2-Norm und Frobenius-Norm, vgl. Gleichungen (4.90) und (4.91), invariant gegen Translation und Rotation des Netzes sind. Die  $\infty$ -Norm, die dem maximalen Intervallradius der Koordinaten entspricht, ist allerdings von der Orientierung des Koordinatensystems abhängig.

Als *lokale Zielfunktionen* eignen sich Kenngrößen der Zonotope wie das Volumen  $Vol(\mathcal{W}_{\hat{\mathbf{x}}_P})$ , der Umfang  $Peri(\mathcal{W}_{\hat{\mathbf{x}}_P})$ , die maximale Breite  $\max\{d_{max}, w_{max}\}$ , die in Abschnitt 4.3.3 vorgestellt wurden. Der Abstand gegenüberliegender Ecken wird mit  $w_i$  bezeichnet und der Abstand paralleler Kanten bzw. Randflächen mit  $d_i$ .

$$\Phi_{Vol} = Vol(\mathcal{W}_{\hat{\mathbf{x}}_P}), \quad (5.40)$$

$$\Phi_B = \max\{d_{max}, w_{max}\}, \quad (5.41)$$

$$\Phi_I = |w_{max} - w_{min}|, \quad (5.42)$$

$$\Phi_{Peri} = Peri(\mathcal{W}_{\hat{\mathbf{x}}_P}). \quad (5.43)$$

Durch die Zielfunktionen (5.40) - (5.43) kann die Form des Zonotops beeinflusst werden. Die Forderung (5.42) nach Isotropie wird für reguläre Zonotope erfüllt, deren Kanten alle gleich lang sind. Dadurch wird gleichzeitig der Abstand  $|d_{max} - d_{min}|$  minimiert. Für weitere Extremaleigenschaften vergleiche beispielsweise LINHART (1986, 1988) oder FILLIMAN (1988). Allerdings eignen sich diese Zielfunktionen wenig für eine analytische Behandlung der Optimierungsaufgaben, da die Formeln entweder nicht differenzierbar sind, vgl. Volumen (4.83), oder von der zyklischen Ordnung der Kantenvektoren in der Matrix  $\mathbf{K}$  abhängen wie der Durchmesser und Breite, vgl. Gleichungen (4.84)-(4.88). Die zyklische Ordnung kann für jeden Netzpunkt unterschiedlich sein. Darüber hinaus ist für eine numerische Optimierung die Berechnung dieser Zielfunktionen zu aufwändig, um effiziente Algorithmen bereit zu stellen.



### 5.3.2 Reaktion globaler Zielfunktionen auf Variationen der Parameter

Im folgenden Abschnitt wird untersucht, wie die Zielfunktionen der Impräzision auf eine Variation des Netzdatums, der Netzgeometrie oder Gewichtung reagieren. Ziel ist es, anhand einfacher Netzbeispiele einerseits Kriterien für optimale Netzkonfigurationen in Bezug auf Zielfunktionen der Impräzision abzuleiten. Andererseits sollen die Ergebnisse mit entsprechenden Erkenntnissen für Zielfunktionen der Stochastizität verglichen werden. Lassen sich Ähnlichkeiten im Verhalten erkennen, so führt die Optimierung einer der beiden Komponenten auch zur Optimierung der Gesamtunsicherheit. Bei den Untersuchungen beschränken wir uns auf die Diskussion globaler Zielfunktionen, die aus den spektralen Kenngrößen der Matrix  $\mathbf{K}$  der Kantenvektoren abgeleitet werden.

#### Einfluss des Netzdatums

Die Kenngrößen des Wertebereichs hängen wie die Form und Größe der Konfidenzellipsen bzw. -ellipsoide in der klassischen Netzoptimierung von der Datumsgebung ab. Eine spezielle Wahl der Datumspunkte verändert die Kenngrößen, was am Beispiel des Linach-Netzes gezeigt wird.

**Beispiel 27** Betrachten wir als Beispiel die zweidimensionale Ausgleichung des Linach-Netzes.

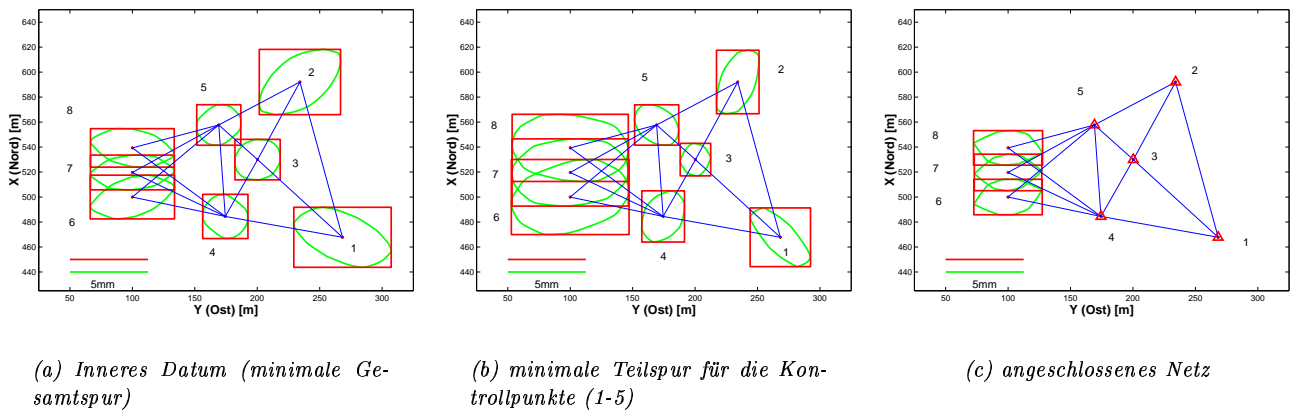


Abbildung 5.1: Vergleich der Intervallboxen und Wertebereiche für unterschiedliche Datumsgebungen beim Linach-Netz. Bestimmung der Intervallradien nach den Faustformeln (3.2), (3.4) und (3.6).

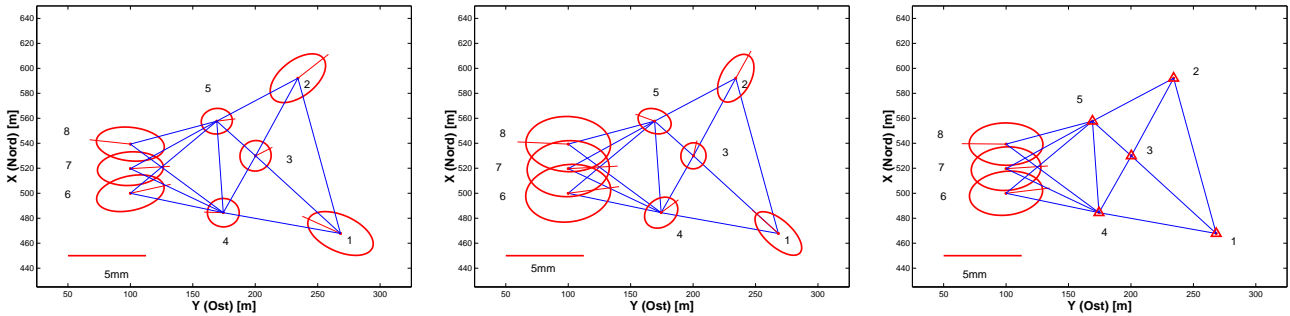
Teilabbildung 5.1(a) zeigt die Zonotope der Punktpositionen für das Linach-Netz bei einer Minimierung der Norm des Vektors der Parameterzuschläge (inneres Datum). Im Gegensatz dazu wird in der Teilabbildung 5.1(b) die Situation dargestellt, wenn eine Teilnormminimierung über die Kontrollpunkte (Nr. 1-5) durchgeführt wird. Die Impräzision der Position der Kontrollpunkte wird dadurch verringert, wohingegen die der Objektpunkte erhöht wird. Die Wertebereiche bei Minimierung der Norm des gesamten Vektors der Koordinatenzuschläge zeigen somit eine homogenere Situation. Dies entspricht dem Verhalten der Konfidenzellipsen, die in der Abbildung 5.2(a,b) als Vergleich gegenübergestellt werden. Für weitere Diskussionen des Einflusses der Datumsgebung auf die Größe und Form der Konfidenzellipsen vergleiche beispielsweise WELSCH ET AL. (2000, S.206ff) oder NIEMEIER (2002, S.230ff).

Die visuelle Interpretation der Abbildungen soll durch Werte der globalen Zielfunktionen belegt werden, die als Matrixnormen der Matrix der Kantenvektoren darstellbar sind. Zum Vergleich sind Matrixnormen der Varianz-Kovarianz-Matrix mit angegeben.

Datum	$\ \mathbf{K}\ _2$ [mm]	$\ \mathbf{K}\ _F$ [mm]	$\ \mathbf{K}\ _\infty$ [mm]	$\ \mathbf{C}_{\hat{x}\hat{x}}\ _2$ [mm <sup>2</sup> ]	$\ \mathbf{C}_{\hat{x}\hat{x}}\ _F$ [mm <sup>2</sup> ]	$\text{spur}\mathbf{C}_{\hat{x}\hat{x}}$ [mm <sup>2</sup> ]
minimale Gesamtspur	1.1	1.8	3.1	1.8	2.3	5.8
minimale Teilspur	1.3	2.1	3.7	2.5	3.1	7.4
angeschlossenes Netz	0.8	1.4	2.2	1.2	1.7	3.6

Tabelle 5.1: Vergleich der Werte der Zielfunktionen in Funktion der Datumsgebung für das Linach-Netz

Teilabbildung 5.1(c) gibt ergänzend die Zonotope der Objektpunkte an, wenn die Kontrollpunkte als bekannt vorausgesetzt werden. In diesem Fall werden die kleinsten Zonotope für die Objektpunkte erhalten. Dies kann folgendermaßen erklärt werden: Bei der Intervallauswertung liefern alle Messungen einen Beitrag zur Form der Zonotope. Diese wird durch die gesamte Netzgeometrie sowie durch die Größe der Intervallradien der Messungen (Impräzision) bestimmt. Die Ergebnisse sind nicht direkt mit der Datumsgebung in freien Netzen vergleichbar, da einerseits weniger Messung zur Bestimmung der Objektpunkte beitragen und andererseits die Netzgeometrie eine andere ist: Es würde eine Betrachtung des Subnetzes aus den Punkten (4-8) ausreichen.



(a) Inneres Datum (minimale Gesamtspur)

(b) minimale Teilspur für die Kontrollpunkte (1-5)

(c) angeschlossenes Netz

Abbildung 5.2: Vergleich der 95%-Konfidenzellipsen für unterschiedliche Datumsgebungen beim Linach-Netz. Für die Beobachtungen wurden folgende Standardabweichungen angenommen:  $\sigma_S = 2 \text{ mm} + 2 \text{ ppm}$ ,  $\sigma_R = 0.5 \text{ mgon}$ .

In Abschnitt 5.2.3 wurde motiviert, warum die innere Datumsgebung als Lösung für das ZOD-Problem für Zielfunktionen der Stochastizität angesehen wird. Im Folgenden wird gezeigt, dass eine innere Datumsgebung bei freien Netzen auch für Zielfunktionen der Impräzision optimal ist. Betrachten wir dazu die 2-Norm der Matrix der Kantenvektoren.

Jeder Vektornorm  $\|\mathbf{y}\|$  ist eine durch den folgenden Ausdruck definierte Matrixnorm zugeordnet, vgl. NOBLE UND DANIEL (1988, S.263ff) oder CASPARY UND WICHMANN (1994, S.60ff):

$$\|\mathbf{B}\| := \max_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{B}\mathbf{y}\|}{\|\mathbf{y}\|}. \quad (5.44)$$

Weiter gilt, dass der Vektor der ausgeglichenen Parameter bei freien Netzen und Verwendung der Pseudoinversen minimale Norm aufweist, vgl. ILLNER (1983) oder KOCH (1997, S.203):

$$\|\hat{\mathbf{x}}(\mathbf{N}^+, \mathbf{l})\| < \|\hat{\mathbf{x}}(\mathbf{N}^-, \mathbf{l})\|, \quad \text{für alle } \mathbf{l} \in \mathbb{R}^n \setminus \{\mathbf{0}\}, \quad (5.45)$$

wobei  $\mathbf{N}^+$  die Pseudoinverse und  $\mathbf{N}^-$  eine reflexive generalisierte Inverse der Normalgleichungsmatrix (4.5) bedeutet. Die Schreibweise  $\hat{\mathbf{x}}(\mathbf{N}^+, \mathbf{l})$  soll verdeutlichen, welche generalisierte Inverse und welcher Beobachtungsvektor  $\mathbf{l}$  zur Berechnung des Vektors der Schätzwerte der Parameter verwendet wurden.

Es soll gezeigt werden, dass die Matrix-Norm der Matrix der Kantenvektoren bei Verwendung der Pseudoinversen kleiner ist als bei Berechnung mit jeder anderen reflexiven generalisierten Inversen:

$$\|\mathbf{K}(\mathbf{N}^+)\| < \|\mathbf{K}(\mathbf{N}^-)\|. \quad (5.46)$$

Der Vektor  $\mathbf{l}$  in (5.45) lässt sich mit der positiven Diagonalmatrix der Intervallradien der Messungen  $\mathbf{L}_r$  in folgender Form ausdrücken:

$$\mathbf{l}' := \mathbf{L}_r \mathbf{l}_K, \quad \text{für alle } \mathbf{l}_K \in \mathbb{R}^n \setminus \{\mathbf{0}\}. \quad (5.47)$$

Da Gleichung (5.45) für alle  $\mathbf{l} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$  gilt, ist auch folgende Gleichung erfüllt:

$$\|\hat{\mathbf{x}}(\mathbf{N}^+, \mathbf{l}')\| < \|\hat{\mathbf{x}}(\mathbf{N}^-, \mathbf{l}')\|, \quad \text{für alle } \mathbf{l}' \in \mathbb{R}^n \setminus \{\mathbf{0}\}. \quad (5.48)$$

Mit der Definition der Matrix der Kantenvektoren nach (4.73) folgt aus (5.48):

$$\begin{aligned} \|\mathbf{K}(\mathbf{N}^+)\mathbf{l}_K\| &< \|\mathbf{K}(\mathbf{N}^-)\mathbf{l}_K\| \quad \text{für alle } \mathbf{l}_K \in \mathbb{R}^n \setminus \{\mathbf{0}\} \\ \Leftrightarrow \frac{\|\mathbf{K}(\mathbf{N}^+)\mathbf{l}_K\|}{\|\mathbf{l}_K\|} &< \frac{\|\mathbf{K}(\mathbf{N}^-)\mathbf{l}_K\|}{\|\mathbf{l}_K\|}. \end{aligned} \quad (5.49)$$

Das Maximum der linken Seite von (5.49) werde an der Stelle  $\mathbf{l}_K^*$  angenommen und habe den Wert:

$$\max_{\mathbf{l}_K \neq \mathbf{0}} \frac{\|\mathbf{K}(\mathbf{N}^+)\mathbf{l}_K\|}{\|\mathbf{l}_K\|} := \frac{\|\mathbf{K}(\mathbf{N}^+)\mathbf{l}_K^*\|}{\|\mathbf{l}_K^*\|}. \quad (5.50)$$

Da (5.49) für alle  $\mathbf{l}_K \in \mathbb{R}^n \setminus \{\mathbf{0}\}$  gilt, folgt:

$$\begin{aligned} \frac{\|\mathbf{K}(\mathbf{N}^+)\mathbf{l}_K^*\|}{\|\mathbf{l}_K^*\|} &< \frac{\|\mathbf{K}(\mathbf{N}^-)\mathbf{l}_K^*\|}{\|\mathbf{l}_K^*\|} \leq \max_{\mathbf{l}_K \neq \mathbf{0}} \frac{\|\mathbf{K}(\mathbf{N}^-)\mathbf{l}_K\|}{\|\mathbf{l}_K\|}, \\ \Rightarrow \|\mathbf{K}(\mathbf{N}^+)\| &< \|\mathbf{K}(\mathbf{N}^-)\|, \quad \text{womit die Behauptung gezeigt ist.} \end{aligned} \quad (5.51)$$

### Einfluss der Netzgeometrie

In diesem Abschnitt wird untersucht, wie eine Variation der Netzgeometrie (Punktlage) bei Beibehaltung der Netztopologie die Werte der Zielfunktionen der Impräzision beeinflusst. Im Allgemeinfall hängt dieser Einfluss vom gewählten Netzdesign ab und muss somit für jedes Netz getrennt untersucht werden. Eine erste Diskussion des FOD für Zielfunktionen der Impräzision wurde von SCHÖN UND KUTTERER (2001b) für das Linach-Netz durchgeführt. Im Folgenden werden synthetischen Netzbeispiele (Diagonalenvierecksketten, flächenhaftes Netz aus gleichseitigen Dreiecken) betrachtet, die eine regelmäßige geometrische Struktur aufweisen. Einerseits werden diese Netzstrukturen verwendet, um allgemeinere Tendenzen der Reaktion von Kenngrößen der Impräzision auf eine Variation der Punktlage abzuleiten, die zu einer Minimierung der globalen Zielfunktionen  $\Phi_2$ ,  $\Phi_F$  und  $\Phi_\infty$  führen. Wegen der Größe der Netze wird die Diskussion qualitativ geführt. Andererseits erlaubt die Wahl dieser Netze Vergleiche zum Verhalten bei der Optimierung klassischer Zielfunktionen. Entsprechende Untersuchungen für diese Zielfunktionen liegen beispielsweise von BILL ET AL. (1984), JÄGER (1988) und KALTENBACH (1992) für Diagonalenvierecksketten bzw. SCHMITT (1997) für flächenhafte Netze vor.

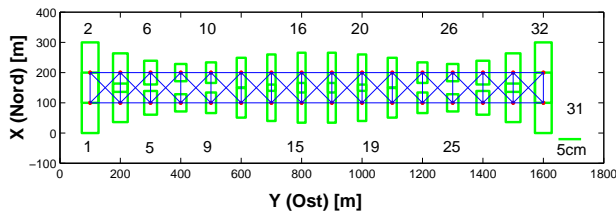
**Beispiel 28** Betrachten wir als erstes Beispiel eine Diagonalenviereckskette von 32 Punkten als reines Streckennetz. Es wird ein Gewichtsansatz gewählt, der von der Streckenlängen unabhängig ist. Für die achsparallelen Strecken ergibt sich:  $S = 100.000 \text{ m}$ ,  $S_r = 1.6 \text{ mm}$ ,  $\sigma_S = 2.0 \text{ mm}$ . Für die diagonal verlaufenden Strecken erhält man  $S = 141.42 \text{ m}$ ,  $S_r = 1.7 \text{ mm}$ ,  $\sigma_S = 2.0 \text{ mm}$ . Es werden maximalen Änderungen der Punktkoordinaten von  $\pm 50 \text{ m}$  zugelassen.

Abbildung 5.3 ist in zwei Spalten unterteilt. In der linken Spalte werden das Ausgangsdesign und die Punktlageverschiebung gezeigt. Diese ist qualitativ für alle drei untersuchten Zielfunktionen gleich. In der rechten Spalte werden die optimalen Netzkonfigurationen nach einer FOD-Optimierung der globalen Zielfunktionen für die Impräzision. Die Berechnungen wurden mit der MATLAB Optimization-Toolbox durchgeführt. Zur Übersichtlichkeit sind nur die Intervallboxen dargestellt.

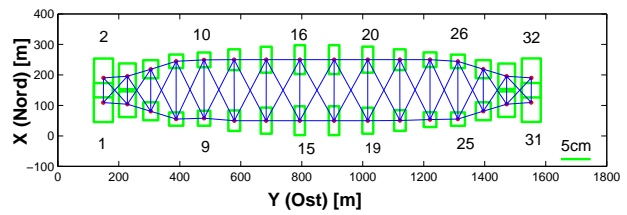
In der Teilabbildung 5.3(a) wird das Ausgangsdesign gezeigt. Die Intervallboxen zeigen ein „Aufschwingen“ in der Intervallmitte und an den Intervallrändern in Querrichtung. Es ist zu erwarten, dass durch die Optimierung diesem Verhalten entgegen gewirkt wird.

In der linken Spalte der Abbildung 5.3 sind die optimalen Netzkonfigurationen dargestellt. Diese ähnelt sich für alle drei Zielfunktionen. Die Variation der Punktlage weist zwei Haupteffekte auf: Zum Einen findet in Längsrichtung eine starke Kontraktion an den NetZRändern statt und eine Ausdehnung in der Netzmitte. Senkrecht dazu dehnt sich das Netz in der Netzmitte aus. Die maximalen Punktvariationsbereiche von  $\pm 50 \text{ m}$  werden dabei ausgeschöpft. Das Netz hat die Tendenz, eine flächenhaftere Form anzunehmen.

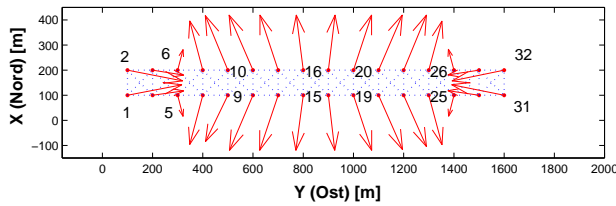
Die Punktbeziehung vom Ausgangsdesign zum optimalen Design ist nicht maßstäblich in der Teilabbildung 5.3(c) dargestellt. Die einzelnen Zielfunktionen unterscheiden sich dabei lediglich in der Größenordnung der Längskontraktion an den NetZRändern. Für dieses Beispiel gilt: Die Kontraktion ist bei Minimierung der Zielfunktion  $\Phi_2$  am schwächsten ausgeprägt und bei der Minimierung der Zielfunktion  $\Phi_\infty$  am stärksten.



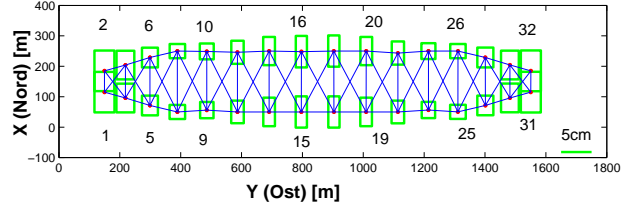
(a) A Priori Design



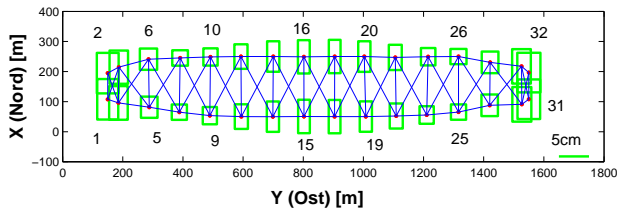
(b) Optimales Design bei Zielfunktion  $\Phi_2$



(c) Variation der Punktlage



(d) Optimales Design bei Zielfunktion  $\Phi_F$



(e) Optimales Design bei Zielfunktion  $\Phi_\infty$

Abbildung 5.3: Vergleich der Intervallboxen beim FOD-Design für globale Zielfunktionen der Impräzision am Beispiel einer Diagonalenviereckskette (Streckennetz)

Die folgende Tabelle gibt eine Übersicht über die Werte der einzelnen Zielfunktionen und ihrer Reduktion. Zum Vergleich sind Zielfunktionen für Kenngrößen der Genauigkeit und der Zuverlässigkeit mit angegeben.

	$\text{spur} \mathbf{Q}_{\hat{x}\hat{x}}$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{x}\hat{x}}\ _2$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{x}\hat{x}}\ _F$ [m <sup>2</sup> ]	$\ \mathbf{K}\ _F$ [mm]	$\ \mathbf{K}\ _2$ [mm]	$\ \mathbf{K}\ _\infty$ [mm]	$r_{max}$ [-]	$r_{min}$ [-]	$r_{max} - r_{min}$ [-]
a priori	2909.6	2211.3	2238.4	44.0	38.3	67.3	0.25	0.12	0.13
Optimierung von $\Phi_2$	889.9	490.1	511.1	24.5	18.1	32.2	0.34	0.04	0.30
Reduktion in [%]	70	78	77	44	52	52	-36	66	-131
Optimierung von $\Phi_F$	900.3	517.8	536.4	24.6	18.6	32	0.34	0.05	0.29
Reduktion in [%]	69	76	76	44	51	52	-36	61	-123
Optimierung von $\Phi_\infty$	956.0	565.4	583.0	25.4	19.4	28.7	0.34	0.01	0.33
Reduktion in [%]	67	74	73	42	49	57	-36	91	-153

Tabelle 5.2: Vergleich der Veränderung der Werte der Zielfunktionen bei der FOD-Optimierung einer Diagonalenviereckskette (Streckennetz) unter Verwendung dreier unterschiedlicher Zielfunktionen:  $\Phi_2$ ,  $\Phi_F$  und  $\Phi_\infty$ .

Die erste Spalte zeigt, welche Zielfunktion der Impräzision zur Berechnung des FOD verwendet wurde. In der zweiten bis vierten Spalte sind die Werte der Zielfunktionen für die Netzgenauigkeit angegeben, in der fünften bis siebten Spalte die Werte der Zielfunktionen der Impräzision. In den letzten drei Spalten sind der maximale und minimale Redundanzanteil als Maße der Zuverlässigkeit aufgelistet sowie die Spannweite zwischen maximalen und minimalem Redundanzanteil, die als Maß ihrer Homogenität angesehen werden kann. Die prozentuale

Reduktion berechnet sich nach:

$$\text{Reduktion}[\%] := - \frac{\text{Kenngröße a posteriori} - \text{Kenngröße a priori}}{\text{Kenngröße a priori}} 100. \quad (5.52)$$

Wird ein negativer Wert für die Reduktion erzielt, so ist der Wert der Kenngröße nach der Optimierung größer als beim a priori Design. Es lässt sich feststellen, dass sich bei einer FOD-Optimierung mit Zielfunktionen der Impräzision die Werte der Kenngrößen der Netzgenauigkeit um 60 – 75%, die Werte der Kenngrößen der Impräzision um 35 – 55% verbessern. Die Netzzuverlässigkeit wird hingegen verschlechtert.

Zum Abschluss dieses Beispiels soll auf die optimalen Netzkonfigurationen kurz eingegangen werden, die man bei der Optimierung des maximalen Eigenwertes der Kofaktormatrix erhält. KALTENBACH (1992, S.67ff) untersucht dazu eine Diagonalenviereckskette mit 26 Netzpunkten, JÄGER (1988, S.124) eine Diagonalenviereckskette mit 20 Netzpunkten. Es bleibt festzuhalten, dass die optimalen Netzkonfigurationen im FOD für die Zielfunktion der Impräzision denen ähneln, die mit spektralen Zielfunktionen, vgl. KALTENBACH (1992, S.71), erhalten werden. Dort lässt sich ebenfalls eine Netzkontraktion in Längsrichtung an den Netzenenden nachweisen und eine Längs- und Querausdehnung in der Netzmitte.

**Beispiel 29** Als zweites Beispiel betrachten wir ein flächenhaftes Netz, das aus gleichseitigen Dreiecken aufgebaut ist, die eine Seitenlänge von  $S = 100 \text{ m}$  haben. Das Design entspricht dem Netz, das von SCHMITT (1997) bezüglich spektraler Kriterien der klassischen Optimierung untersucht wurde. Die Intervallradien und Standardabweichungen der Streckenmessung sind im vorherigen Beispiel gegeben. Für die Richtungsbeobachtungen erhält man:  $R_r = 0.35 \text{ mgon}$  und  $\sigma_R = 0.5 \text{ mgon}$ .

Abbildung 5.4 zeigt die Zonotope und Intervallboxen des A-Priori-Designs. Ihre Größe nimmt mit wachsendem Abstand von der Netzmitte zu. In der Abbildung 5.5 sind die optimalen Netzkonfigurationen nach einer FOD-Optimierung der globalen Zielfunktionen für die Impräzision dargestellt. Die Berechnungen wurden mit der MATLAB Optimization-Toolbox durchgeführt. Es lassen sich leichte Abweichungen der optimierten Designs von einer vollständig symmetrischen Punktlage feststellen, die durch eine Beschränkung der Anzahl der Iterationsschritte bei der Optimierung bedingt sind.

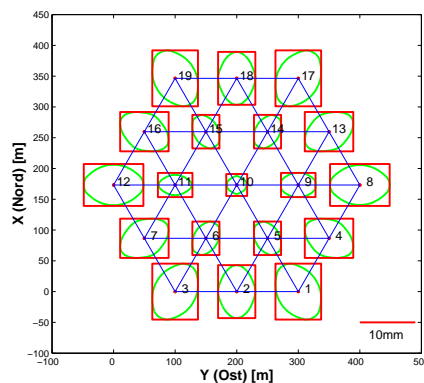


Abbildung 5.4: Intervallboxen und Zonotope des A-Priori-Design eines flächenhaften Netzes

In der oberen Zeile der Abbildung 5.5 sind die Punktlagevariationen qualitativ wiedergegeben. In der unteren Zeile der Abbildung sind jeweils die optimale Netzkonfiguration und die Zonotope und Intervallboxen angegeben. Teilabbildung 5.5(d) zeigt diese bei einer Optimierung bezüglich der Zielfunktion  $\Phi_2$ , Teilabbildung 5.5(e) bezüglich  $\Phi_F$  und Teilabbildung 5.5(e) bezüglich  $\Phi_\infty$ . Die Netzgesamtstruktur weist durch ihre hohe Symmetrie bereits gewisse optimale Eigenschaften auf. Für die Optimierung der Zielfunktion  $\|\mathbf{K}\|_2$  wird ein ähnliches optimales Design erreicht, wie SCHMITT (1997, Fig. 8) für die gezielte Reduktion um 10% des maximalen Eigenwertes der Varianz-Kovarianz-Matrix erhalten hat.

Der Übergang vom A-Priori-Design auf das optimierte Design wird durch zwei Mechanismen erreicht. Als dominante Punktbeziehung werden die inneren Netzpunkte zur Netzmitte kontrahiert, die selbst konstant bleibt. Die äußeren Netzpunkte verändern geringfügig ihre Lage auf dem Rand des Sechsecks zu einer sternförmigen Randlage. Da durch diese Variation der Punktlage der Einfluss der Hauptschwachform reduziert wird, wird die Stabilität des Netzes verbessert (SCHMITT 1997). Die Werte der Zielgrößen  $\|\mathbf{K}\|_F$  und  $\text{spur} \mathbf{Q}_{\hat{x}\hat{x}}$  erhöhen sich im Vergleich zum Ausgangsdesign, ebenso die Werte der Kenngrößen der Zuverlässigkeit, vgl. Tabelle 5.3.

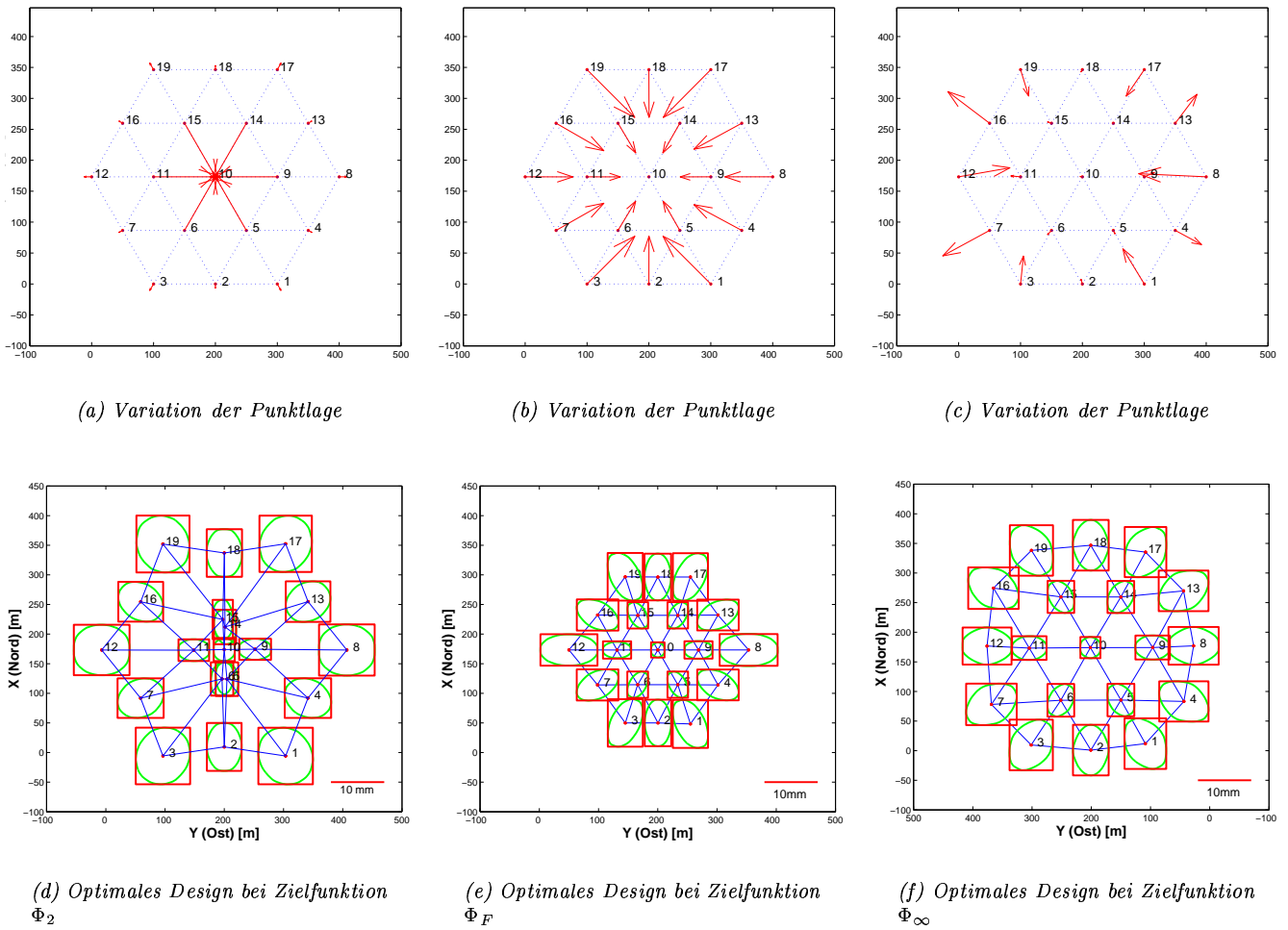


Abbildung 5.5: Vergleich der Intervallboxen und Zonotope der optimalen Netzkonfigurationen beim FOD-Design für ein regelmäßiges flächenhaftes Netz

	$\text{spur } \mathbf{Q}_{\hat{x}\hat{x}}$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{x}\hat{x}}\ _2$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{x}\hat{x}}\ _F$ [m <sup>2</sup> ]	$\ \mathbf{K}\ _F$ [mm]	$\ \mathbf{K}\ _2$ [mm]	$\ \mathbf{K}\ _\infty$ [mm]	$r_{max}$ [-]	$r_{min}$ [-]	$r_{max} - r_{min}$ [-]
a priori	15.1	4.6	5.4	2.9	1.7	5.4	0.91	0.32	0.59
Optimierung von $\Phi_2$	16.7	3.6	5.0	3.2	1.5	5.3	0.97	0.29	0.68
Reduktion in [%]	-10	21	7	-10	11	2	-6	9	-15
Optimierung von $\Phi_F$	10.4	4.4	4.7	2.4	1.6	4.4	0.94	0.27	0.67
Reduktion in [%]	31	4	12	17	5	18	-3	15	-13
Optimierung von $\Phi_\infty$	14.7	4.5	5.3	2.9	1.7	4.7	0.90	0.29	0.61
Reduktion in [%]	2	2	2	0	0	12	1	9	-3

Tabelle 5.3: Veränderung der Werte der Kenngrößen für Impräzision, Stochastizität und Zuverlässigkeit bei einem flächenhaften Netz unter Verwendung dreier unterschiedlicher Zielfunktionen:  $\Phi_2$ ,  $\Phi_F$  und  $\Phi_\infty$ . Berechnung der Reduktion nach (5.52).

Die Zielfunktion  $\Phi_F$  führt zu einer zentrischen Kontraktion des gesamten Netzes. Dieser Mechanismus lässt sich folgendermaßen erklären: Da alle Intervallradien direkt von der Streckenlänge abhängen, ist so eine starke Reduktion der Werte der Zielfunktionen zu erreichen. Die Netzstruktur wird dabei nicht verändert und weist sich somit als optimal für diese spektrale Zielgröße aus. Bei dieser Optimierung werden ebenfalls die Werte der Kenngrößen der Netzgenauigkeit reduziert, wohingegen das Spektrum der Redundanzanteile sich leicht verbreitert.

Bei der Optimierung mit der Zielfunktion  $\Phi_\infty$  wurde im Wesentlichen nur eine Minimierung des maximalen Intervallradius der Koordinaten erreicht. Die Werte aller anderen Zielfunktionen bleiben hingegen fast unverändert. Die Aussagen einer Optimierung mit der Zielfunktion  $\Phi_\infty$  gelten aber immer nur für den aktuell gewählten Koordinatenrahmen. In einem gedrehten Koordinatenrahmen sind andere Effekte zu erwarten, da sich die Größenverhältnisse der Intervallradien ändern.

### Einfluss der Netztopologie

Im Folgenden wird die Abhängigkeit der Werte der Kenngrößen bzw. Zielfunktionen von der Netztopologie untersucht. Erste Überlegungen zu diesem Themenkomplex wurden von SCHÖN UND KUTTERER (2001c) vorgestellt. Die Idee ist, anhand der Strategie der Netzerweiterung zu zeigen, wie das Hinzufügen redundanter Messungen die Werte der Kenngrößen der Impräzision, der Netzgenauigkeit und Zuverlässigkeit beeinflusst.

**Beispiel 30** Diese Effekte sollen an einem einfachen Netzbeispiel erläutert werden, das sich an das in SCHÖN UND KUTTERER (2001c) gezeigte Design anlehnt.

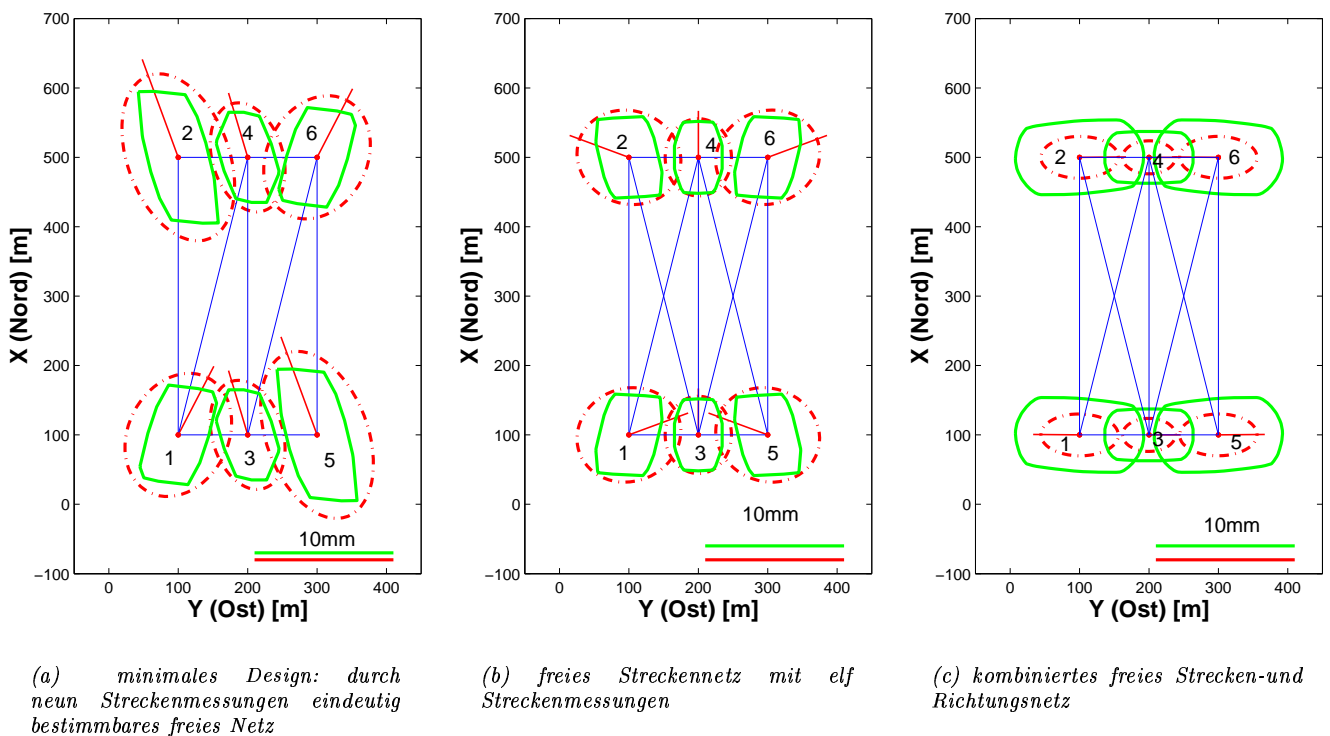


Abbildung 5.6: Vergleich der 95%-Konfidenzellipsen (gestrichelt) und Zonotope für drei Netzkonfigurationen eines Diagonalenvierecks

Abbildung 5.6 zeigt die 95%-Konfidenzellipsen und Zonotope für drei Messanordnungen einer Diagonalenviereckskette von sechs Punkten. Die Standardabweichungen der Messungen betragen  $\sigma_S = 2 \text{ mm} + 2 \text{ ppm}$  und  $\sigma_R = 0.5 \text{ mgon}$ . Die Intervallradien wurden nach den Faustformeln (3.2), (3.4) und (3.6) bestimmt und ergeben  $1.6 \text{ mm}$  bzw.  $0.3 \text{ mgon}$  für kurze und  $1.9 \text{ mm}$  bzw.  $0.7 \text{ mgon}$  für lange Strecken.

In der Teilabbildung 5.6(a) ist das minimale Design gezeigt, d.h. eine durch neun Streckenmessungen eindeutig bestimmte, freie Netzkonfiguration. Im ersten Schritt der Netzerweiterung wird das minimale Design um zwei Streckenmessungen ergänzt. Das sich dabei ergebende Netz wird in der Teilabbildung 5.6(b) gezeigt. Im zweiten Schritt der Netzerweiterung werden Richtungsmessungen hinzugefügt, vgl. SCHÖN UND KUTTERER (2001c). Teilabbildung 5.6(c) stellt die Zonotope und 95%-Konfidenzellipsen dieses dritten Netzdesigns dar.

Ein graphischer Vergleich des ersten und zweiten Designs zeigt, dass sowohl der Werte der Kenngrößen der Impräzision als auch der Netzgenauigkeit verbessert werden können. Zusätzlich kann die Zuverlässigkeit durch das Hinzufügen der zwei redundanten Strecken gesteigert werden, vgl. SCHÖN UND KUTTERER (2001c). Werden im

zweiten Schritt der Netzerweiterung noch Richtungsmessungen zum vollständigen Streckendesign hinzugenommen, so können die Konfidenzellipsen weiter verkleinert und die Zuverlässigkeit gesteigert werden. Die Werte der Kenngrößen der Impräzision der Punktpositionen nehmen hingegen wieder zu, vgl. Teilabbildung 5.6(c).

Allgemein gilt, dass durch die Hinzunahme von Messungen die Zuverlässigkeit gesteigert wird und die Werte der Kenngrößen der Stochastizität verringert werden. Grundlage für diese Verringerung ist eine Verbesserung der Bestimmungsgüte der Punktposition durch Überbestimmung: Übergang von Design 1 auf Design 2 bzw. auf Design 3. Für die Werte der Kenngrößen der Impräzision ist hingegen neben diesem ersten, ein zweiter gegenläufiger Mechanismus zu berücksichtigen. Bis zu einem gewissen Grad führt die Verbesserung der Bestimmungsgüte der Punktposition auch zu einem günstigeren Übertragungsverhältnis der Impräzision der Messungen und damit zu kleineren Werten der Kenngrößen der Impräzision für die Punktposition. Andererseits erhöht das Hinzufügen jeder Messung die Gesamtimpräzision, die auf die Koordinaten fortgepflanzt wird. Je nach Netzausdehnung bzw. Größe der Intervallradien der Messungen überwiegt der eine oder andere Mechanismus. Als Konsequenz existiert eine Netzkonfiguration, ab der das weitere Hinzufügen von Messungen die Werte der Kenngrößen der Impräzision verschlechtert. Diese „Grenzkonfiguration“ liefert die optimale Topologie bei einer gemeinsamen Betrachtung der Qualitätsmerkmale Netzgenauigkeit, Zuverlässigkeit und Impräzision.

### Einfluss des Gewichtung

Im Folgenden wird der Einfluss der Gewichtung auf die Werte der Kenngrößen der Impräzision untersucht. Im Gegensatz zu Kenngrößen der Stochastizität sind die Zielfunktionen der Impräzision invariant in Bezug auf eine Skalierung

$$\mathbf{P}'_{ll} := \mu \mathbf{P}_{ll}, \quad \mu > 0. \quad (5.53)$$

Die Matrix der Kantenvektoren  $\mathbf{K}$  und somit alle daraus abgeleiteten Kenngrößen sind von der Wahl des Gewichtniveaus unabhängig, was im hier kurz angedeutet werden soll.

- Betrachten wir zuerst angeschlossene Netze. Hier ergibt sich die Matrix der Kantenvektoren zu:

$$\mathbf{K}(\mathbf{P}'_{ll}) = (\bar{\mathbf{A}}^T \mathbf{P}'_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}} \mathbf{P}'_{ll} \mathbf{L}_r. \quad (5.54)$$

Die Matrix  $\bar{\mathbf{A}} = \mathbf{A}_K - \mathbf{A}_Z (\mathbf{A}_Z^T \mathbf{P}_{ll} \mathbf{A}_Z)^{-1} \mathbf{A}_Z^T \mathbf{P}_{ll} \mathbf{A}_K$  ist invariant gegen eine Transformation vom Typ (5.53) und damit auch das Matrixprodukt (5.54).

- Für freie Netzausgleichungen lässt sich die Kofaktor-Matrix der ausgeglichenen Koordinaten in der folgenden Form darstellen (KOCH 1997, S.65):

$$\mathbf{Q}_{\hat{x}\hat{x}} = (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}} + \mathbf{D}\mathbf{D}^T)^{-1} - \mathbf{G}(\mathbf{G}^T \mathbf{D}\mathbf{D}^T \mathbf{G})^{-1} \mathbf{G}^T. \quad (5.55)$$

Da eine Skalierung des Nullraums keinen Einfluss hat, erhält man mit  $\bar{\mathbf{A}}\mathbf{G} = \mathbf{0}$  für die Matrix der Kantenvektoren:

$$\begin{aligned} \mathbf{K}(\mathbf{P}'_{ll}) &= (\bar{\mathbf{A}}^T \mathbf{P}'_{ll} \bar{\mathbf{A}} + \mathbf{D}\mathbf{D}^T)^{-1} \bar{\mathbf{A}} \mathbf{P}'_{ll} \mathbf{L}_r \\ &= (\mu \bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}} + \mu \mathbf{D}\mathbf{D}^T)^{-1} \bar{\mathbf{A}} \mu \mathbf{P}_{ll} \mathbf{L}_r \\ &= (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}} + \mathbf{D}\mathbf{D}^T)^{-1} \bar{\mathbf{A}} \mathbf{P}_{ll} \mathbf{L}_r = \mathbf{K}(\mathbf{P}_{ll}). \end{aligned} \quad (5.56)$$

Die Matrix der Kantenvektoren und daraus abgeleitete Größen sind somit invariant gegen skalare Transformation (5.53). Die Beantwortung der Ausgangsfragestellung nach dem Einfluss der Gewichtung zielt also nicht auf die Festlegung eines optimalen absoluten Wertes für jedes Gewicht, sondern vielmehr auf die Bestimmung eines optimalen Verhältnisses zwischen Gewichten unterschiedlicher Messungstypen.

Eine erste Lösung für dieser Frage nach einem optimalen Verhältnis der Gewichte wurde von SCHÖN UND KUTTERER (2001b) vorgestellt. Die wichtigsten Ergebnisse sollen hier kurz skizziert werden. Zielsetzung der Untersuchungen ist, ein optimales Gesichtsverhältnis zwischen den einzelnen Messungstypen (Richtungen, Strecken, Zenitdistanzen) zu finden und dieses in einem zweiten Schritt auf konkrete Werte für verschiedene Instrumente zurückzuführen.

Da sich die Gewichte für Messungen, die mit modernen Instrumenten gemessen wurden, für jeden Messungstyp (Richtung, Strecke oder Zenitdistanz) je einheitlich funktional durch die Sensor-Performance-Parameter  $a$ ,  $b$ ,  $q_{r_i}$  und  $q_\zeta$  beschreiben lassen, wurden diese als variable Parameter der Optimierung eingeführt. Als Randbedingungen wurden maximale Variationsbereiche der Performance-Parameter vorgegeben, die sich beispielsweise



aus ihren Werten für Instrumente unterschiedlicher Instrumentenklassen angeben lassen. Die Lösung des SOD-Problems ergibt optimale Werte dieser Parameter, vgl. SCHÖN UND KUTTERER (2001b).

Existiert kein Instrumentarium, das direkt durch die optimalen Performance-Parameter charakterisiert ist, so können in einem zweiten Schritt durch gemeinsame Skalierung der optimalen Performance-Parameter Werte bestimmt werden, die mit Instrumentarien umgesetzt werden können. Dies ist möglich, da einerseits die Zielfunktionen unabhängig von einer Skalierung der Gewichte sind, vgl. (5.56) und andererseits die Gewichte (alle) quadratisch von den Performance-Parametern abhängen.

### 5.3.3 Zusammenfassung und Wertung

Die Beispiele in den vorherigen Abschnitten haben die Reaktion der Netze auf Optimalitätsforderungen bezüglich Zielfunktionen der Impräzision, Netzgenauigkeit und Zuverlässigkeit gezeigt. Hierbei wurden jeweils unterschiedliche Parameter (Datumsgebung, Netzgeometrie, Netztopologie oder Gewichtung der Messungen) variiert. Dabei sind zwei wesentliche Eigenschaften besonders aufgefallen. Sie sollen in diesem Abschnitt näher diskutiert werden: Zum Einen konnte eine *Ähnlichkeit* festgestellt werden, mit der die Werte der Kenngrößen der Impräzision und Stochastizität auf Variationen der Parameter reagieren. Zum Anderen scheint die Behandlung der Impräzision im Kontext der klassischen Designs (FOD, SOD) eine *größere Komplexität* aufzuweisen als die Betrachtung klassischer Zielfunktionen.

#### Ähnlichkeit der Reaktion von Impräzision und Stochastizität bei der Netzoptimierung

Bei der Optimierung der in den vorherigen Abschnitten betrachteten Netzbeispielen haben sich Ähnlichkeiten im Reduktionsverhalten der Kenngrößen der Impräzision und Stochastizität der Koordinaten gezeigt. Im Folgenden soll deshalb untersucht werden, wie diese Ähnlichkeit mathematisch beschrieben und gefasst werden kann.

Ausgangspunkt für die Überlegungen ist ein Vergleich der spektralen Zielfunktionen für die Impräzision und Stochastizität. Betrachten wir dazu die Singulärwertzerlegung des folgenden Matrixproduktes:

$$\bar{\mathbf{A}}^T \mathbf{P}_{\bar{u}}^{\frac{1}{2}} = \mathbf{U}^* \mathbf{S}^* \mathbf{V}^{*T}, \quad (5.57)$$

mit einer  $u \times u$  orthonormierten Matrix  $\mathbf{U}^*$  der Linkseigenvektoren, einer  $n \times n$  orthonormierten Matrix  $\mathbf{V}^*$  der Rechtseigenvektoren und einer  $u \times n$  Matrix  $\mathbf{S}^*$  der Singulärwerte. Für  $\mathbf{S}^*$  gilt:

$$\mathbf{S}^* = (\mathbf{S}_1^*, \mathbf{0}), \quad \mathbf{S}^{*+} := (\mathbf{S}_1^{*+}, \mathbf{0})^T, \quad (5.58)$$

wobei  $\mathbf{S}_1^*$  die  $u \times u$  Diagonalmatrix bezeichnet, deren Diagonalelemente die Singulärwerte sind. Mit (5.58) lässt sich die Eigenwertzerlegung der Kofaktormatrix der ausgeglichenen Koordinaten bei pseudoinverser Lösung in einer zweiten Darstellung angeben:

$$\mathbf{Q}_{\bar{x}\bar{x}} = \mathbf{U}^* (\mathbf{S}^* \mathbf{S}^{*T})^+ \mathbf{U}^{*T} = \mathbf{M} \mathbf{\Lambda} \mathbf{M}^T, \quad (5.59)$$

wobei  $\mathbf{M} = \mathbf{U}^*$  die  $u \times u$  orthonormale Matrix der Eigenvektoren ist und  $\mathbf{\Lambda} = (\mathbf{S}^* \mathbf{S}^{*T})^+$  die  $u \times u$  Diagonalmatrix der Eigenwerte. Die Matrix der Kantenvektoren lässt sich folgendermaßen darstellen:

$$\begin{aligned} \mathbf{K} &= \mathbf{U}^* (\mathbf{S}^* \mathbf{S}^{*T})^+ \mathbf{S}^* \mathbf{V}^{*T} \mathbf{P}_{\bar{u}}^{\frac{1}{2}} \mathbf{L}_r \\ &= \mathbf{U}^* (\mathbf{S}^{*+})^T \mathbf{V}^{*T} \mathbf{P}_{\bar{u}}^{\frac{1}{2}} \mathbf{L}_r. \end{aligned} \quad (5.60)$$

Ein Schlüssel zur Lösung der Aufgabenstellung ist ein Vergleich der Gleichungen (5.59) und (5.60).

- Ist das Verhältnis zwischen der Impräzision, die durch die Intervallradien angegeben ist und der Gewichtung für alle Beobachtungen gleich groß, so gilt:

$$\mathbf{P}_{\bar{u}}^{\frac{1}{2}} \mathbf{L}_r = \eta \mathbf{I}, \quad \eta > 0. \quad (5.61)$$

Mit (5.61) lässt sich die Matrix der Kantenvektoren direkt in der Basis  $\mathbf{U}^*$  der Kofaktormatrix ausdrücken:

$$\mathbf{K} = \eta \mathbf{U}^* (\mathbf{S}^{*+})^T \mathbf{V}^{*T} \quad (5.62)$$

Die Spaltenvektoren der Matrix der Linkseigenwerte  $\mathbf{U}^* = \mathbf{M}$  geben an, welche Linearkombinationen des Koordinatenvektors am schlechtesten bestimmt sind. Da sowohl die Varianz-Kovarianz-Matrix als auch

die Matrix der Kantenvektoren in dieser Basis  $\mathbf{U}^*$  darstellbar sind, sind diese Linearkombinationen für beide Anteile der Gesamtunsicherheit identisch. Die Spaltenvektoren von  $\mathbf{U}^*$  zeigen somit in die Richtung, in der das Netz die größte Gesamtunsicherheit aufweist. Für die Interpretation ist zu beachten, dass der jeweilige Einfluss durch die Größe der Eigenwerte bzw. Singulärwerte gewichtet wird. Die Ausprägung von Impräzision und Stochastizität unterscheidet sich dabei durch den Faktor  $\eta$ : Die Singulärwerte aus (5.62) lassen sich mit den Eigenwerten  $\lambda_i$  der Kofaktormatrix ausdrücken:

$$s_i(\mathbf{K}) = \eta\sqrt{\lambda_i}. \quad (5.63)$$

Da durch die Monotonie der Wurzelfunktion in (5.63) die Reihenfolge der Eigenwerte erhalten bleibt, lässt sich eine Proportionalität zwischen Zielfunktionen der Impräzision und der Stochastizität angeben:

$$\begin{aligned} \|\mathbf{K}\|_F &= \sqrt{\sum_{i=1}^u s_i^2} = \eta\sqrt{\text{spur}\mathbf{Q}_{\hat{x}\hat{x}}}, \\ \|\mathbf{K}\|_2 &= s_1 = \eta\sqrt{\|\mathbf{Q}_{\hat{x}\hat{x}}\|_2}, \end{aligned} \quad (5.64)$$

wobei  $s_1$  den maximalen Singulärwert bezeichnet. Für die Netzoptimierung bedeutet dies, dass eine Minimierung der Eigenwerte von  $\mathbf{Q}_{\hat{x}\hat{x}}$  äquivalent zur Minimierung der Singulärwerte von  $\mathbf{K}$  ist, wenn die Relation (5.60) bei der Optimierung durchweg erhalten bleibt. Eine Netzkonfiguration, die optimal bezüglich der in (5.64) genannten Zielfunktionen der Stochastizität ist, ist auch optimal bezüglich der entsprechenden Kriterien der Impräzision.

KUTTERER (1994, S.33ff) hat das Verhältnis von Kenngrößen der Impräzision und Stochastizität für das Beispiel eines angeschlossenen Höhennetzes dargestellt. Allerdings beschränkt er sich auf eine Diskussion der Intervallradien, für die keine funktionale Verbindung zu Kenngrößen der Stochastizität wie in (5.63) hergestellt werden kann.

- Ist Gleichung (5.61) nicht erfüllt, so kann die Ähnlichkeit dennoch abgeschätzt werden. Dabei soll der Einfluss der Matrix  $\mathbf{P}_{ii}^{\frac{1}{2}}\mathbf{L}_r$  geeignet auf die Rotationsmatrix  $\mathbf{U}^*$  und die Matrix  $\mathbf{S}^*$  aufzuteilen. Dies kann erreicht werden, indem die Singulärwertzerlegung der Matrix  $\mathbf{K} = \mathbf{USV}^T$  nach (4.89) auf die Basis  $\mathbf{U}^*$  transformiert wird. Mit (5.60) bzw. (4.89) ergeben sich die beiden folgenden Darstellung des Produktes  $\mathbf{U}^{*T}\mathbf{K}$ :

$$\begin{aligned} \mathbf{U}^{*T}\mathbf{K} &= (\mathbf{S}^{*+})^T\mathbf{V}^{*T}\mathbf{P}_{ii}^{\frac{1}{2}}\mathbf{L}_r \\ &= \mathbf{U}^{*T}\mathbf{USV}^T \end{aligned} \quad (5.65)$$

wobei  $\mathbf{S}$  die Matrix der Singulärwerte und  $\mathbf{V}$  die Matrix der Rechtseigenwerte der Matrix der Kantenvektoren sind, vgl. (4.89). Mit der Abkürzung

$$\mathbf{U}^{**} := \mathbf{U}^{*T}\mathbf{U} \quad (5.66)$$

kann die Matrix der Kantenvektoren in der Basis der Kofaktormatrix  $\mathbf{U}^*$  angegeben werden:

$$\mathbf{K} = \mathbf{U}^*\mathbf{U}^{**}\mathbf{SV}^T. \quad (5.67)$$

Die orthonormierte  $u \times u$  Matrix  $\mathbf{U}^{**}$  beschreibt die Rotation der Basis der Kofaktormatrix  $\mathbf{U}^*$  auf die Basis der Matrix der Kantenvektoren  $\mathbf{U}$ . Dadurch kann die Veränderung der Orientierung und Amplituden, die die spektralen Kenngrößen der Matrix  $\mathbf{K}$  im allgemeinen Fall im Vergleich zur Basis  $\mathbf{U}^*$  der Varianz-Kovarianz-Matrix der ausgeglichenen Koordinaten besitzen, abgeleitet werden. Die Singulärwerte von  $\mathbf{K}$  geben die Magnituden in der gedrehten Basis an.

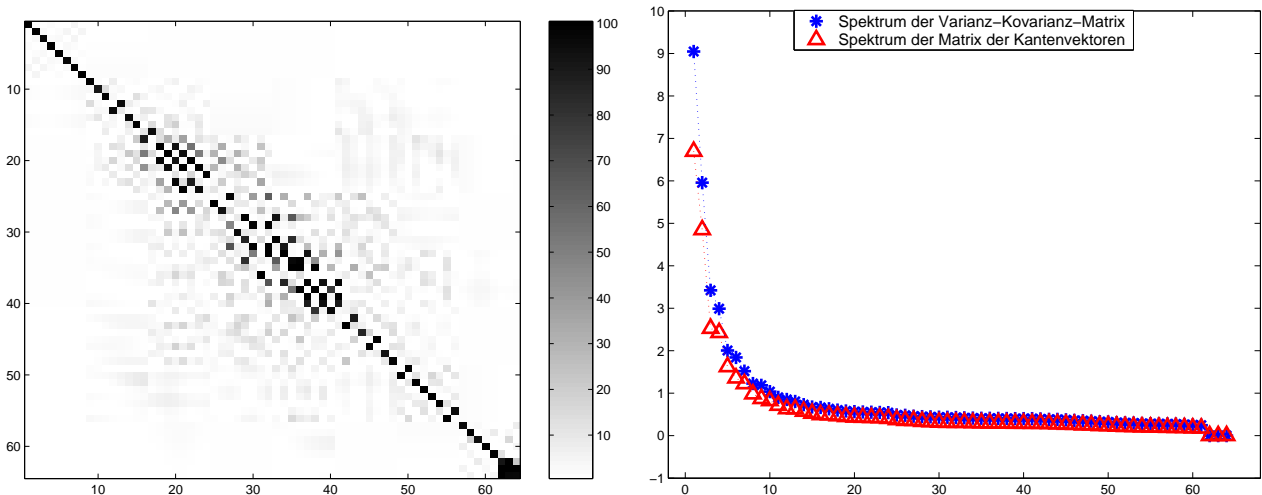
Je stärker die Matrix  $\mathbf{U}^{**}$  zur Einheitsmatrix tendiert, desto ähnlicher verhalten sich alle Hauptrichtungen von Stochastizität und Impräzision. Dieses Kriterium fordert allerdings, dass einerseits die Reihenfolge der Eigenvektoren eingehalten werden muss. Andererseits spielt die Gewichtung der einzelnen Eigenvektoren (Richtungen) mit den entsprechenden der Größe nach sortierten Eigenwerten bzw. Singulärwerten eine entscheidende Rolle. Existieren Eigenwerte, deren Beitrag zum gesamten Spektrum dieses dominiert, so kann in Analogie zu Hauptschwachformen die Diskussion auf die ersten, den größten Eigenwerten zugeordneten Eigenvektoren beschränkt werden. Liegen mehrfache Eigenwerte vor, so hat eine Vertauschung der entsprechenden Eigenvektoren keinen Einfluss auf die Interpretation.

Soll umgekehrt die Basis der Varianz-Kovarianz-Matrix exakt beibehalten werden, so können die Singulärwerte von  $\mathbf{K}$  mit der Matrix  $\mathbf{U}^{**}$  auf diese Hauptrichtungen transformiert (projiziert) werden.

**Beispiel 31** Betrachten wir als erläuterndes Beispiel die um Richtungsmessungen erweiterte Diagonalenvier-eckskette aus Beispiel 28, die als freies Netz im inneren Datum ausgeglichen wurde.

Die Teilabbildung 5.7(a) zeigt eine Grauwertdarstellung der Rotationsmatrix  $U^{**}$ . Je dunkler ein Element ist, desto näher liegt sein Betrag an der Eins. Die Teilabbildung 5.7(b) stellt die Wurzel der Eigenwerte der Varianz-Kovarianz-Matrix den Singulärwerten der Matrix der Kantenvektoren gegenüber. Das Spektrum der Kofaktormatrix wird von den ersten beiden Eigenwerten dominiert, die 50% bzw. 22% des Spektrums ausmachen. Für das Spektrum der Matrix der Kantenvektoren zeigt sich eine homogenere Situation. Hier trägt der erste Singulärwert 16.5% der Information, der zweite 12%.

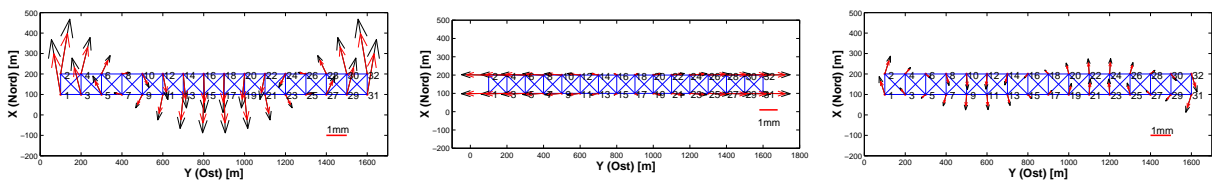
Die graphische Darstellung 5.7(a) zeigt eine dominante Diagonalstruktur. An diesen Stellen entsprechen sich die Eigenvektoren der Basis der Kofaktormatrix und die der Matrix der Kantenvektoren. Richtungen maximaler Impräzision fallen mit Richtungen minimaler Netzgenauigkeit zusammen. Diese Eigenvektoren geben somit Richtungen an, in denen die Gesamtunsicherheit groß ist. In der Mitte der Matrix  $U^{**}$  bilden sich „Cluster“ aus, und die Diagonalstruktur wird aufgebrochen. Dies bedeutet, dass der 20. bis 40. Eigenvektor der Basis der Kofaktormatrix durch eine Linearkombination der entsprechenden Eigenvektoren der Matrix der Kantenvektoren darstellbar ist. Prinzipiell ergeben sich dadurch andere Richtungen für die Auswirkung der Impräzision als für die Stochastizität. Bei dieser Interpretation muss allerdings beachtet werden, dass die entsprechenden Eigenvektoren mit den Singulärwerten bzw. mit der Wurzel der Eigenwerte skaliert werden. Wenn diese Eigen- bzw. Singulärwerte im Vergleich zum ersten Eigen- bzw. Singulärwert klein sind, spielen diese Variationen eine untergeordnete Rolle.



(a) Grauwertdarstellung der Rotationsmatrix  $U^{**}$

(b) Darstellung des Spektrums

Abbildung 5.7: Darstellung der Ähnlichkeit zwischen spektralen Kenngrößen für Impräzision und Stochastizität am Beispiel einer Diagonalenvier-eckskette



(a) 1. Eigenvektor

(b) 2. Eigenvektor

(c) 3. Eigenvektor

Abbildung 5.8: Vergleich der ersten drei Schwachformen für die Diagonalenvier-eckskette

Die Abbildung 5.8 zeigt die ersten drei Schwachformen für die Diagonalenvier-eckskette, die sich aus dem Produkt der ersten Eigenvektoren mit der Wurzel der entsprechenden Eigenwerten ergeben. In allen drei Fällen

stimmen die Richtungen für die spektralen Kenngrößen der Impräzision mit denen der Stochastizität überein. Die Teilabbildungen 5.8(a) und (c) zeigen ein „Ausschwingen“ der Netzform an den Netzen und in der Netzmitte senkrecht zur Netzausdehnung. Teilabbildung 5.8(b) stellt die durch den zweiten Eigenvektor generierte Längsausdehnung dar. Bei der Netzoptimierung im FOD wird versucht, diese Ausprägungen (Schwachformen) zu homogenisieren, indem beispielsweise der Wert des ersten Eigenwerts minimiert wird. Dadurch soll eine bessere Stabilität der Netzform erreicht werden. Es sollen keine ausgeprägten Richtungen mehr existieren, in denen das Netz am schlechtesten bestimmt ist, d.h. die größte Gesamtunsicherheit aufweist, vgl. beispielsweise die Diskussion in BILL ET AL. (1984) oder JÄGER UND KALTENBACH (1990).

### Komplexität der Behandlung der Impräzision

Im folgenden Abschnitt soll erklärt werden, warum die Behandlung der Impräzision auf den ersten Blick komplexer scheinen mag als die Betrachtung klassischer Zielgrößen. Ein kritischer Überblick der klassischen Designklassen zeigt auf Grenzen dieser Klassifizierung auf und legt Lösungskonzepte nahe, die auf eine klare Strukturierung zielen.

Die Beispiele in Abschnitt 5.3.2 haben gezeigt, dass bei der Betrachtung der Zielfunktionen für die Impräzision vor dem Hintergrund der klassischen Designs scheinbar vielschichtiger Problemstellungen auftreten, als man dies insbesondere vom Second Order Design gewohnt sein mag. Ein Grund hierfür liegt in der komplexeren Struktur der Intervallradien. Im Vergleich zu unabhängigen Gewichten, die beim SOD betrachtet werden, weisen die Intervallradien über die Streckenlänge und die Zenitdistanz per Definition eine besonders ausgeprägte Abhängigkeit von der Netzkonfiguration auf, vgl. die Faustformeln (3.2), (3.4) und (3.6).

Historisch gesehen bieten sich die „Grundbausteine“ eines geodätischen Netzes als freie Parameter in der Netzoptimierung an: Netztopologie, Netzgeometrie, Netzdatum und Gewichtung. Sie erweisen sich aber bei einer strengen Betrachtung als abhängige Parameter. Je nach Ansatz verändert beispielsweise eine Variation der Punktpositionen nicht nur die Werte der Messelemente zwischen den Punkten, sondern auch die Gewichtung, wenn ein Ansatz eingeführt wird, der die Abhängigkeit der Gewichte von der Streckenlängen berücksichtigt. Ebenso verändern sich die Werte der Matrix  $\mathbf{G}$ , die den Nullraum der aktuellen Varianz-Kovarianz-Matrix der Koordinaten beschreibt. Als Folge können direkte und indirekte Effekte entstehen, die die Problemstellung der Optimierung verkomplizieren. Diese Komplexität weist auf Grenzen der strikten Einteilung in die verschiedenen klassischen Designs hin.

Die Vermengung von direkten und indirekten Effekten kann umgangen werden, wenn die Fragestellungen so formuliert werden, dass sie durch unabhängige Parameter gelöst werden können. Die Einführung der Impräzision als neue Zielgröße legt dieses Vorgehen nahe. Die Impräzision hängt einerseits explizit vom Netzdesign (Streckenlängen bzw. Zenitdistanzen) als auch vom verwendeten Instrumentarium und den Mess- und Auswertemethoden ab. Als unabhängige Parameter bieten sich deshalb einerseits die Koordinaten der Netzknoten an. Andererseits müssen die Performance des verwendeten Instrumentariums und die angewandten Mess- und Auswertemethoden geeignet parametrisiert werden. Die Beobachtungsgewichte lassen sich auf die Performance-Parameter der Sensoren zurückführen. Sie beschreiben die stochastischen Eigenschaften der mit diesen Sensoren gemessenen Beobachtungen. Die Intervallradien hängen ebenfalls von den verwendeten Sensoren ab, vgl. Kapitel 3, aber auch von den Mess- und Auswertemethoden. Als eine gemeinsame Parametrisierung und Quantifizierung der unterschiedlichen Einflüsse sind die *Koeffizienten der Faustformeln* für die Intervallradien anzusehen, vgl. (3.7) in Abschnitt 3.3.

Zusammenfassend lässt sich mit diesen Überlegungen ein Satz unabhängiger Parameter oder „Grundbausteine“ für die geodätische Netzoptimierung angeben: *Netztopologie, Punktkoordinaten, Performance-Parameter der Sensoren und entsprechende Koeffizienten der Faustformeln*. Diese Parameterwahl erlaubt eine klare Strukturierung ohne indirekte Effekte. Die Netzkoordinaten wären dabei einem neuen „Konfigurationsdesign“ zugeordnet, die Performance-Parameter und entsprechenden Koeffizienten der Faustformeln einem „Unsicherheit-Design“.

## 5.4 Optimierung der Impräzision: indirekte Fragestellung

In diesem Abschnitt soll der indirekten Fragestellung nachgegangen werden: Wie müssen Instrumentarien, Messmethoden und Auswerteverfahren gewählt werden, damit Vorgaben an die Kenngrößen der Impräzision nicht überschritten werden?

In Kapitel 2 wurde gezeigt, dass Impräzision durch Intervalle geeignet beschrieben werden kann und somit systematisch wirkende Effekte mathematisch handhabbar werden. Für viele Anwendungen wäre es vereinfachend,

wenn diese Effekte möglichst so klein wären, dass sie vernachlässigbar sind. Im Folgenden soll beispielhaft untersucht werden, inwieweit durch geschickte Wahl der Mess- und Auswertemethoden und Instrumentarien eine solche Reduktion der Impräzision erreicht werden kann. Diese Aufgabenstellung soll als *Intervalldesign* bezeichnet werden. Schranken, ab denen die Impräzision vernachlässigbar ist, sollen mit Hilfe der stochastischen Kenngrößen der Punktpositionen, beispielsweise durch 95 %-Konfidenzbereiche, ausgedrückt werden.

### 5.4.1 Idee und Zielsetzung

Das Ziel der folgenden Untersuchungen ist, bei vorgegebenen oberen Schranken für die tolerierbare Impräzision  $\hat{x}_{r,i,soll}$  maximale Werte für die Intervallradien  $\overline{l_{r,j}}$  der Messungen zu finden. Ein erster Ansatz zur Lösung dieser Fragestellung ist eine unabhängige Variation der Intervallradien aller Messungen. Dieses Vorgehen ist eher von theoretischer Bedeutung, da die Größe der Intervallradien der Messungen funktional von den Einflussparametern abhängt, vgl. Gleichung (2.71). Wegen dieser Abhängigkeit ist keine beliebige Variation möglich. Vereinfacht lässt sich dieser Zusammenhang durch die Faustformeln, vgl. (3.2), (3.4) oder (3.6) beschreiben. In der vorliegenden Arbeit wird deshalb exemplarisch für die terrestrischen Messungstypen Strecke, Richtung und Zenitdistanz untersucht, wie diese Koeffizienten geeignet gewählt werden können, um die Zielerfordernisse zu erfüllen. Diese Überlegungen lassen sich auf alle anderen Messungstypen und Aufgabenstellungen übertragen.

Die Koeffizienten der Faustformeln, vgl. Abschnitt 3.3, sollen als freie Parameter einer Optimierungsaufgabe angesehen werden. Die Intervallradien der Koordinaten lassen sich in Abhängigkeit dieser Koeffizienten bei spaltenregulärer reduzierter Designmatrix  $\bar{\mathbf{A}}$  durch folgenden linearen Zusammenhang beschreiben:

$$\hat{\mathbf{x}}_r = |(\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} | \mathbf{L} \boldsymbol{\kappa} \quad (5.68)$$

mit der nicht-negativen  $n \times n_\kappa$  Matrix  $\mathbf{L}$  und dem  $n_\kappa \times 1$  Vektor  $\boldsymbol{\kappa}$  der Koeffizienten, vgl. Abschnitt 3.3, der in folgender Reihenfolge (Strecken - Richtungen - Zenitdistanzen) besetzt ist:

$$\boldsymbol{\kappa} = \left( a_S [mm] \quad c_S [ppm], \quad a_R [mgon], \quad b_R [mgon\ m], \quad c_R \left[ \frac{10^{-3} mgon}{m} \right], \quad d_R [mgon], \quad \dots \right. \\ \left. \dots, \quad a_\zeta [mgon], \quad b_\zeta [mgon\ m], \quad c_\zeta \left[ \frac{10^{-3} mgon}{m} \right] \right)^T. \quad (5.69)$$

In (5.69) sind die entsprechenden Einheiten der Koeffizienten angegeben. Eine Zeile der Matrix  $\mathbf{L}$  lässt sich für den zweidimensionalen Fall darstellen:

$$\mathbf{l}^i = \left( \sin \zeta_i, \quad \sin \zeta_i S_i, \quad 1, \quad \frac{1}{S_i}, \quad S_i \quad | \quad \cos \zeta_i, \quad \frac{|\cos \zeta_i| S_i}{\rho}, \quad \frac{|\cos \zeta_i|}{\rho}, \quad \frac{|\cos \zeta_i| S_i^2}{\rho} \right) \\ \text{bzw. im Dreidimensionalen:} \quad (5.70)$$

$$\mathbf{l}^i = \left( 1, \quad S_i, \quad 1, \quad \frac{1}{S_i}, \quad S_i, \quad | \quad \cot \zeta_i, \quad 1, \quad \frac{1}{S_i} \quad S_i \right).$$

Durch die (approximative) Beschreibung der Intervallradien der Messungen

$$\mathbf{l}_r = \mathbf{L} \boldsymbol{\kappa} \quad (5.71)$$

ist direkt eine Trennung in einen *Geometrieanteil*, der durch die Matrix  $\mathbf{L}$  ausgedrückt wird und in einen Anteil  $\boldsymbol{\kappa}$ , der von den verwendeten Instrumentarien sowie den Mess- und Auswertemethoden abhängt, möglich.

Die Lösungsstrategie ist zweistufig: Die erste Stufe besteht aus dem *Lösbarkeitstest* des Intervalldesigns, der im Detail in Abschnitt 5.4.3 beschrieben wird. Schranken für die Lösbarkeit sind einerseits die obere Grenzen für die Intervallradien der Koordinaten, die als Zielvorgaben im nächsten Abschnitt entwickelt werden. Andererseits lassen sich die Koeffizienten der Faustformeln nicht beliebig reduzieren. Aus der Auflösbarkeit der Sensoren und der Repräsentativität der verwendeten Korrekptionsmodelle ergeben sich minimale Werte  $\underline{\boldsymbol{\kappa}}$ , die nicht unterschritten werden können. Wie diese Werte abgeschätzt werden können, wurde in Abschnitt 3.3 gezeigt.

In der zweiten Stufe ergeben sich folgende Optimierungsaufgaben: Ist der Lösbarkeitstest erfüllt, so können größere Koeffizientenwerte als die Minimalwerte  $\underline{\boldsymbol{\kappa}}$  zugelassen werden. Die Wahl dieser Werte soll durch eine Optimierungsaufgabe gelöst werden. Als Ergebnis erhält man optimale Koeffizienten der Faustformeln  $\boldsymbol{\kappa}^*$ . Diese implizieren Empfehlungen für die zu verwendenden Instrumente sowie Mess- und Auswertemethoden. Ist der Lösbarkeitstest nicht erfüllt, kann durch eine Variation der Netzgeometrie (FOD) versucht werden, die Netzkonfiguration so zu verbessern, dass der Lösbarkeitstest wenigstens mit den Minimalwerten  $\underline{\boldsymbol{\kappa}}$  erfüllt ist. Ist dies ebenfalls nicht möglich, so ist die Impräzision nicht vernachlässigbar.

Lösbarkeitstest: $ (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll}   \mathbf{L} \underline{\boldsymbol{\kappa}} \leq \hat{\mathbf{x}}_{r,soll}$		
lösbar	<ul style="list-style-type: none"> <li>▶ Bestimmung von <math>\bar{\boldsymbol{\kappa}}</math></li> <li>▶ Stufe II: optimale Wahl der Koeffizienten <math>\boldsymbol{\kappa}^*</math></li> </ul>	Impräzision vernachlässigbar bei Realisierung von $\boldsymbol{\kappa}^*$
nicht lösbar	<ul style="list-style-type: none"> <li>▶ FOD: Ziel: ein Design zu finden, das wenigstens für <math>\underline{\boldsymbol{\kappa}}</math> den Lösbarkeitstest erfüllt</li> </ul>	<div style="display: flex; flex-direction: column; align-items: center;"> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>lösbar</span> <span>Impräzision vernachlässigbar bei Realisierung von <math>\underline{\boldsymbol{\kappa}}</math></span> </div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>nicht lösbar</span> <span>Impräzision nicht vernachlässigbar</span> </div> </div>

Tabelle 5.4: Lösungsstrategie des Intervalldesigns

### 5.4.2 Formulierung von Zielvorgaben

Für die Untersuchungen wird gefordert, dass die Impräzision im Vergleich zur Stochastizität „klein“ sein soll, damit sie vernachlässigbar ist. Für die Formulierung bieten sich vier Alternativen an:

#### Individuelle Anpassung auf Basis der Standardabweichung

$$\hat{x}_{r,i} \leq \frac{\sigma_0}{f} \sqrt{\chi_{2,1-\gamma}^2(\mathbf{Q}_{\hat{x}\hat{x}})_{ii}} := \hat{x}_{r,i,soll}, \quad \text{für } i = 1, \dots, u$$

bzw. im Dreidimensionalen (5.72)

$$\hat{x}_{r,i} \leq \frac{\sigma_0}{f} \sqrt{\chi_{3,1-\gamma}^2(\mathbf{Q}_{\hat{x}\hat{x}})_{ii}} := \hat{x}_{r,i,soll}, \quad \text{für } i = 1, \dots, u,$$

wobei  $f$  das Verhältnis von Stochastizität zu Impräzision angibt und  $1 - \gamma$  die Konfidenzwahrscheinlichkeit. Der Vergleich wird auf Ebene der Koordinaten bzw. punktweise durchgeführt. Dabei wird die Größe der Intervallbox mit der Größe der Box verglichen, deren Kantenlängen durch  $2\sigma_y, 2\sigma_x$  im Zweidimensionalen bzw.  $2\sigma_x, 2\sigma_y, 2\sigma_z$  im Dreidimensionalen gegeben sind, vgl. Abbildung 5.9. Der Faktor sollte beispielsweise zu  $f = 10$  gewählt werden. Dadurch wird die Größe der Intervallradien der Koordinaten (Impräzision) auf 10% der entsprechenden  $(1 - \gamma)$ -Standardabweichung der Koordinaten beschränkt. Für eine Wahl von  $\gamma < 5\%$  scheint es gerechtfertigt, die Anteile der Impräzision zu vernachlässigen. Abweichend von dieser Empfehlung kann  $f$  den Anforderungen der jeweiligen Anwendungen angepasst werden. Die Forderung (5.72) ist abhängig von der Orientierung des Koordinatensystems. Diese Abhängigkeit ist dann zu berücksichtigen, wenn die Orientierung des Koordinatensystems frei ist. In diesem Fall bieten sich alternative Zielvorgaben auf Basis der umschließenden Ellipse an, die weiter unten besprochen werden.

#### Individuelle Anpassung auf Basis der Konfidenzellipsen

Alternativ zu einem Kriterium, das auf den Standardabweichungen der Koordinaten beruht, kann die gesamte Information der  $2 \times 2$  bzw.  $3 \times 3$  Submatrizen der Varianz-Kovarianz-Matrix der Koordinaten verwendet werden. Für die Koordinaten eines beliebigen Punktes P wird gefordert:

$$\hat{x}_{r,P,soll} := \frac{\sigma_0}{f} b_{P,1-\gamma}, \quad (5.73)$$

wobei  $b_{P,1-\gamma}$  die Achslänge der kleinen Halbachse der  $(1 - \gamma)$ -Konfidenzellipse bzw. des  $(1 - \gamma)$ -Konfidenzellsoids angibt.

### Punktweise Anpassung der umschließenden Ellipse an die Konfidenzellipse

Soll das Kriterium unabhängig von der Orientierung des Koordinatensystems sein, weil diese beispielsweise in der Planungsphase noch nicht festgelegt werden kann, so eignet sich eine Forderung auf Grundlage der umschließenden Ellipsen, vgl. (4.96) Abschnitt 4.3.6:

$$s_{max}(\mathbf{K}_P) \leq \frac{\sigma_0}{f} b_{P, 1-\gamma}, \quad (5.74)$$

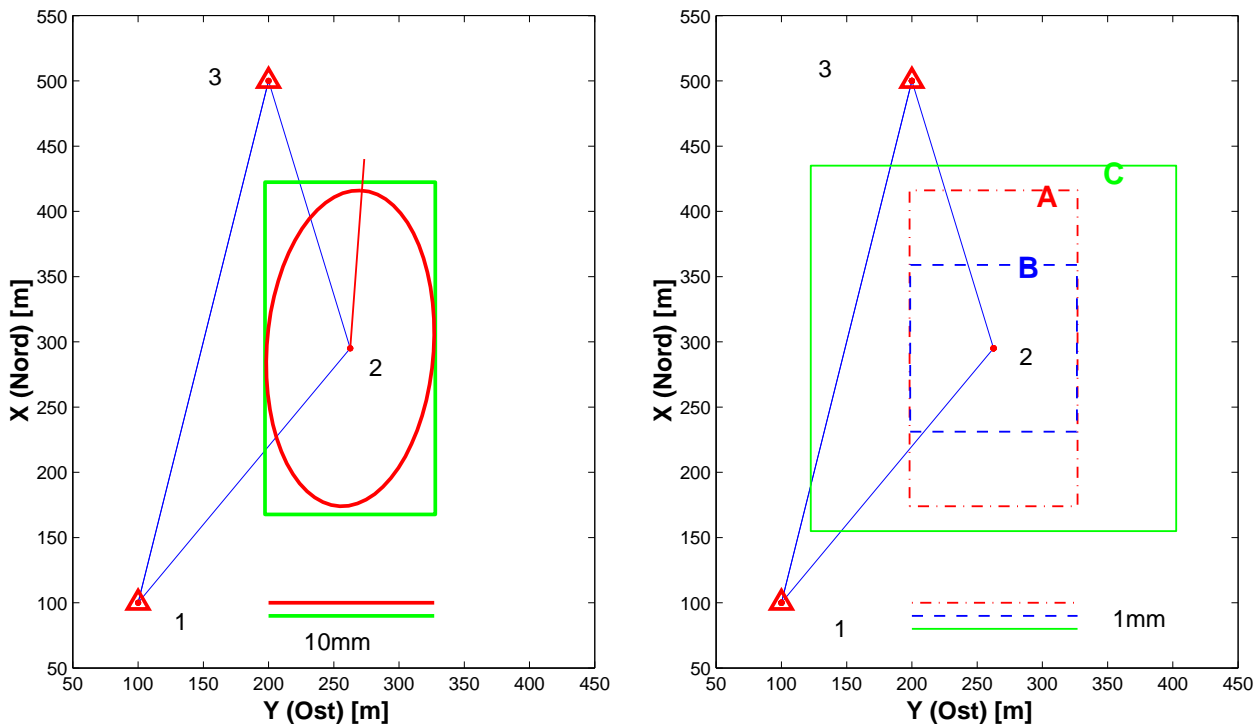
wobei  $s_{max}(\mathbf{K}_P)$  der größte Singulärwert der Matrix der Kantenvektoren des Zonotops des Punktes P ist und  $b_{P, 1-\gamma}$  die Halbachslänge der kleinen Halbachse der Konfidenzellipse bzw. des Konfidenzellipsoids. Im Gegensatz zu (5.72) und (5.73) stellt das Kriterium (5.74) einen nichtlinearen Zusammenhang zwischen den Sollvorgaben und den Intervallradien der Messungen dar.

### Integrale Anpassung auf Basis des durchschnittlichen Punktfehlers

Ist eine individuelle Anpassung für alle Netzpunkte nicht erforderlich, so kann das folgende Kriterium angewandt werden, das sich am durchschnittlichen Punktfehler orientiert. Sei  $n_{pkte}$  die Anzahl der Punkte, so gilt:

$$\hat{x}_{r,i} \leq \frac{\sigma_0}{f^*} \sqrt{\frac{spur \mathbf{Q}_{\hat{x}\hat{x}}}{n_{pkte}}}. \quad (5.75)$$

Orientiert man sich an der Konfidenzwahrscheinlichkeit der Punktfehlerellipse von ca 39.4% bzw. des Punktfehlerellipsoids von 17% so kann als Anhaltspunkt für den zweidimensionalen Fall  $f^* = 4$  und für den dreidimensionalen Fall  $f^* = 2$  gewählt werden. Dieser Ansatz eignet sich insbesondere, wenn homogene Punktfehlerellipsen im Netz vorliegen.



(a) Intervallbox und 95%-Konfidenzellipse des A-priori-Designs

(b) Zielvorgaben A: individuelle Anpassung auf Basis der Standardabweichung, B: individuelle Anpassung auf Basis der Konfidenzellipsen, C: integrale Anpassung auf Basis des durchschnittlichen Punktfehlers

Abbildung 5.9: Darstellung der verschiedenen Zielvorgaben im Intervalldesign für das Beispiel des Vorwärtsschnittes

In der Teilabbildung 5.9(a) wird der Zustand dokumentiert, wenn die Intervallradien mit Standardmess- und Auswerteverfahren erhalten werden, vgl. Faustformel (3.4). Ebenso ist die 95% Konfidenzellipse des Neupunktes

angegeben. Auf ihrer Basis werden die Kriterien zur Anpassung (5.72) - (5.75) berechnet. In Teilabbildung 5.9(b) sind die Zielvorgaben dargestellt, wobei ein um Faktor 10 kleinerer Maßstab vorliegt.

Das Kriterium der individuellen Anpassung auf Basis der Standardabweichung liefert die Box **A**, innerhalb derer die Intervallradien der Koordinaten des Neupunktes liegen müssen, um die Impräzision zu vernachlässigen. Für das Kriterium der individuellen Anpassung auf Basis der Konfidenzellipsen wird die Box **B** als Sollvorgabe gefordert. Für die punktweise Anpassung der umschließenden Ellipse an die Konfidenzellipse muss die umschließende Ellipse innerhalb der Box **B** liegen. Am einfachsten sind die Zielvorgaben der globalen Anpassung auf Basis des mittleren Punktfehlers (5.75) zu erfüllen (Box **C**).

Aus Abbildung 5.9(a) und den Beispielen im Kapitel 4 wird deutlich, dass sich bei Anwendung von Standardmess- und Auswerteverfahren Intervallboxen in der Größe der 95% Konfidenzellipsen ergeben. Es ist daher zu erwarten, dass eine Reduktion um den Faktor 10 nur für wenige Netze möglich sein wird. Je nach Anwendung können sich auch andere Anforderungen als zweckmäßig erweisen.

### 5.4.3 Lösung des Intervalldesigns

Die Lösung des Intervalldesign lässt sich in zwei Stufen unterteilen, vgl. Übersichtstabelle 5.4. In einer ersten Stufe ist die Lösbarkeit zu untersuchen. Hierzu bieten die Überlegungen des Abschnitts 3.3 wichtige Anhaltspunkte, wo minimal erreichbare Koeffizienten  $\underline{\kappa}$  berechnet wurden. Falls die Lösbarkeit gegeben ist, werden maximale individuelle Werte  $\overline{\kappa}$  für die Koeffizienten unter Berücksichtigung der gewählten Zielvorgaben berechnet. Nach diesen Vorarbeiten werden in einer zweiten Stufe die Koeffizienten bestimmt, mit denen die Kriterien „bestmöglich“ erreicht werden können.

#### Stufe I: Lösbarkeitstest und Berechnung maximaler individueller Werte

In der Stufe I wird getestet, ob das Intervalldesign lösbar ist, d.h. ob es Koeffizienten  $\kappa$  gibt, die die Anforderungen, die in Abschnitt 5.4.1 formuliert wurden, erfüllen. Notwendige und hinreichende Bedingung ist, dass:

$$\left| (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} \mid \mathbf{L} \underline{\kappa} \leq \hat{\mathbf{x}}_{r,soll} \right. \quad (5.76)$$

erfüllt ist, wobei die Ungleichung komponentenweise zu verstehen ist. Aus der Betrachtung von (5.76) ergeben sich zwei wesentliche Folgerungen:

- Kann (5.76) für alle Koeffizienten erfüllt werden, so dass es - zumindest theoretisch - möglich ist, die Impräzision zu vernachlässigen. Zusätzlich kann für jeden Koeffizienten individuell eine obere Schranke  $\overline{\kappa}_i$  berechnet werden, die der Wert nicht übersteigen darf, um den Anforderungen zu genügen. Man erhält dann mit der unteren Grenze  $\underline{\kappa}$ , vgl. Abschnitt 3.3, eine Intervallbox  $[\kappa]$ , in der alle zulässigen Koeffizienten enthalten sind. In der zweiten Stufe des Intervalldesigns werden Ansätze vorgestellt, die es ermöglichen, einen „besten“ Satz an Koeffizienten zu wählen.
- Kann (5.76) nicht für alle Koeffizienten erfüllt werden, so ist die Impräzision in der gegebenen Netzkonfiguration nicht zu vernachlässigen. Gegebenenfalls kann durch eine Variation der Netzkonfiguration (FOD) erreicht werden, dass (5.76) erfüllt wird. Ansonsten müssen Impräzision und Stochastizität und getrennt behandelt werden. Unsicherheitsmaße für die Punktposition sind nur in Form der erweiterten Konfidenzbereiche, vgl. Abschnitt 4.4, adäquat repräsentiert.

Wir betrachten zunächst den Fall, dass der Lösbarkeitstest erfüllt wird. Dann kann für jeden Koeffizienten  $\kappa_i$  individuell eine obere Schranke durch die Lösung des folgenden Optimierungsproblems bestimmt werden:

$$\begin{aligned} \max_{i=1,\dots,9} \kappa_i \\ \left| (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} \mid \mathbf{L} \kappa \right. &\leq \hat{\mathbf{x}}_{r,soll}, & (a) \\ \kappa_j &\geq \underline{\kappa}_j. & (b) \end{aligned} \quad (5.77)$$

Die Nebenbedingungen (a), (b) bestimmen den zulässigen Bereich  $\Omega$  der Optimierungsaufgabe. Dieser ist durch ein konvexes Polyeder beschrieben, das im positiven Halbraum liegt, da  $\underline{\kappa}_j \geq 0$ . Da das Matrixprodukt  $\left| (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} \mid \mathbf{L} \right.$  eine nicht-negative Matrix ergibt (alle Elemente  $\geq 0$ ), wird der maximale Wert für einen Parameter  $\kappa_i$  dann angenommen, wenn alle anderen Parameter gleich dem Wert ihrer unteren Schranke sind, d.h. (5.77) kann direkt (ohne Optimierung) gelöst werden.



**Beispiel 32** Betrachten wir als erläuterndes Beispiel den überbestimmten zweidimensionalen Bogenschnitt, wobei der Neupunkt von drei Festpunkten aus durch Messung jeweils einer Strecke eingeschnitten wird. Folgende a priori Werte wurden angesetzt: die Standardabweichung beträgt  $\sigma_S = 2 \text{ mm} + 2 \text{ ppm } S$ . Die Koeffizienten der Faustformeln ergeben sich für das Instrument TPS1101 nach Tabelle (3.3):  $\kappa_{\text{a priori}} = (1.5[\text{mm}], 1[\text{ppm}])^T$ . Für das Netz, das in Abbildung 5.10(a) veranschaulicht ist, wurden für die drei Strecken folgende Intervallradien  $S_{1,3,r} = 1.7 \text{ mm}$ ,  $S_{2,3,r} = 1.8 \text{ mm}$  und  $S_{4,3,r} = 1.7 \text{ mm}$  berechnet. Mit  $\underline{\kappa} = (0.15[\text{mm}], 0.75[\text{ppm}])^T$  liefert der Lösbarkeitstest (5.76) folgende Werte:

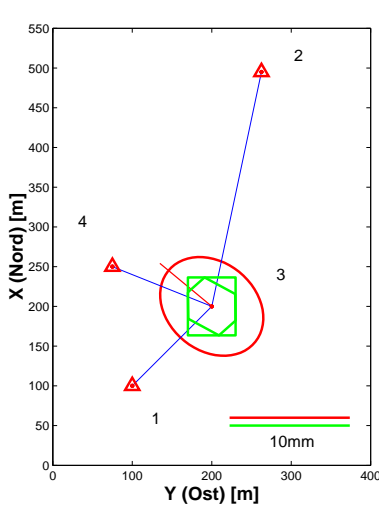
Koordinate	Zielvorgaben		erreichbare Werte		
	$\hat{x}_{r,i,soll}$ individuelle Anpassung	$\hat{x}_{r,i,soll}$ integrale Anpassung	$\hat{x}_{r,i}(\underline{\kappa})$ Minimalwerte	$\hat{x}_{r,i}(\bar{\kappa})$ Maximalwerte	$\hat{x}_{r,i}(\kappa_{\text{a priori}})$ a priori Werte
$y_3$	0.42	0.60	0.30	1.02	1.94
$x_3$	0.41	0.60	0.42	1.33	2.36

Tabelle 5.5: Lösbarkeitstest für das Kriterium der individuellen Anpassung auf Basis der Standardabweichung und der integralen Anpassung auf Basis des durchschnittlichen Punktfehlers, alle Werte in [mm].

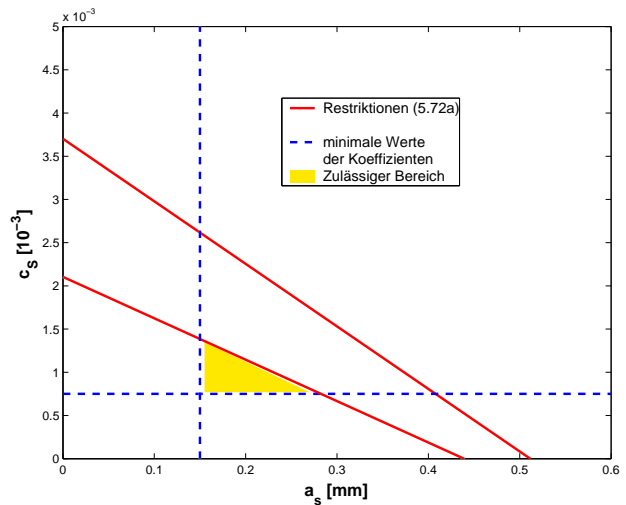
In der zweiten Spalte sind die Zielvorgaben für die Intervallradien der Koordinaten bei individueller Anpassung auf Basis der Standardabweichung angegeben, in der dritten Spalte bei integraler Anpassung auf Basis des durchschnittlichen Punktfehlers. Der Lösbarkeitstest liefert minimale Intervallradien, die erreicht werden können, vgl. vierte Spalte. Für die  $x$ -Koordinate liegt dieser Wert um 4% über den Zielvorgaben der individuellen Anpassung, d.h. der Lösbarkeitstest ist nicht erfüllt. Für die  $y$ -Koordinate können die Kriterien jedoch eingehalten werden. Die Kriterien der integralen Anpassung werden für beide Koordinaten erfüllt. In der fünften und in der letzten Spalte sind die Intervallradien der Koordinaten angegeben, die sich aus den maximalen Werten für die Koeffizienten  $\bar{\kappa}$  und den a priori Werten berechnen lassen.

Betrachten wir die zwei Koeffizienten der Streckenmessung, so zeigt Abbildung 5.10(b) den zulässigen Bereich der Koeffizienten  $a_S$  und  $c_S$ . Der Koeffizient  $a_S$  beschreibt den konstanten Beitrag zum Intervallradius der Streckenmessung,  $c_S$  den Beitrag, der linear mit der Streckenlänge zunimmt, vgl. Abschnitt 3.3.2. Da die Matrix  $|(\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll}| \mathbf{L}$  nicht-negativ ist, sind die Steigungen der Geraden, die die Restriktionen beschreiben, immer negativ. Die obere Grenze ist gegeben durch  $\bar{\kappa} = (0.28[\text{mm}], 1.4[\text{ppm}])^T$ . Die individuelle Behandlung der einzelnen Koeffizienten führt allerdings dazu, dass nicht gleichzeitig die Extremwerte angenommen werden können, d.h.

$$|(\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll}| \mathbf{L} \bar{\kappa} > \hat{\mathbf{x}}_{r,soll}. \tag{5.78}$$



(a) Intervallbox, Wertebereich und 95%-Konfidenzellipse des Startdesigns



(b) Zulässiger Bereich für eine Optimierung im Intervall-design bei Vorgabe integraler Maximalwerte

Abbildung 5.10: Intervalldesign für den überbestimmten zweidimensionalen Bogenschnitt

Eine Darstellung des Randes der Restriktionsmenge in Abhängigkeit der Parameter liefert den funktionalen Zusammenhang, wie sich eine Variation eines Koeffizienten auf alle anderen auswirkt. Für alle neun Koeffizienten der Faustformeln ist der Zusammenhang durch ein neundimensionales konvexes Polyeder gegeben, so dass eine explizite Parametrisierung zu aufwändig ist. Für den einfachen Fall zweier Koeffizienten bei der Streckenmessung ergibt sich:

$$c_S(a_S) = 2.1 \cdot 10^{-3} - 4 \cdot 10^{-3} a_S, \quad \text{für } a_S \in [0.15, 0.28]. \quad (5.79)$$

Es lassen sich die folgenden Schlussfolgerungen ziehen:

1. Bei Anwendung des individuellen Kriteriums (5.72) (was zu bevorzugen ist), ist eine Vernachlässigung der Impräzision im obigen Beispiel knapp nicht möglich, da die Zielvorgaben mit den Minimalwerten  $\underline{\kappa}$  der Koeffizienten nicht einzuhalten sind.
2. Bei Anwendung des integralen Kriteriums (5.75) lassen sich Spielräume für die Koeffizienten erkennen, die im Folgenden ausgenutzt werden sollen. Man hat somit eine gewisse Wahl, die Koeffizienten  $a_S$  und  $c_S$  optimal zu bestimmen, d.h. ein passendes Instrumentarium zu wählen bzw. den Aufwand der Erfassung der Refraktion zu begrenzen.

### Stufe II:

In der Stufe II wird ein Vektor der Koeffizienten bestimmt, der die Anforderungen mit minimalem Aufwand erfüllt, d.h. ein Vektor, der einerseits möglichst nahe an der oberen Grenze  $\bar{\kappa}$  liegt und andererseits die Sollvorgaben für die Intervallradien der Koordinaten bestmöglich (von unten) annähert. Durch die funktionale Abhängigkeit der Intervallradien von den Koeffizienten wird man es im Allgemeinen nicht erreichen, dass der Abstand

$$\Delta \hat{x}_{r,i} = \hat{x}_{r,i} - \hat{x}_{r,i,soll} \quad (5.80)$$

für alle Koordinaten  $i$  gleichzeitig verschwindet. Für eine maximal tolerierbare Impräzision sollte der Vektor  $\hat{x}_{r,i,soll}$  „bestmöglich“ angepasst werden.

Ein erstes (automatisches) Verfahren kann durch die Forderung einer quadratischen Anpassung realisiert werden. Hierbei wird im Sinn der Vektoroptimierung (Normskalarisierung mit identischen Gewichten) versucht, beiden Zielvorgaben gerecht zu werden. Sei  $\mathbf{f} = (\hat{\mathbf{x}}_r, \boldsymbol{\kappa})^T$  der  $(u+9) \times 1$  Vektor und

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{22} \end{pmatrix} \quad \text{mit} \quad \mathbf{M}_{11} = \text{diag}\left(\frac{1}{\hat{x}_{r,i,soll}}\right) \quad \text{und} \quad \mathbf{M}_{22} = \text{diag}\left(\frac{1}{\kappa_i}\right) \quad (5.81)$$

die  $(u+9) \times (u+9)$  Diagonalmatrix zur Homogenisierung der Einheiten bzw. Größenordnungen und folgenden Vektor der Sollwerte

$$\mathbf{f}_{soll} = (\hat{\mathbf{x}}_{r,soll}, \bar{\boldsymbol{\kappa}})^T, \quad (5.82)$$

so wird eine bestmögliche Anpassung an die Sollwerte im quadratischen Mittel durch folgende Vektoroptimierungsaufgabe erzielt:

$$\begin{aligned} \min_{\boldsymbol{\kappa}} (\mathbf{f} - \mathbf{f}_{soll})^T \mathbf{M}^T \mathbf{M} (\mathbf{f} - \mathbf{f}_{soll}) \\ \left| (\bar{\mathbf{A}}^T \mathbf{P}_{ll} \bar{\mathbf{A}})^{-1} \bar{\mathbf{A}}^T \mathbf{P}_{ll} \right| \mathbf{L} \boldsymbol{\kappa} &\leq \hat{\mathbf{x}}_{r,soll}, & (a) \\ \boldsymbol{\kappa} &\geq \underline{\boldsymbol{\kappa}}. & (b) \end{aligned} \quad (5.83)$$

Die Lösung  $\boldsymbol{\kappa}^*$  von (5.83) liefert eine bestmögliche Anpassung der Maximalwerte der Intervallradien der Koordinaten. Die Variation der einzelnen Koeffizienten ist mit unterschiedlichen Kosten bzw. Aufwand verbunden: Die Verwendung eines Instrumentes einer höherwertigen Instrumentenklasse verteuert (lediglich) die Messung. Das Messen in zwei Lagen verdoppelt den Aufwand der Anzielungen für die Richtungen und somit die Messzeit. Der Einfluss der Prismen kann ganz ausgeschaltet werden, wenn eine entsprechende Ausrichtung der Prismen vorgenommen wird. Die Reduktion des Einflusses der Refraktion ist schließlich mit dem höchsten messtechnischen und finanziellen Aufwand verbunden, wobei immer die Problematik der Repräsentativität der Werte zu berücksichtigen ist, vgl. die Diskussion in Abschnitt 3.3. Die quadratische Anpassung berücksichtigt diese Randbedingungen nicht, was durch die Fortsetzung des Beispiels 32 verdeutlicht werden soll.

**Beispiel 33** Wir setzen das Beispiel 32 fort. Für die quadratische Anpassung nach (5.83) erhält man den folgenden Lösungsvektor des Intervalldesigns:

$$\boldsymbol{\kappa}^* = (0.22[\text{mm}], 1.0[\text{ppm}])^T. \quad (5.84)$$

$x_{r,\text{soll}}$ [mm]	$x_r(\boldsymbol{\kappa}^*)$ [mm]	$\Delta x_r$ [mm]	$\Delta x_r$ [%]
0.60	0.43	0.17	28.5
0.60	0.60	0.00	0.0

Tabelle 5.6: Vergleich der Anpassung bei quadratischer Zielfunktion am Beispiel des überbestimmten Bogenschnittes

Die obige Tabelle zeigt, wie die Maximalanforderungen  $x_{r,\text{soll}}$  durch den Lösungsvektor  $\boldsymbol{\kappa}^*$  exakt eingehalten bzw. unterschritten werden. Mit den Minimalwerten verschiedener Instrumente nach Tabelle 3.3 würde dieses Ergebnis am Besten mit dem Instrument TCA2003 ( $a_{S,TCA2003} = 0.17$  und  $c_{S,TCA2003} = 0.9$ ) realisiert werden. Allerdings ist hier noch ein gewisser Spielraum des Koeffizienten  $a_S$ , der zugunsten von  $c_S$  verringert werden könnte, ohne dass ein höherwertiges und somit teureres Instrument zum Einsatz kommen muss. Dadurch kann der Koeffizient  $c_S$  einen größeren Wert („mehr Impräzision“) annehmen. Für die Messanordnung heißt dies, dass der Einfluss der Refraktion mit einem geringeren Aufwand erfasst werden kann, was die Kosten der Messanordnung reduziert. Das zweite Verfahren, das im Folgenden vorgestellt wird, eignet sich, direkt diese Rahmenbedingungen zu berücksichtigen.

Als Alternative zum Verfahren der quadratischen Anpassung wird ein *iteratives und interaktives Vorgehen* vorgeschlagen, das dem Anwender Spielraum lässt, welcher Koeffizient wie stark von den Maximalwerten  $\overline{\boldsymbol{\kappa}}$  reduziert werden soll. Die Lösung  $\boldsymbol{\kappa}^*$  des Verfahrens der quadratischen Anpassung kann dabei einen Anhaltspunkt geben, welche Größenordnungen der Reduktionen realistisch sind. Der Ansatz basiert auf dem Optimierungsproblem (5.77), wobei in jedem Iterationsschritt  $l$  eine Restriktion in der Form

$$\mathbf{b}_k^T \boldsymbol{\kappa} = \overline{\boldsymbol{\kappa}}_k^{(l)} \quad \text{oder alternativ:} \quad \boldsymbol{\kappa}_k = \mathbf{c}_k^T \overline{\boldsymbol{\kappa}}_k^{(l)} \quad (5.85)$$

für einen weiteren Koeffizienten  $\boldsymbol{\kappa}_k$  hinzugefügt wird. Durch den Zeilenvektor  $\mathbf{b}_k^T$  wird der Wert des  $k$ -ten Koeffizienten auf das Niveau  $\frac{\overline{\boldsymbol{\kappa}}_k^{(l)}}{b_k} = c_k \overline{\boldsymbol{\kappa}}_k^{(l)}$ ,  $0 < c_k \leq 1$ ,  $b_k \geq 1$  fixiert.

$$\begin{aligned} & \max_{i \in \mathcal{I}} \boldsymbol{\kappa}_i \\ & | (\overline{\mathbf{A}}^T \mathbf{P}_{ll} \overline{\mathbf{A}})^{-1} \overline{\mathbf{A}}^T \mathbf{P}_{ll} | \mathbf{L} \boldsymbol{\kappa} \leq \hat{\mathbf{x}}_{r,\text{soll}}, \quad (a) \\ & \boldsymbol{\kappa}_j \geq \overline{\boldsymbol{\kappa}}_j, \quad (b) \\ & \mathbf{B}^{(l)} \boldsymbol{\kappa} = \overline{\boldsymbol{\kappa}}_k^{(l)}, \quad (c) \end{aligned} \quad (5.86)$$

wobei  $\mathbf{B}^{(l)}$  eine  $l \times 9$  Matrix ist. Für eine iterative Anwendung von (5.86) sind folgende Schritte nötig:

1. Die Berechnung von (5.77) liefert einen Startvektor der individuellen oberen Schranken  $\overline{\boldsymbol{\kappa}}^{(1)}$ .
2. Durch die Wahl eines Koeffizienten  $\boldsymbol{\kappa}_k$  und eines Reduktionsniveaus  $b_k$  bzw.  $c_k$  ist durch (5.85) der optimale Wert  $\boldsymbol{\kappa}_k^{**}$  für diesen Koeffizienten durch den Anwender fixiert.
3. Die Menge  $\mathcal{I}$  bezeichne die noch nicht fixierten Koeffizienten. Die Berechnung von (5.86) für alle Koeffizienten  $\boldsymbol{\kappa}_i$ ,  $i \in \mathcal{I}$ , liefert für diese neue individuelle obere Schranken  $\overline{\boldsymbol{\kappa}}_i^{(l+1)}$  und neue Sollvorgaben  $\hat{\mathbf{x}}_{r,\text{soll}}^{(l+1)}$ :

$$\hat{\mathbf{x}}_{r,\text{soll}}^{(l+1)} := \hat{\mathbf{x}}_{r,\text{soll}}^{(l)} - | (\overline{\mathbf{A}}^T \mathbf{P}_{ll} \overline{\mathbf{A}})^{-1} \overline{\mathbf{A}}^T \mathbf{P}_{ll} | \mathbf{L} \boldsymbol{\kappa}', \quad \boldsymbol{\kappa}' := (0, \dots, 0, \boldsymbol{\kappa}_k^{**}, 0, \dots, 0). \quad (5.87)$$

**Beispiel 34** Wir setzen das Beispiel 32 fort, um das zweite Verfahren zu erläutern. Um eine gezielte Anpassung zu erreichen und um somit das gesamte Potential auszuschöpfen, d.h. den Aufwand der Messung zu minimieren, wird der zweite Ansatz verwendet. In der ersten Iterationsstufe wird  $a_S = a_{S,TCA2003} = 0.17 \text{ mm}$  auf den Wert des Instrumentariums TCA2003, vgl. Tabelle 3.3, festgesetzt. Der Reduktionsfaktor hat dabei den Wert  $c_i = 0.6$ , d.h. der optimale Wert für den Koeffizienten  $\boldsymbol{\kappa}_1^{**} = a_{S,TCA2003}$  entspricht 60% des Maximalwertes  $\overline{\boldsymbol{\kappa}}_i$ , vgl. Beispiel 32, der aufgrund der Zielvorgaben möglich wäre. Dadurch wird der maximale Spielraum für den Koeffizienten  $c_S$  gelassen. Als Ergebnisvektor erhält man:

$$\boldsymbol{\kappa}^{**} = (0.17[\text{mm}], 1.3[\text{ppm}])^T. \quad (5.88)$$

Durch den größeren Wert  $\kappa_2^{**}$  kann einerseits der eingeschränkten Repräsentativität des Brechungsindex für den gesamten Lichtweg besser Rechnung getragen und andererseits der Aufwand der Erfassung reduziert werden. Wird das TCA2003 verwendet, so muss der Brechungsindex der Atmosphäre auf  $\pm 0.5ppm$  bestimmt werden. Mit diesen Werten werden die Anforderungen nach (5.72) eingehalten bzw. unterschritten, was in der folgenden Tabelle dargestellt ist.

$x_{r,soll}$ [mm]	$x_r(\kappa^{**})$ [mm]	$\Delta x_r$ [mm]	$\Delta x_r$ [%]
0.60	0.41	0.19	31.9
0.60	0.60	0.00	0.0

Tabelle 5.7: Vergleich der Anpassung beim zweiten Ansatz im Intervalldesign

Es muss an dieser Stelle ausdrücklich angemerkt werden, dass eine weitaus größere Auswahl an Instrumenten verschiedenster Hersteller zur Verfügung steht und ggf. andere Werte ausgewählt werden können. Die Ergebniswerte der Beispiele 32-34 sind daher rein exemplarisch zu verstehen. Da im Rahmen der vorliegenden Arbeit keine Instrumentenuntersuchungen, sondern primär theoretische Überlegungen durchgeführt wurden, kann die obige Wahl keine Präferenz für einen bestimmten Hersteller ausdrücken.

#### 5.4.4 Beispiel und Wertung der Lösung des Intervalldesigns

Zusammenfassend lässt sich festhalten, dass eine individuelle und auch integrale Anpassung nur schwer zu erreichen ist. Durch die Ähnlichkeit zwischen Stochastizität und Impräzision, die durch Gleichung (5.61) und (5.62) ausgedrückt ist, fehlt ein Spielraum, die Anpassung durch Variation der Netzkonfiguration zu verbessern.

Betrachten wir als abschließendes Beispiel die Anpassung der Koeffizienten beim dreidimensionalen Linach-Netz. Da bei einem guten Ausgangsdesign die Intervallboxen und Konfidenzellipsen ungefähr in der selben Größenordnung liegen, ist eine Verringerung gemäß der Vorgaben in Abschnitt 5.4.2 nicht immer möglich. Im Folgenden soll deshalb zur Demonstration des Verfahrens ein abgeschwächtes Kriterium verwendet werden:

$$\hat{x}_{r,i} \leq \sqrt{\frac{\text{spur } \mathbf{Q}_{\hat{x}\hat{x}}}{n_{pkte}}} \quad (5.89)$$

Der Lösbarkeitstest (5.76) ist erfüllbar und es wird eine obere Schranke  $\bar{\kappa}$  für die Koeffizienten berechnet, vgl. Tabelle 5.8, vierte Spalte. In der zweiten Phase des Intervalldesign geht es darum, aus dem Intervall  $[\kappa]$  einen Vektor der Koeffizienten auszuwählen, der einerseits (5.89) erfüllt und andererseits eine Umsetzung mit möglichst geringem Aufwand ermöglicht. Für einen ersten Anhaltspunkt, wie ein solcher Vektor aussehen kann und welche Spielräume für die Reduktion der einzelnen Koeffizienten möglich sind, wird die Optimierungsaufgabe (5.83) gelöst. In der fünften und sechsten Spalte der Tabelle 5.8 ist der Ergebnisvektor  $\kappa^*$  angegeben und die prozentualen Reduktionsbeträge, auf die die oberen Grenzen  $\bar{\kappa}$  dabei verringert wurden. In einem zweiten Schritt ist mit Hilfe der Ergebnisse der in Abschnitt 3.3 geführten Diskussion zu untersuchen, welche Instrumentarien, Mess- und Auswertemethoden, die exemplarisch in dieser Arbeit betrachtet wurden, verwendet werden müssten, um diese Koeffizienten zu realisieren.

- Für die Wahl der Instrumente folgt: die Forderungen an die Streckenmessung sind nur mit dem Mekometer 5000 einzuhalten, die Anforderungen für die Richtungs- bzw. Zenitdistanzmessung mit dem TCA2003. Dieses Ergebnis ist unbefriedigend, da das Netz mit zwei verschiedenen Instrumenten gemessen werden muss, was den Aufwand verdoppelt.
- Für die Messmethoden folgt, dass die Richtungen und Zenitdistanzen in zwei Lagen gemessen werden müssen. Eine geräteinterne Korrektur der Instrumentenfehler nach (3.1) bzw. (3.4) ist für diese Anwendung nicht ausreichend.
- Für die Prismenausrichtung liegt ein gewisser Spielraum vor. Daher müssten nicht unbedingt Prismen verwendet werden, die das Konstruktionskriterium (3.8) erfüllen. Umgekehrt sind bei dieser Staumauer die Prismen nicht fest installiert, so dass eine individuelle Ausrichtung möglich ist. Dadurch könnte der Spielraum auf andere Koeffizienten umverteilt werden.
- An die Refraktionserfassung werden höchste Anforderungen gestellt, diese erhöhen den Messungsaufwand erheblich. Hier wäre es wünschenswert, größere Spielräume zur Verfügung zu haben.

Koeffizient	Startdesign			quadratische Anpassung		interaktive Anpassung	
	$\kappa_{\text{apriori}}$	$\underline{\kappa}$	$\overline{\kappa}$	$\kappa^*$	$\frac{\kappa^*}{\overline{\kappa}}$ in [%]	$\kappa^{**}$	$\frac{\kappa^{**}}{\overline{\kappa}}$ in [%]
$a_S$ [mm]	1.50	0.15	0.22	0.15	65.5	0.17	75.0
$c_S$ [ppm]	1.00	0.75	0.16	0.8	45.6	0.10	72.0
$a_R$ [mgon]	0.10	0.01	0.03	0.01	28.0	0.01	26.5
$b_R$ [mgon m]	8.80	0.0	2.37	1.08	45.9	0.00	0.0
$c_R$ $\frac{\text{mgon m}}{\text{m}}$	0.16	0.16	0.19	0.16	85.2	0.16	85.0
$d_R$ [mgon]	0.20	0.1	0.37	0.12	34.6	0.10	26.9
$a_\zeta$ [mgon]	0.45	0.1	0.24	0.13	45.7	0.10	42.0
$b_\zeta$ [mgon m]	8.70	0.0	12.5	9.48	24.1	0.00	0.0
$c_\zeta$ $\frac{\text{mgon m}}{\text{m}}$	0.16	0.16	0.31	0.16	48.8	0.30	96.8

Tabelle 5.8: Übersicht über die optimalen Koeffizienten  $\kappa^*$  bei der quadratischen Anpassung bzw.  $\kappa^{**}$  bei der interaktiven Anpassung

Als Alternative soll deshalb die interaktive Anpassung verwendet werden, die dem Anwender erlaubt, die Verteilung der Impräzision auf die Koeffizienten zu steuern, vgl. Optimierungsaufgabe (5.86). In den ersten zwei Iterationsschritten werden die Koeffizienten  $b_R$  und  $b_\zeta$  auf ihre Minimalwerte fixiert, d.h. es wird davon ausgegangen, dass die Prismen exakt ausgerichtet werden können. Dadurch verbleiben für alle anderen Koeffizienten die maximal möglichen Variationsbereiche.

In einem zweiten Schritt soll versucht werden, dass das Netz nur mit einem Messinstrument, dem TCA2003, gemessen werden soll. Dies kann durch Reduktion der Koeffizienten  $a_S$ ,  $c_S$  bzw.  $a_R$ ,  $d_R$ ,  $a_\zeta$  auf die Werte des TCA2003 erreicht werden. Der dann noch verbleibende Spielraum soll für die Impräzision der Refraktionserfassung verwendet werden. Tabelle 5.8 zeigt in der vorletzten Spalte die so erhaltenen Koeffizienten  $\kappa^*$ . Die letzte Spalte gibt an, auf welchen Prozentsatz dabei die maximalen Werte  $\overline{\kappa}$  reduziert wurden.

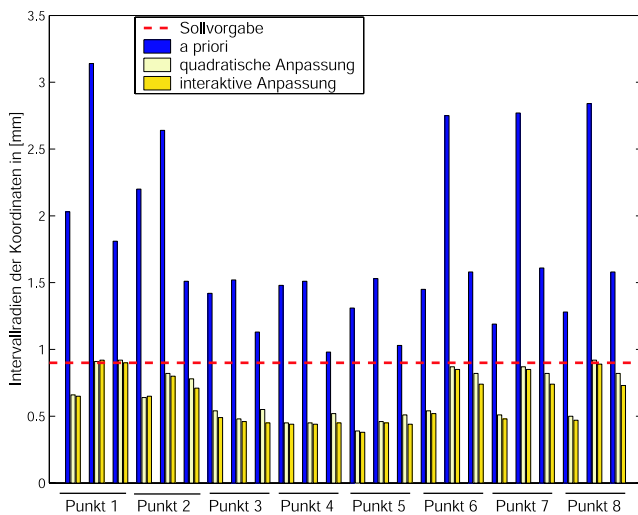


Abbildung 5.11: Vergleich der Sollvorgaben  $\mathbf{x}_{r,soll}$  der Intervallradien der Koordinaten mit den Intervallradien  $\mathbf{x}_r(\kappa^*)$  bzw.  $\mathbf{x}_r(\kappa^{**})$ , die sich mit den optimalen Koeffizienten  $\kappa^*$  der quadratischen Anpassung bzw.  $\kappa^{**}$  der interaktiven Anpassung ergeben.

philie“ der Intervallmethode widerspricht. Diese sieht vor, dass die Sollvorgaben feste obere Schranken darstellen, die auf keinen Fall überschritten werden dürfen.

Beide Verfahren liefern somit vergleichbare Lösungen, bei denen die Sollwerte der Intervallradien der Koordinaten eingehalten oder unterschritten werden. Die interaktive Anpassung weist den großen Vorteil auf, dass die Präferenzen des Nutzers berücksichtigt werden können, wie das Impräzisionsbudget auf die Koeffizienten der Faustformeln verteilt werden soll. Dadurch werden der Messaufwand und die Kosten stark reduziert.

Abbildung 5.11 gibt einen Vergleich der Intervallradien der Koordinaten  $x$ ,  $y$ ,  $z$  pro Punkt vor und nach der Optimierung im Intervalldesign an. In Schwarz sind die vorgegebenen Sollwerte  $\mathbf{x}_{r,soll}$  der Intervallradien pro Koordinate dargestellt, in einem helleren Grau, die Intervallradien  $\mathbf{x}_r(\kappa^*)$ , die sich bei einer quadratischen Anpassung ergeben und in einem dunkleren Grau die Intervallradien  $\mathbf{x}_r(\kappa^{**})$ , die aus einer quadratischen Anpassung resultieren.

Es wird deutlich, dass für viele Koordinaten die Zielvorgaben weit unterschritten werden können. Durch die Koppelungsmechanismen zwischen den Koeffizienten der Faustformeln und der Intervallradien der Messungen einerseits sowie den Intervallradien der Messungen und denen der Koordinaten andererseits können nicht alle Anforderungen gleichzeitig exakt erreicht werden. Um die Abweichungen zu verringern und somit eine bessere Anpassung zu erreichen, könnte eine quadratische Anpassung der Intervallradien an die Sollvorgaben ohne die Nebenbedingung (a):  $\mathbf{x}_r(\kappa) \leq \mathbf{x}_{r,soll}$  berechnet werden. Die Radien der Ergebnisintervalle würden in diesem Fall die Sollvorgaben auch überschreiten dürfen. Dieser Ansatz ist als Lösungsstrategie wenig geeignet, da er der „Philosophie“ der Intervallmethode widerspricht.

In diesem Beispiel wurden abgeschwächte Zielvorgaben verwendet, um das Prinzip des Intervalldesigns zu demonstrieren. Die „scharfe“ Formulierung der Zielvorgaben nach (5.72) - (5.75) können beispielsweise für das Linach-Netz sowie die Diagonalenviereckskette nicht eingehalten werden. Als Konsequenz heißt dies, dass in Bezug auf die genannten Zielvorgaben und die betrachtete Netzkonfigurationen die *Impräzision nicht vernachlässigt werden darf*. Beide Anteile der erweiterten Unsicherheit müssen für eine realistische Unsicherheitsbehandlung betrachtet werden, wie dies in Abschnitt 2 vorgeschlagen wurde.

Ergänzend soll kurz diskutiert werden, ob durch eine Variation der Netzkonfiguration ein neues Netzdesign erhalten werden kann, mit dem die Kriterien einzuhalten sind. Die Netzpunkte, die die Staumauer repräsentieren, können nicht variiert werden, sondern lediglich die Kontrollpunkte sind als variabel anzusehen. Die Problematik der Lösung dieser Aufgabenstellung liegt in der Ähnlichkeit des Verhaltens von Stochastizität und Impräzision, die in Abschnitt 5.3.3 gezeigt wurde. Es müsste eine Netzkonfiguration gefunden werden, bei der das Verhältnis zwischen den Kenngrößen der Stochastizität und der Impräzision deutlich vergrößert werden kann.

Ein alternativer Ansatz wäre ein gezieltes Einführen weiterer Messungen. Dadurch wird die Bestimmungsgeometrie und das Übertragungsverhältnis der Impräzision der Messungen auf die ausgeglichenen Koordinaten verbessert. Gleichzeitig wirkt sich dieser Mechanismus aber auch auf die Varianz-Kovarianz-Matrix aus, so dass die Sollvorgaben ebenfalls verringert werden. Andererseits wird zusätzliche Impräzision eingeführt, die die Verbesserungen teilweise kompensieren kann oder die Impräzision der Koordinaten sogar erhöhen kann, vgl. die Diskussion in Abschnitt 5.3.2. Weitere Untersuchungen dieser Fragestellung wären wünschenswert, insbesondere den Einfluss einzelner Beobachtungen auf die Impräzision der Koordinaten analytisch angeben zu können.

Umgekehrt kann die Anpassung der Impräzision an gewisse Zielvorgaben unter einem zweiten Blickwinkel betrachtet werden. Neben der Optimierung im Intervalldesign mit der Zielsetzung der Vernachlässigung der Impräzision kann eine zweite Ausrichtung verfolgt werden. Hierbei steht das Einhalten von Maximalvorgaben für die Impräzision im Vordergrund, die beispielsweise in einem Lastenheft für eine konkrete Anwendung gefordert werden, vgl. BRAEMS (2002, S.155ff) oder BRAEMS ET AL. (2003). Es werden analog zu (5.72) - (5.75) obere Schranken für die Impräzision der Koordinaten oder Punkte festgelegt, die nicht überschritten werden dürfen. Somit wird der Worst-Case der Impräzision der Punkte eingeschränkt. In Verbindung mit einer unteren Schranke für die Impräzision, die sich mit den Minimalwerten der Koeffizienten der Faustformeln angeben lässt, ist eine innere und äußere Einschließung der Impräzision möglich. BRAEMS (2002, S.165ff) zeigt eine Anwendung für die Gesamtunsicherheit verschiedener Thermometer. Durch gekoppelte Minimax-Optimierungsprobleme werden die innere und äußere Einschließungen berechnet. Ist die Lösungsmenge leer, so erfüllt das untersuchte Thermometer nicht die Vorgaben im Lastenheft und ist für die Lösung der Aufgabenstellung nicht geeignet.

## 5.5 Gemeinsame Optimierung von Impräzision und klassischen Zielvorgaben

Nachdem einzelne Optimalitätskriterien bezüglich Genauigkeit, Zuverlässigkeit und Impräzision in den vorangegangenen Abschnitten getrennt für verschiedene Netze untersucht wurden, sollen diese Zielgrößen in einer Vektoroptimierung zusammengeführt und kombiniert werden. In einem ersten Schritt soll die erweiterte Unsicherheit (Stochastizität und Impräzision) minimiert werden. In einem zweiten Schritt wird das Kriterium der Zuverlässigkeit zu diesen beiden Zielvorgaben hinzu genommen und die Aufgabenstellung durch Vektoroptimierung gelöst.

### 5.5.1 Optimierung der erweiterten Unsicherheit

Die erweiterte Unsicherheit ergibt sich aus der Überlagerung der Stochastizität und Impräzision der Punktkoordinaten, vgl. Abschnitt 4.4. Im Folgenden sollen die Ideen, die bereits in Abschnitt 5.3.3 vorgestellt wurden, aufgegriffen werden. Dabei wird gezeigt, dass eine Reduktion der erweiterten Unsicherheit entweder durch eine Optimierung spektraler Zielfunktionen für die Impräzision oder für die Stochastizität erreicht wird.

In Abschnitt 5.3.3 wurde die Ähnlichkeit der Reaktion von spektralen Zielfunktionen der Stochastizität und der Impräzision bei einer Variation des Netzdatums, der Netzgeometrie oder der Gewichtung der Beobachtungen herausgearbeitet. Der Grad der Ähnlichkeit wird dabei durch das Verhältnis  $\mathbf{P}_{ii}^{\frac{1}{2}} \mathbf{L}_r$  der Wurzel der Gewichte zu den Intervallradien der Messungen bestimmt. Für kleinräumige Netze ist diese Relation nahezu für alle Beobachtungen eines Messungstyps konstant, da sich die Variation der Intervallradien mit der Streckenlänge nach den Faustformeln (3.2), (3.4) und (3.6) nur gering ausprägt. Es werden allerdings für unterschiedliche

Messungstypen verschiedene Werte für das Verhältnis erreicht. Lediglich für sehr heterogene Netze, in denen sowohl geringe ( $< 100\text{ m}$ ) als auch große Punktabstände ( $>> 1000\text{ m}$ ) vorkommen, kann das Verhältnis zwischen Gewichtung und Intervallradien auch für Beobachtungen eines Messungstypen stärker variieren.

Für den Spezialfall  $\mathbf{P}_{ii}^{\frac{1}{2}}\mathbf{L}_r = \eta\mathbf{I}$  wird die Ähnlichkeit zur Äquivalenz der Zielfunktionen:

$$\begin{aligned}\|\mathbf{K}\|_F &= \sqrt{\sum_{i=1}^u s_i^2} = \eta\sqrt{\text{spur}\mathbf{Q}_{\hat{x}\hat{x}}}, \\ \|\mathbf{K}\|_2 &= s_1 = \eta\sqrt{\|\mathbf{Q}_{\hat{x}\hat{x}}\|_2}.\end{aligned}\quad (5.90)$$

In beiden Fällen (Ähnlichkeit oder Äquivalenz) führt eine Reduktion der Werte einer spektralen Zielfunktion der Impräzision auch zur Reduktion der Werte der spektralen Zielfunktionen der Stochastizität und umgekehrt, vgl. Tabelle 5.2 für das Beispiel der Diagonalenviereckskette oder Tabelle 5.3 für ein flächenhaftes Netz. Da kein gegenläufiges Verhalten von Impräzision und Stochastizität vorliegt, vereinfacht sich die Optimierungsaufgabe für die erweiterte Unsicherheit. Als Zielfunktionen kann direkt auf die spektralen Kenngrößen der Impräzision oder Stochastizität zurückgegriffen werden. Dadurch wird eine aufwändige Formulierung von Kenngrößen oder Zielfunktionen, vgl. Abschnitt 4.4, elegant umgangen.

### 5.5.2 Gemeinsame Optimierung von Impräzision, Stochastizität und Zuverlässigkeit

Abschließend soll die gemeinsame Optimierung von Impräzision, Genauigkeit und Zuverlässigkeit vorgestellt werden. Hierzu werden Methoden der Vektoroptimierung verwendet, wobei jeweils eine der Zielfunktionen für die einzelnen Kriterien, die in Abschnitt 5.2.2 vorgestellt wurden, ausgewählt werden soll. Es ist zu erwarten, dass sich die Zielfunktionen  $\Phi_A$  und  $\Phi_F(\mathbf{K})$  bzw.  $\Phi_S$  und  $\Phi_2(\mathbf{K})$  je ähnlich verhalten werden, vgl. Diskussion in Abschnitt 5.3.3. Um alle Eigenschaften abzudecken, kann deshalb beispielsweise  $\Phi_S$  als Zielfunktion für die Genauigkeit angesetzt werden und  $\Phi_F(\mathbf{K})$  als Zielfunktion für die Impräzision. Als Kriterium für die Zuverlässigkeit wird eine homogene Verteilung der Redundanzanteile im Netz angestrebt, was beispielsweise durch Minimierung des maximalen Redundanzanteils als Zielfunktion erreicht werden kann. Der Vektor der Zielfunktionen habe folgende Gestalt:

$$\Phi = (\|\mathbf{Q}_{\hat{x}\hat{x}}\|_2, \|\mathbf{K}\|_F, \|\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_U)\|_\infty)^T, \quad (5.91)$$

wobei  $\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_U)$  den Vektor darstellt, der die Diagonalelemente des Matrixproduktes  $\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_U$  enthält. Das Vektoroptimierungsproblem für das First Order Design lässt sich folgendermaßen darstellen:

$$\begin{cases} \min_{\mathbf{x}} \Phi \\ \underline{\mathbf{x}} < \mathbf{x} < \bar{\mathbf{x}}, \end{cases} \quad (5.92)$$

wobei  $\underline{\mathbf{x}}$  bzw.  $\bar{\mathbf{x}}$  die untere bzw. obere Grenze des Variationsbereichs der Koordinaten angeben. Im Folgenden soll (5.92) durch die Wahl einer Zielfunktion gelöst werden, wobei die Werte der anderen Zielfunktionen unter gewissen Schranken liegen sollen. Als alternativer Ansatz wird eine Lösung mit Normskalarisierung vorgestellt.

#### Lösung der Vektoroptimierung durch Wahl einer Zielfunktion

Betrachten wir als ersten Ansatz zur Lösung der Vektoroptimierung die Wahl einer Zielfunktion. Die Werte der anderen Zielfunktionen werden als Nebenbedingungen eingeführt, so dass garantiert werden kann, dass diese gewisse Schranken nicht übersteigen. Für das folgende Beispiel soll als Zielfunktion der maximale Redundanzanteil minimiert werden, wodurch eine Homogenisierung der Werte erreicht werden kann. Die Werte der Zielfunktionen der Unsicherheit sollen die Startwerte nicht überschreiten, vgl. Abbildungen 5.13. Als zusätzliche Restriktion wird gefordert, dass der minimale Redundanzanteil nach der Optimierung nicht kleiner sein darf als beim Startdesign.

$$\begin{cases} \min_i \max_i \text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_U) \\ \underline{\mathbf{x}} < \mathbf{x} < \bar{\mathbf{x}}, & (a) \\ \|\mathbf{K}\|_F < \|\mathbf{K}^\circ\|_F, & (b) \\ \|\mathbf{Q}_{\hat{x}\hat{x}}\|_2 < \|\mathbf{Q}_{\hat{x}\hat{x}}^\circ\|_2, & (c) \\ \{\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_U)^\circ\} < \min\{\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_U)\}. & (d) \end{cases} \quad (5.93)$$

Verschiedene Untersuchungen haben gezeigt, dass synthetische Netzformen wie Diagonalevierecksketten bereits eine fast optimale Situation bezüglich der Homogenität der Redundanzanteile aufweisen, so dass nur eine geringe Reduktion von  $r_{max} - r_{min}$ , ( $< 4\%$ ) möglich ist. Deshalb wird im Folgenden das Linach-Netz als Beispiel betrachtet. Für die Untersuchungen steht die didaktische Komponente der Netzoptimierung im Vordergrund. Die Ergebnisse sollen eher sekundär als ein Vorschlag für ein optimales Design eines Überwachungsnetzes verstanden werden.

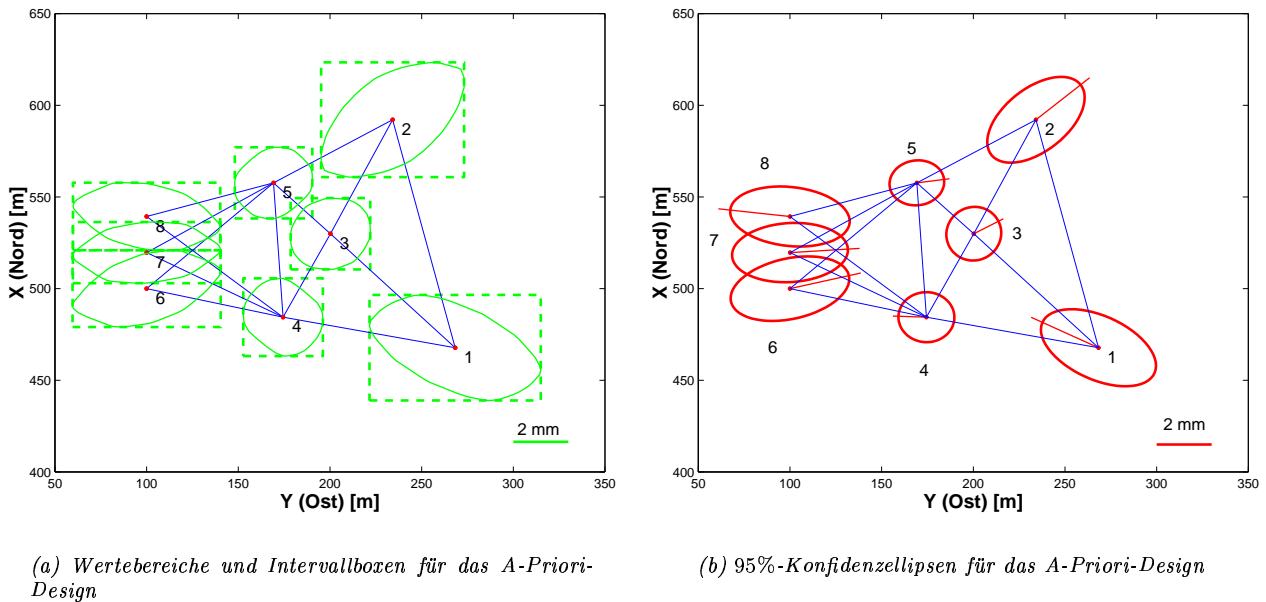


Abbildung 5.12: Darstellung der Punktunsicherheitsbereiche für das Linach-Netz im Ausgangsdesign

Abbildung 5.12 zeigt die Punktunsicherheitsbereiche für das A-Priori-Design des Linach-Netzes bei einer zweidimensionalen Auswertung.

**Beispiel 35** Betrachten wir die Optimierung des Linach-Netzes mit der Zielfunktion (5.93). Die Punktpositionen der Objektpunkte an der Staumauer (Nr.6-8) wurden festgehalten und eine Variation von  $\pm 50 m$  für die Koordinaten der Kontrollpunkte (Nr.1-5) zugelassen.

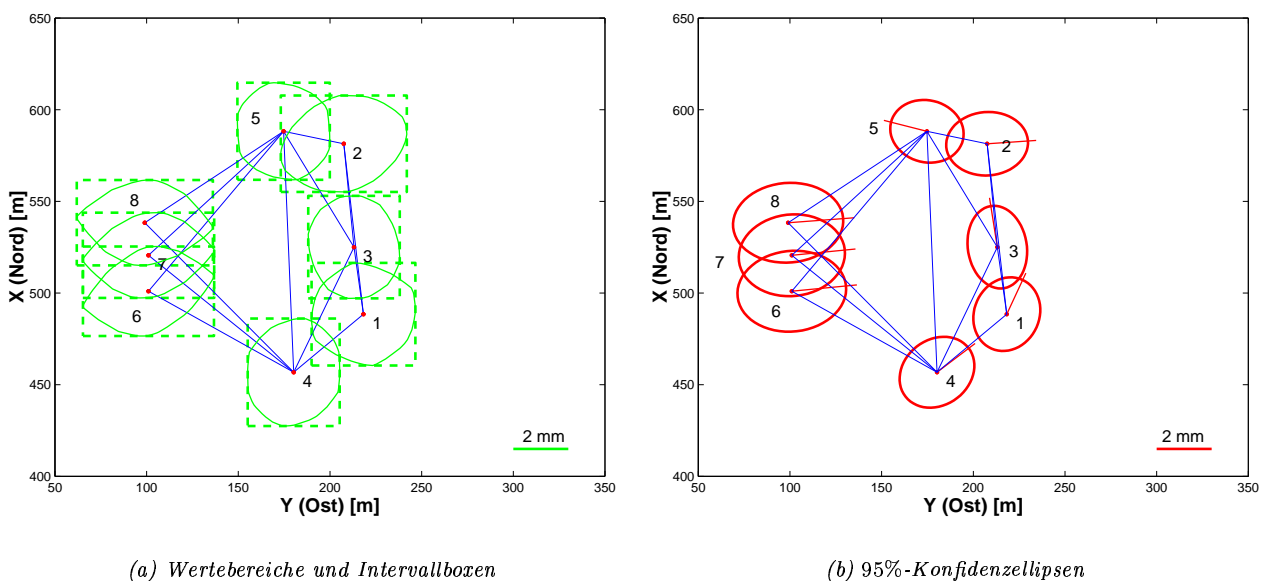


Abbildung 5.13: Vektoroptimierung des Linach-Netzes bei Wahl einer Zielfunktion



Teilabbildung 5.13(a) zeigt die Wertebereiche und Intervallboxen, Teilabbildung 5.13(b) die 95%-Konfidenzellipsen für das optimierte Design. Es lassen sich zwei Effekte feststellen: zum Einen findet eine Bewegung der Kontrollpunkte zur Position des Punktes 3 statt (Kontraktionsanteil). Zum Anderen bewegen sich die Kontrollpunkte (4,5) senkrecht dazu nach außen (Dilatationsanteil). Beide Mechanismen führen dazu, dass für das optimierte Netzdesign die Form und Größe der Zonotope und Konfidenzellipsen homogener werden. Durch die Dilationsbewegung werden günstigere Schnittbedingungen für die Objektpunkte erzielt, vgl. beispielsweise die Abnahme der Elliptizität der Konfidenzellipsen und die Reduktion der Kenngrößen in der Tabelle 5.9. Die Kontraktionsbewegung verbessert die Kenngrößen der Impräzision, deren Größe von der Streckenlänge abhängt, vgl. Tabelle 5.9, fünfte bis siebte Spalte. Die Spannweite der Redundanzanteile kann um 15% reduziert werden, so dass das optimierte Design eine homogenere Verteilung der Redundanzanteile aufweist.

	$\text{spur } \mathbf{Q}_{\hat{\delta}\hat{\delta}}$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{\delta}\hat{\delta}}\ _2$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{\delta}\hat{\delta}}\ _F$ [m <sup>2</sup> ]	$\ \mathbf{K}\ _F$ [mm]	$\ \mathbf{K}\ _2$ [mm]	$\ \mathbf{K}\ _\infty$ [mm]	$r_{max}$ [-]	$r_{min}$ [-]	$r_{max} - r_{min}$ [-]
a priori	5.83	1.80	2.32	1.88	1.12	3.11	0.92	0.09	0.83
aposteriori	5.76	1.31	2.01	1.88	0.96	2.42	0.79	0.09	0.70
Reduktion in [%]	1	27	13	0	14	22	14	0	15

Tabelle 5.9: Veränderung der Werte der Zielfunktionen bei der Vektoroptimierung des Linach-Netzes mit Wahl einer Zielfunktion

### Lösung mit Normskalarisierung

Als zweiter Ansatz sollen zwei Lösungsvarianten mittels Normskalarisierung vorgestellt werden. Die gewichtete Norm des Vektors der Zielfunktionen liefert dabei ein skalares Ersatzproblem für das Vektoroptimierungsproblem (5.92). Zusätzlich gilt, dass jede Lösung des Ersatzproblems auch Lösung des Ausgangsproblems ist, d.h. pareto minimal ist, vgl. GÖPFERT UND NEHSE (1990, S.100ff) bzw. die Diskussion in Abschnitt 5.1.2.

$$\begin{cases} \min_{\mathbf{x}} w_1 \frac{\|\mathbf{Q}_{\hat{\delta}\hat{\delta}}\|_2}{\|\mathbf{Q}_{\hat{\delta}\hat{\delta}}^{\circ}\|_2} + w_2 \frac{\|\mathbf{K}\|_F}{\|\mathbf{K}^{\circ}\|_F} + w_3 \frac{\|\text{vecd}(\mathbf{Q}_{\hat{\delta}\hat{\delta}}\mathbf{P}_{II})\|_\infty}{\|\text{vecd}(\mathbf{Q}_{\hat{\delta}\hat{\delta}}^{\circ}\mathbf{P}_{II})^{\circ}\|_\infty} \\ \underline{\mathbf{x}} < \mathbf{x} < \bar{\mathbf{x}}. \end{cases} \quad (5.94)$$

Die Gewichte  $w_1$ ,  $w_2$ ,  $w_3$  sollen den Einfluss der einzelnen Zielfunktionen steuern. Daher ist es wichtig, die Einheiten bzw. Größenordnungen zu vereinheitlichen, da diese Differenzen sonst implizit in den Gewichten enthalten sind, vgl. XU (1993). Praktisch kann diese Harmonisierung beispielsweise durch Division mit den Werten der Zielfunktionen im Startdesign erhalten werden, vgl. (5.94). Gleichung (5.94) stellt eine ganze Familie an Optimierungsproblemen dar. KUANG (1996, S.250ff) sieht deshalb in Ansätzen der Vektoroptimierung wie (5.94) das allgemeine Modell, aus dem alle anderen Optimierungsprobleme abgeleitet werden können.

**Beispiel 36** Betrachten wir die Optimierung des Linach-Netzes mit der Zielfunktion (5.94). Die Punktpositionen der Objektpunkte an der Staumauer (Nr.6-8) werden festgehalten, eine Variation von  $\pm 50$  m für die Koordinaten der Kontrollpunkte (Nr.1-5) wird jedoch zugelassen. Die Gewichte wurden zu  $w_1 = 3$ ,  $w_2 = 1$  und  $w_3 = 4$  gewählt.

Teilabbildung 5.14(a) zeigt die Wertebereiche und Intervallboxen, Teilabbildung 5.13(b) die 95%-Konfidenzellipsen für das optimierte Design. Es lassen sich zwei Effekte feststellen: zum Einen findet eine Bewegung aller Kontrollpunkte in  $y$ -Richtung hin zur Staumauer statt (Gesamtkontraktion). Zum Anderen bewegen sich die Kontrollpunkte (4,5) in senkrechter Richtung dazu nach außen. Die Grundtendenz dieser Mechanismen wurde bereits im vorherigen Abschnitt beobachtet, allerdings unterscheiden sich die exakten Punktbewegungen. Der erste Effekt ist hauptsächlich durch die Zielfunktionen, die die Unsicherheit (Stochastizität und Impräzision) der Punktpositionen beschreibt, bedingt. Da die Intervallradien streckenlängenabhängig sind, führt eine Reduktion der Punktabstände zu einer Verbesserung dieser Zielfunktionen, vgl. Tabelle 5.10, zweite bis siebte Spalte. Durch die Außenbewegung der Kontrollpunkte kann die Spannweite der Redundanzanteile verringert werden, vgl. die letzten drei Spalten der Tabelle 5.10. Da bei diesem Optimierungsansatz keine Restriktion für den minimalen Redundanzanteil eingeführt wurde, sinkt dieser ebenfalls.

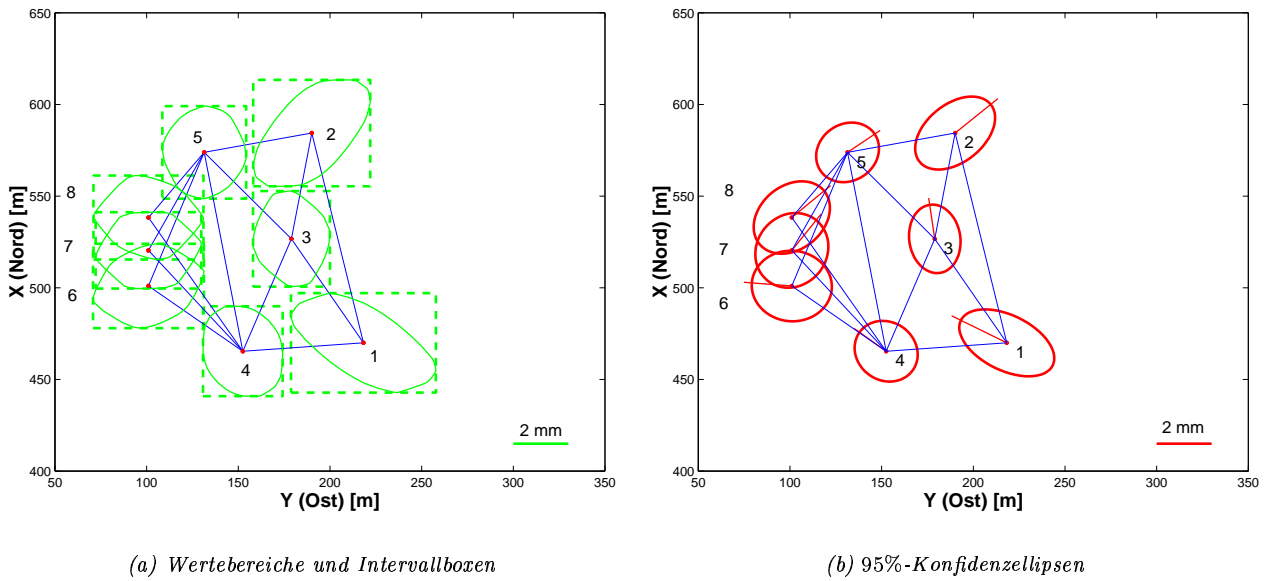


Abbildung 5.14: Vektoroptimierung mittels Normskalarisierung für das Linach-Netz

	$\text{spur } \mathbf{Q}_{\hat{x}\hat{x}}$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{x}\hat{x}}\ _2$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{x}\hat{x}}\ _F$ [m <sup>2</sup> ]	$\ \mathbf{K}\ _F$ [mm]	$\ \mathbf{K}\ _2$ [mm]	$\ \mathbf{K}\ _\infty$ [mm]	$r_{max}$ [-]	$r_{min}$ [-]	$r_{max} - r_{min}$ [-]
a priori	5.83	1.80	2.32	1.88	1.12	3.11	0.92	0.09	0.83
a posteriori	4.61	1.35	1.75	1.66	0.97	2.97	0.83	0.04	0.79
Reduktion in [%]	21	25	24	12	14	5	10	56	5

Tabelle 5.10: Veränderung der Zielfunktionen bei der Vektoroptimierung des Linach-Netzes mit Normskalarisierung

Aus dem Vergleich der Werte der Zielfunktionen a priori und a posteriori wird deutlich, dass eine Verbesserung aller Zielfunktionen erreicht werden kann. Bei der getroffenen Wahl der Gewichte werden die Zielfunktionen, die die Gesamtunsicherheit beschreiben, favorisiert. Die Homogenität der Redundanzanteile kann zwar leicht gesteigert werden, allerdings auf Kosten der Kontrollierbarkeit einiger Beobachtungen, die ebenfalls reduziert wird. Es wäre deshalb sinnvoll, eine Restriktion einzuführen, die ein Mindestniveau der Redundanzanteile garantiert.

Die Diskussion im obigen Beispiel 36 verdeutlicht die Komplexität der Vektoroptimierung. KUANG (1996, S.250ff) schlägt im Gegensatz zum Ansatz der freien Minimierung (5.94) eine „beste“ Anpassung an gegebene Idealwerte vor, vgl. hierzu auch die Diskussion in XU (1993). Dadurch wird die Problemstellung vereinfacht und übersichtlicher. Für die folgenden Untersuchungen soll die Zielfunktion

$$\begin{cases} \min_{\mathbf{x}} w_1 \frac{(\|\mathbf{Q}_{\hat{x}\hat{x}}\|_2 - \|\mathbf{Q}_{\hat{x}\hat{x}}^*\|_2)^2}{\|\mathbf{Q}_{\hat{x}\hat{x}}^*\|_2^2} + w_2 \frac{(\|\mathbf{K}\|_F - \|\mathbf{K}^*\|_F)^2}{\|\mathbf{K}^*\|_F^2} + w_3 \frac{(\|\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_{II})\|_\infty - \|\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_{II})^*\|_\infty)^2}{\|\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_{II})^*\|_\infty^2} \\ \underline{\mathbf{x}} < \mathbf{x} < \bar{\mathbf{x}}. \\ \min\{\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_{II})^*\} < \min\{\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_{II})\} \end{cases} \quad (a) \quad (5.95)$$

betrachtet werden, wobei die Nebenbedingung (b) eingeführt wurde, um ein Mindestniveau der Redundanzanteile zu garantieren. Die Idealwerte werden mit  $\|\mathbf{Q}_{\hat{x}\hat{x}}^*\|_2$ ,  $\|\mathbf{K}^*\|_F$  bzw.  $\|\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_{II})^*\|_\infty$  bezeichnet. Das folgende Beispiel stellt diesen Ansatz vor.

**Beispiel 37** Betrachten wir die Optimierung des Linach-Netzes mit der Zielfunktion (5.95). Die Punktpositionen der Objektpunkte an der Staumauer (Nr.6-8) werden wieder festgehalten und eine Variation von  $\pm 50$  m für die Koordinaten der Kontrollpunkte (Nr.1-5) ist zugelassen. Die Gewichte wurden zu  $w_1 = 1$ ,  $w_2 = 1$  und  $w_3 = 1$  gewählt, d.h. alle Zielfunktionen sollen gleichberechtigt sein. Als Idealwerte  $\|\mathbf{Q}_{\hat{x}\hat{x}}^*\|_2$ ,  $\|\mathbf{K}^*\|_F$  bzw.  $\|\text{vecd}(\mathbf{Q}_{\hat{v}\hat{v}}\mathbf{P}_{II})^*\|_\infty$  wird 75% der Werte der Zielfunktionen im Ausgangsdesign angesetzt.

Teilabbildung 5.15(a) zeigt die Wertebereiche für das Startdesign, Teilabbildung 5.15(b) die entsprechenden 95%-Konfidenzellipsen. Die Punktbewegungen lassen sich wieder in zwei Effekte unterteilen. Im Gegensatz zur gemeinsamen Bewegung aller Kontrollpunkte (1-5) zur Staumauer in  $y$ -Richtung, vgl. Beispiel 36, findet nur eine

Konktraktion der Netzes der Kontrollpunkte statt. Die Werte der  $y$ -Koordinaten der Punkte 4, 5 und damit der Abstand zur Staumauer bleiben konstant. Durch die Konktraktion werden die Punktabstände verringert, was zu einer Reduktion der Impräzision führt. Als zweiter Effekt lässt sich eine Dilatationsbewegung der Kontrollpunkte 4, 5 in  $x$ -Richtung feststellen. Dadurch kann die Bestimmungsgüte der Objektpunkte verbessert, die Impräzision reduziert und die Punktgenauigkeit gesteigert werden. Tabelle 5.11 stellt die Ergebnisse gegenüber.

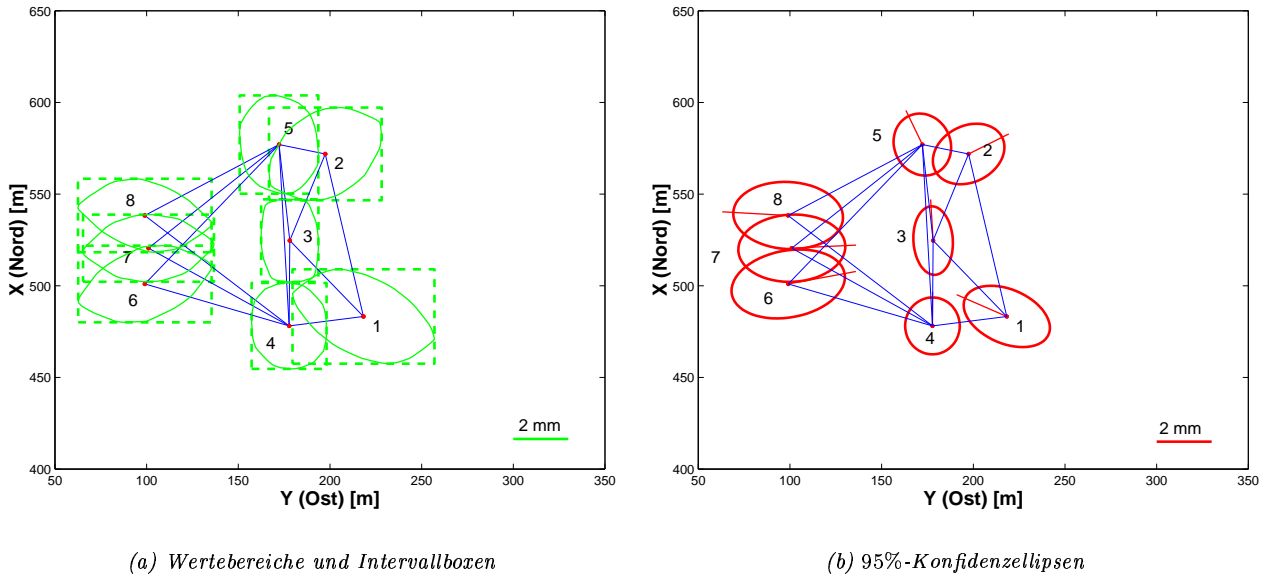


Abbildung 5.15: Vektoroptimierung mittels Normskalarisierung für das Linach-Netz (Anpassung an Idealwerte)

	$spur \mathbf{Q}_{\hat{x}\hat{x}}$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{x}\hat{x}}\ _2$ [m <sup>2</sup> ]	$\ \mathbf{Q}_{\hat{x}\hat{x}}\ _F$ [m <sup>2</sup> ]	$\ \mathbf{K}\ _F$ [mm]	$\ \mathbf{K}\ _2$ [mm]	$\ \mathbf{K}\ _\infty$ [mm]	$r_{max}$ [-]	$r_{min}$ [-]	$r_{max} - r_{min}$ [-]
a priori	5.83	1.80	2.32	1.88	1.12	3.11	0.92	0.09	0.83
aposteriori	5.13	1.38	1.92	1.76	0.98	2.48	0.85	0.09	0.76
Reduktion in [%]	12	23	18	6	12	20	8	0	8

Tabelle 5.11: Veränderung der Werte der Zielfunktionen bei der Vektoroptimierung des Linach-Netzes (Anpassung an Idealwerte)

Durch die Restriktion (b) wird garantiert, dass der Wert des minimalen Redundanzanteils im A-Priori-Design nicht unterschritten werden kann. Es lassen sich alle Werte der Zielfunktionen reduzieren. Ein Vergleich der Reduktionsbeträge, die in den letzten beiden Zeilen der Tabelle 5.11 angegeben sind, zeigt, dass bei einer Gleichgewichtung aller Zielfunktionen der Beitrag des maximalen Eigenwertes der Varianz-Kovarianz-Matrix überwiegt. Hier ist eine Reduktion von über 20% möglich. Die Anteile von  $\|\mathbf{K}\|_F$  und  $spur \mathbf{Q}_{\hat{x}\hat{x}}$  können hingegen nur wenig verbessert werden. Erstaunlich ist, dass der maximale Intervallradius der Koordinaten  $\|\mathbf{K}\|_\infty$  auch um 20% reduziert werden kann, obwohl er explizit weder durch eine Zielfunktion noch durch Nebenbedingungen in (5.92) berücksichtigt wird.

In diesem Abschnitt wurden erste Ansätze einer gemeinsamen Optimierung von Zielfunktionen der Netzgenauigkeit, Zuverlässigkeit und Impräzision mit Verfahren der Vektoroptimierung vorgestellt. Die Beispiele 35-37 sind primär zur Veranschaulichung der verschiedenen, teils gegenläufigen Mechanismen der einzelnen Zielfunktionen zu verstehen. Weitere Untersuchungen wären von Interesse, um das Zusammenspiel der einzelnen Zielfunktionen besser zu verstehen und Empfehlungen für die Wahl des einen oder anderen Optimierungsverfahrens auszusprechen.

## 6. Zusammenfassung

In der vorliegenden Arbeit wurde das Konzept der Beschreibung von Impräzision durch Intervalle, das einen Spezialfall der Beschreibung durch Fuzzy-Sets darstellt, als ein „Baustein für eine neue Fehlertheorie“ (KUTTERER 2002b) vertieft und konkretisiert. Ansatzpunkt ist dabei, dass sich die Unsicherheit geodätischer Messungen aus zwei Hauptbeiträgen, der Stochastizität und der Impräzision, zusammensetzt. Stochastizität beschreibt die zufällige Variabilität der Messwerte. Impräzision wird durch unbekannt systematische Abweichungen zwischen Modell und Daten hervorgerufen. Das Konzept erweitert das klassische der rein stochastischen Beschreibung und liefert durch eine getrennte Betrachtung beider Komponenten realistischere Unsicherheitsmaße. Diese bilden das Fundament für eine aussagekräftige Interpretation von Ergebnissen der geodätischen Datenanalyse. Das inhaltliche Vorgehen lässt sich in drei große Themenkomplexe gliedern, die die Bereiche „Messung“, „Auswertung und Analyse geodätischer Netze“ und „Optimierung geodätischer Messanordnungen“ umfassen und die logische Abfolge der Kapitel der vorliegenden Arbeit vorgeben.

Zunächst wurden im Bereich „Messung“ alle Unsicherheitsbeiträge zur Gesamtunsicherheit geodätischer Messungen gemäß den Vorschlägen des *Guide to the Expression of Uncertainty in Measurements (GMU)*, vgl. DEUTSCHES INSTITUT FÜR NORMUNG (1995), analysiert. Es wurde dabei motiviert, warum nicht für alle Beiträge eine stochastische Modellierung geeignet ist und dass in diesen Fällen die Unsicherheit besser durch Intervalle beschrieben werden kann. Dies gilt insbesondere für Beiträge, deren stochastische Eigenschaften nicht oder nur wenig bekannt sind, wenn z.B. Rundungsfehler betrachtet werden sollen oder man sich für maximale Effekte (worst case) interessiert. Ein weiterer Ansatzpunkt für die Verwendung von Intervallen ist die oft mangelnde Repräsentativität einzelner Messwerte, beispielsweise „zusätzlicher“ Messungen, die zur Erfassung meteorologischer Rahmendaten dienen und für die Korrektur von Refraktionseffekten auf die Messungen benötigt werden. Einerseits kann dabei durch einen Messwert am Instrumentenstandpunkt der Gesamteffekt entlang des Messstrahls nur unvollständig erfasst werden. Andererseits liegen oftmals nur wenige Ablesungen für eine Messung vor, so dass statistische Kenngrößen wie Mittelwerte oder Varianzen nicht aussagekräftig sind. Nach der Bestimmung der Impräzision der einzelnen Beiträge und ihrer Beschreibung durch Intervalle wird diese auf die korrigierten Messungen übertragen, die sich nach Anwendung von Mess- und Auswertemethoden und Anbringen verschiedener Korrekturen aus den originären Messwerten ergeben. Man erhält so Ergebnisintervalle für die korrigierte Messung, die deren Impräzision adäquat ausdrücken.

An verschiedenen Beispielen wurde für terrestrische Messungen wie Richtungen, Strecken und Zenitdistanzen gezeigt, wie konkrete Werte für die Intervalle dieser Messungstypen in einem ersten Schritt aus Fachartikeln über Geräteuntersuchungen oder den Angaben der Gerätehandbücher abgeleitet werden können. Die detaillierten Ergebnisse wurden vereinfachend durch *Faustformeln* dargestellt. Diese sind für das konkrete Arbeiten mit Intervallen als Impräzisionsmaße von zentraler Bedeutung, da sie die Abhängigkeiten der Intervallradien von der Geometrie im Netz (Punktabstände, Zenitdistanzen), von den angewandten Mess- und Auswertemethoden oder Korrekturenmodellen sowie vom verwendeten Instrumentarium einfach in einen klar überschaubaren formalen Zusammenhang bringen.

Darüber hinaus wurden die Unterschiede von Stochastizität und Impräzision herausgearbeitet. Beide Komponenten erfassen je einen charakteristischen Aspekt der Unsicherheit. Erst durch eine getrennte Modellierung ist es möglich, diese wesentlichen, aber unterschiedlichen Eigenschaften angemessen zu behandeln. Dies lässt sich besonders gut für das unterschiedliche Verhalten der Unsicherheitsfortpflanzung zeigen. Die Verwendung von Intervallen impliziert ein alternatives lineares Fortpflanzungsgesetz der Unsicherheit. So kann Impräzision durch Differenzbildung reduziert werden, wohingegen Mehrfachmessungen keine Reduktion zulassen. Beim bekannten Varianz-Kovarianz-Fortpflanzungsgesetz wird die Unsicherheit hingegen quadratisch fortgepflanzt, wodurch zu optimistische Unsicherheitsmaße für die Ergebnisse erzielt werden können, insbesondere bei Mehrfachmessungen. Durch eine Überlagerung beider Komponenten der Unsicherheit wurde das Maß der *erweiterten Unsicherheit* abgeleitet. Es repräsentiert die Gesamtunsicherheit adäquat, da die unterschiedlichen Charakteristika beider Komponenten erhalten bleiben und liefert somit den Ausgangspunkt für eine realitätsnahe Interpretation der Ergebnisse.

Ein zweiter Themenkomplex hat sich mit der Übertragung der Impräzision auf die nach der Methode der kleinsten Quadrate geschätzten Punktkoordinaten geodätischer Netze und einer Beschreibung der Impräzision der Punktpositionen befasst. Hierbei wurde das Konzept der Intervallboxen als Impräzisionsmaße für Punktpositionen durch die Beschreibung des Wertebereichs der Intervallauswertung als *Zonotop* erweitert und verallgemeinert. Wichtige Eigenschaften der Zonotope als spezielle Klasse konvexer Polyeder wurden eingeführt und an anschaulichen Beispielen erläutert. Das Konzept der Zonotope bietet drei wesentliche Vorteile. Erstens wird der Wertebereich direkt durch Zonotope angegeben, wohingegen Intervallboxen eine engstmögliche Einschließung des Wertebereichs liefern. Zweitens sind Zonotope von der Orientierung des Koordinatenrahmens, der für die

Netzausgleichung verwendet wird, unabhängig. Schließlich beschreiben Zonotope die Punktimpräzisionsbereiche detaillierter als Intervallboxen.

Die Unabhängigkeit von der Orientierung der Koordinatensysteme ist für Kenngrößen der Unsicherheit wünschenswert, falls, wie oftmals bei Anwendungen in der Ingenieurgeodäsie, die Orientierung des Koordinatensystems frei wählbar ist. Anhand verschiedener zwei- und dreidimensionaler Beispielnetze wurde veranschaulicht, dass durch Zonotope Detailfragen zur Form und Größe der Punktimpräzisionsbereiche beantwortet werden können. So können beispielsweise Richtungen angegeben werden, in denen die Punktposition die größte oder die geringste Impräzision aufweist.

Abschließend wurden die Unsicherheitsmaße, die getrennt von den Messungen auf die Punktkoordinaten übertragen wurden, zu einem Maß kombiniert. Für Punktpositionen wurden so mit Methoden der Fuzzy-Theorie oder durch einfache Überlagerung von Konfidenzellipsen mit Zonotopen *erweiterte Punkt-Unsicherheitsbereiche* abgeleitet. Alternativ zur Verwendung von Zonotopen wurde ein Konzept von Ellipsoiden als Ersatzformen entwickelt. Dieses basiert auf Löwner-Ellipsoiden oder *umschließenden Ellipsoiden* und ist besonders dann geeignet, wenn die Zonotope aufwändig zu berechnen sind, d.h. wenn im Netz viele Beobachtungen vorkommen.

Ein dritter und letzter Themenbereich hat sich der Optimierung geodätischer Messanordnungen befusst. Hierbei wurden zwei neue Aspekte aufgezeigt. Zum Einen wurden anhand verschiedener Netzbeispiele Fragestellungen der klassischen Netzoptimierung nach optimaler Netzkonfiguration, Gewichtung und Datumswahl für die Kenngrößen der Impräzision beantwortet. Es ist festzuhalten, dass sich die spektralen Kenngrößen der Impräzision in Bezug auf eine Variation der Netzgeometrie (First Order Design) für verschiedene Netze ähnlich verhalten wie globale Kenngrößen der Stochastizität. Der Grad diese *Ähnlichkeit* lässt sich aus der Struktur einer Rotationsmatrix ablesen, die die Transformation der Basis der Varianz-Kovarianz-Matrix auf die Basis der Matrix der Kantenvektoren mathematisch beschreibt. Eine optimale Datumsgebung in Bezug auf die Kenngrößen der Impräzision wird bei freien Netzen ebenfalls für die innere Datumsgebung erreicht.

Zum Anderen wurden Lösungsansätze für das *Intervalldesign* vorgestellt. Im Intervalldesign wird untersucht, wie durch eine geschickte Wahl der Mess- und Auswertemethoden, der Sensoren und Korrekptionsmodelle erreicht wird, Werte für die Impräzision der geschätzten Koordinaten zu erhalten, die kleiner als Maximalvorgaben sind. Hierfür wurde ein zweistufiges Lösungskonzept vorgestellt und exemplarisch für verschiedene Konfigurationen getestet. Neben der Kontrolle oder *Garantie oberer Schranken* für die Impräzision kann versucht werden, die Impräzision so zu reduzieren, dass ihr Wert vernachlässigbar klein wird. Für die meisten Netzbeispiele wurde gezeigt, dass die Impräzision jedoch nicht vernachlässigbar ist und somit beide Komponenten der Unsicherheit betrachtet werden müssen. Abschließend wurden erste Ansätze zur gemeinsamen Optimierung klassischer Zielgrößen wie Netzgenauigkeit und Zuverlässigkeit mit den neuen Zielgrößen der Impräzision mittels Verfahren der Vektoroptimierung vorgestellt. Hier sind weitere Untersuchungen wichtig, um die Reaktionsmuster der einzelnen Komponenten und ihr Zusammenspiel besser zu verstehen. Darüber hinaus kann der Vektor der Zielfunktionen um zusätzliche Kenngrößen ergänzt werden wie beispielsweise die Sensitivität geodätischer Überwachungsnetze.

Ein wichtiges Ergebnis des dritten Themenbereichs ist der Vorschlag zu einer neuen Einteilung der freien Parameter der geodätischen Netzoptimierung in unabhängige Parametergruppen. Hierfür wurde ein *Konfigurationsdesign* und ein *Unsicherheitsdesign* vorgestellt. Ersteres umfasst die Variation der Netztopologie und der Lage der Netzpunkte. Das Unsicherheitsdesign beschreibt den Einfluss der Stochastizität und Impräzision der Beobachtungen auf die Werte der Zielfunktionen. Stochastizität und Impräzision werden dabei durch die Performance-Parameter der Instrumente bzw. durch die Koeffizienten der Faustformeln parametrisiert. Durch diese Trennung treten indirekte Effekte nicht mehr auf, die beispielsweise durch die Abhängigkeit der Gewichte der Beobachtungen vom Netzdesign entstehen.

Für Anwendungen im Bereich der Deformationsanalyse sollte das vorgestellte Konzept der erweiterten Unsicherheit um statistische Tests ergänzt werden. Dadurch können beispielsweise Punktverschiebungen auf Signifikanz in Bezug auf die Gesamtunsicherheit getestet werden. Für eindimensionale Tests lassen sich beispielsweise die für fuzzy-erweiterte Größen formulierten statistischen Tests (KUTTERER 2003) auf (intervall-) erweiterte Unsicherheitsmaße übertragen. Für mehrdimensionale Tests, die z.B. für das Aufdecken von Punktverschiebungen notwendig sind, müssen entsprechende Testverfahren entwickelt werden. Dadurch können Fragestellungen der geodätischen Deformationsanalyse aus einer neuen Perspektive betrachtet werden.

Eine prinzipielle Verbesserung der Qualitätsangaben ist für alle geodätischen Raumverfahren zu erwarten, insbesondere auch für GPS, das kein direkter Gegenstand der Arbeit war. Die Varianz-Kovarianz-Matrix der geschätzten Parameter, die bei einer klassischen Ausgleichung berechnet wird, weist für die genannten Beobachtungsverfahren eher formalen Charakter auf. Dies liegt beispielsweise darin begründet, dass zum Einen in der Regel sehr viele Beobachtungen an der Schätzung beteiligt sind und Korrelationen zwischen den Beobachtungen vernachlässigt werden. Zum Anderen wird oftmals keine durchgängige und konsequente Trennung in systema-

tisch wirkende Effekte und solche mit zufälligem Charakter durchgeführt. Das vorgestellte Vorgehen wird sich deshalb für diese Anwendungen besonders eignen. Zusätzlich können die Repräsentativität der verwendeten Korrektionsmodelle berücksichtigt und der Einfluss unterschiedlicher Auswertetechniken quantifiziert werden.

Weitere Anwendungsgebiete sind im Bereich der Kalibrierung geodätischer Messmittel und bei der Qualitätssicherung zu sehen. Hier liefert das vorgestellte Verfahren neue Möglichkeiten zur Angabe der Unsicherheit und kann als ein Impuls für die aktuellen Diskussionen gesehen werden, wie sie beispielsweise in HEISTER UND STAIGER (2001) geführt werden. Einerseits werden die Empfehlungen des Leitfadens für die Qualifizierung und Identifizierung von Unsicherheit verwendet. Andererseits geht das Verfahren über eine einheitliche stochastische Betrachtung aller Beiträge zur Gesamtunsicherheit hinaus: Durch eine konsequent getrennte Behandlung der identifizierten und qualitativ unterschiedlichen Effekte (Stochastizität und Impräzision) wird eine klare Struktur der Fortpflanzung der Unsicherheit gegeben. Damit wird die Unsicherheit und Repräsentativität der Ergebnisgrößen nachvollziehbar und einschätzbar. Erste Ansätze hierzu sind in den Faustformeln verwirklicht. Die vorliegende Arbeit hat somit ein Gesamtkonzept für die erweiterte Unsicherheit vorgestellt und exemplarisch gezeigt, dass dieses geeignet ist, die Gesamtunsicherheit geodätischer Beobachtungen und Schätzwerte adäquat zu beschreiben. Dies ist eine wesentliche Voraussetzung, um die Ergebnisse geodätischer Messungen und Datenanalyse realistisch einschätzen und interpretieren zu können.

## Literatur

- ALEFELD, G. und J. HERZBERGER (1974): *Einführung in die Intervallrechnung*, Reihe Informatik, Bd. 12, Bibliographisches Institut, Mannheim Wien Zürich.
- ALEFELD, G. und J. HERZBERGER (1983): *Introduction to Interval Computations*, Computer Science and Applied Mathematics, Academic Press, Boston San Diego New York.
- BAARDA, W. (1968): *A Testing Procedure for Use in Geodetic Networks*, Publications on Geodesy - New Series - 2(5), Netherlands Geodetic Commission, Delft.
- BAARDA, W. (1973): *S-Transformations and Criterion Matrices*, Publications on Geodesy - New Series - 5(1), Netherlands Geodetic Commission, Delft.
- BANDEMER, H. (1997): *Ratschläge zum mathematischen Umgang mit Unwissenheit - Reasonable Computing*, B. G. Teubner, Stuttgart Leipzig.
- BANDEMER, H. und S. GOTTWALD (1993): *Einführung in Fuzzy-Methoden - Theorie und Anwendungen unscharfer Mengen*, 4. Aufl., Akademie Verlag, Berlin.
- BEHREND, F. (1938): *Über die kleinste umbeschriebene und die größte einbeschriebene Ellipse eines konvexen Bereichs*, Mathematische Annalen, 115(3): 379–411.
- BILL, R. (1985): *Kriteriummatrizen ebener geodätischer Netze*, DGK Reihe A 102, Deutsche Geodätische Kommission, München.
- BILL, R., R. JÄGER und G. SCHMITT (1984): *Effekte in langgestreckten Netzen und ihre statischen Analogien*, Zeitschrift für Vermessungswesen, 109(10): 526–540.
- BJÖRNER, A., M. L. VERGNAS, B. STURMFELS, N. WHITE und G. M. ZIEGLER (1999): *Oriented Matroids*, 2. Aufl., Encyclopedia of mathematics and its applications, Bd. 46, Cambridge University Press, Cambridge New York Port Chester.
- BÖCKEM, B., P. FLACH, A. WEISS und M. HENNES (2000): *Refraction Influence Analysis and Investigations on Automated Elimination of Refraction Effects on Geodetic Measurements*, In: *Proceedings of the XVI<sup>th</sup> IMEKO World Congress 2000*, Kluwer Academic Publishers, Dordrecht.
- BONNESEN, T. und W. FENCHEL (1934): *Theorie der konvexen Körper*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Bd. 3(1), J. Springer, Berlin.
- BRAEMS, I. (2002): *Méthodes ensemblistes garanties pour l'estimation de grandeurs physiques*, Dissertation - Laboratoire des Signaux et Systèmes, Universität Paris-Süd.
- BRAEMS, I., M. KIEFFER und E. WALTER (2003): *Prior characterization of the performance of software sensors*, In: *Proceedings of the 13<sup>th</sup> IFAG Symposium on System Identification, SysId2003*.
- BROCKS, K. (1939): *Vertikaler Temperaturgradient und terrestrische Refraktion, insbesondere im Hochgebirge*, Veröffentlichungen des Meteorologischen Instituts der Universität Berlin, Bd. 3(4), Reimer, Andrews & Steiner, Berlin.
- BRUNNER, F. K. und H. WOSCHITZ (2001): *Kalibrierung von Messsystemen: Grundlagen und Beispiele*, In: HEISTER, H. und R. STAIGER (Hrsg.), *Qualitätsmanagement in der geodätischen Messtechnik*, 70–90, Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 42, Verlag Konrad Wittwer, Stuttgart.
- CASPARY, W. und K. WICHMANN (1994): *Lineare Modelle - Algebraische Grundlagen und statistische Anwendungen*, Lehr- und Handbücher der Statistik, R. Oldenbourg Verlag, München Wien.
- CIDDOR, P. E. (1996): *Refractive index of air: new equations for the visible and near infrared*, Applied Optics, 35(9): 1566–1573.
- CIDDOR, P. E. und R. J. HILL (1999): *Refractive index of air, 2. Group index*, Applied Optics, 38(9): 1663–1667.
- COLDITZ, S. (1997): *Untersuchung zur Flexibilität und Parameterunsicherheit bei verfahrenstechnischen Prozessen*, Fortschrittsberichte VDI Reihe 3 Verfahrenstechnik Nr.474, VDI-Verlag, Düsseldorf.
- COMBA, J. L. D. und J. STOLFI (1993): *Affine Arithmetic and its Applications to Computer Graphics*, In: *Anais do VI Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens (SIB-GRAPI'93)*, 9–18.
- COOPER, M. A. (1987): *Modern Theodolites and Levels*, 2. Aufl., Oxford BSP Professional Books, Oxford London Edinburgh.
- COXETER, H. (1973): *Regular Polytopes*, 3. Aufl., Dover Publications Inc., New York.
- DANZER, L., D. LAUGWITZ und H. LENZ (1957): *Über das Löwnersche Ellipsoid und sein Analogon unter den einem Eikörper einbeschriebenen Ellipsoiden*, Archiv der Mathematik, 8: 214–219.
- DEICHL, K. (1984): *Der Brechungsindex für Licht und Mikrowellen*, Allgemeine Vermessungsnachrichten, 91(3): 85–100.
- DEUMLICH, F. und R. STAIGER (2002): *Instrumentenkunde der Vermessungstechnik*, 9. Aufl., Herbert Wichmann Verlag, Heidelberg.
- DEUSSEN, D. (2000): *Meßverfahren zur Erfassung der Vertikalrefraktion unter Nutzung der atmosphärischen Turbulenz*, Geodäsie, Bd.4, Shaker, Aachen.
- DEUTSCHES INSTITUT FÜR NORMUNG (Hrsg.) (1995): *Leitfaden zur Angabe der Unsicherheit beim Messen*, Beuth Verlag, Berlin Wien Zürich.
- DEUTSCHES INSTITUT FÜR NORMUNG (1998): *DIN 1319-3 Auswertung von Messungen einer einzelnen Meßgröße*, In: DEUTSCHES INSTITUT FÜR

- NORMUNG (Hrsg.), *DIN-Taschenbuch 111 - Vermessungswesen*, 79–102, Beuth Verlag, Berlin Wien Zürich.
- DINTER, G. (2002): *Generalisierte Orthogonalzerlegungen in der Ausgleichsrechnung*, DGK Reihe C 559, Deutsche Geodätische Kommission, München.
- DRIXLER, E. (1993): *Analyse der Form und Lage von Objekten im Raum*, DGK Reihe C 409, Deutsche Geodätische Kommission, München.
- DUBOIS, D. und H. PRADE (1980): *Fuzzy Sets and Systems*, Plenum Press, New York London.
- DUPRAZ, H. und W. NIEMEIER (1979): *Un critère pour l'analyse des réseaux géodésiques de contrôle*, *Vermessung, Photogrammetrie, Kulturtechnik*, 77(4): 70–76.
- EPPSTEIN, D. (1995): *Zonohedra and Zonotopes*, Tech. Report 95-53, <http://www.ics.uci.edu/eppstein>.
- FANJOR, S. (1985): *Das geodätische Modell und die Theorie der Messung*, *Allgemeine Vermessungsnachrichten*, 92(8-9): 323–329.
- FAVRE, C. und M. HENNES (2000): *Zum Einfluss der geometrischen Ausrichtung von 360°-Reflektoren bei Messungen mit automatischer Zielerfassung*, *Vermessung, Photogrammetrie, Kulturtechnik*, 98(2): 72–78.
- DE FIGUIEREDO, L. H. und J. STOLFI (1997): *Métodos numéricos auto-validados e aplicações - Self-Validated Numerical Methods and Applications*, Instituto de matemática pura e aplicada, Rio de Janeiro.
- DE FIGUIEREDO, L. H., J. STOLFI und L. VELHO (2003): *Approximating Parametric Curves with Strip Trees using Affine Arithmetic*, *Computer Graphics Forum*, 22(2): 171–179.
- FILLIMAN, P. (1988): *Extremum problems for zonotopes*, *Geometriae Dedicata*, 27: 251–262.
- FLACH, P. und D. NATEROP (1999): *Neue Analysetechniken für Deformationsmessungen in permanenten Robotertachymeter-Netzen*, *Allgemeine Vermessungsnachrichten*, 106(8-9): 284–291.
- FUNCKE, G. (1982): *Verfahren zur Parameterelimination im Gauß-Markoff-Modell und deren Einfluss auf ausgeglichene Beobachtungen*, *Allgemeine Vermessungsnachrichten*, 89(3): 112–122.
- GOODMAN, J. E. und J. O'ROURKE (Hrsg.) (1997): *Handbook of Discrete and Computational Geometry*, *Discrete Mathematics and its Applications*, CRC Press, Boca Raton New York London.
- GÖPFERT, A. und R. NEHSE (1990): *Vektoroptimierung - Theorie, Verfahren und Anwendungen*, *Mathematisch-Naturwissenschaftliche Bibliothek*, Bd.74, B. G. Teubner, Leipzig.
- GOTTWALD, R. (1985): *Zur Genauigkeitssteigerung und Erstellung eines automatisierten Datenflusses beim trigonometrischen Nivellement mit kurzen Zielweiten*, Veröffentlichungen des Geodätischen Instituts der Rheinisch-Westfälischen Technischen Hochschule Aachen, Bd. 37.
- GRAFAREND, E. W. (1972): *Genauigkeitsmaße geodätischer Netze*, DGK, Reihe A 73, Deutsche Geodätische Kommission, München.
- GRAFAREND, E. W. (1974): *Optimisation of Geodetic Networks*, *Bollettino di Geodesia e Scienze Affini*, 33(4): 352–406.
- GRAFAREND, E. W., H. HEISTER, R. KELM, H. KROPFF und B. SCHAFFRIN (1979): *Optimierung geodätischer Messoperationen*, Serie Wichmann - Neue Folge, Bd. 11, H. Wichmann Verlag, Karlsruhe.
- GRAFAREND, E. W. und F. SANSÒ (Hrsg.) (1985): *Optimization and Design of Geodetic Networks*, Springer, Berlin Heidelberg New York.
- GRAFAREND, E. W. und B. SCHAFFRIN (1993): *Ausgleichsrechnung in linearen Modellen*, Bibliographisches Institut, Mannheim Wien Zürich.
- GRIMHARDT, H. (1984): *Berechnung minimaler Wertebereiche für die Parameterfunktionen eines Lagenetzes aus den vorgegebenen Toleranzen der Beobachtungen*, DGK Reihe C 300, Deutsche Geodätische Kommission, München.
- GROTEN, E. (2000): *Parameters of Common Relevance of Astronomy, Geodesy, and Geodynamics*, *Journal of Geodesy*, 74(1): 134–140.
- GRÖTSCHHEL, M., L. LOVÁSZ und A. SCHRIJVER (1988): *Geometric Algorithms and Combinatorial Optimization*, 2. Aufl., Springer, Berlin Heidelberg New York.
- GRUBER, P. M. und J. M. WILLIS (Hrsg.) (1993): *Handbook of Convex Geometry - 2 Bde*, Elsevier, Amsterdam, Lausanne, New York.
- GRÜNBAUM, B. (1967): *Convex Polytopes*, *Pure and Applied Mathematics*, Bd. 16, Interscience Publisher, London New York Sydney.
- HADWIGER, H. (1957): *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Bd. XCIII, Springer, Berlin Göttingen Heidelberg.
- HAMPEL, F., E. R. RONCHETTI, P. ROUSSEEUW und W. STAHEL (1986): *Robust Statistics: the approach based on influence functions*, John Wiley & Sons.
- HANSEN, E. R. (1975): *A generalized interval arithmetic*, In: NICKEL, K. (Hrsg.), *Interval Mathematics*, 7–18, *Lecture Notes in Computer Sciences*, Bd. 29, Springer, Berlin Heidelberg New York.
- HECK, B. (1981): *Der Einfluss einzelner Beobachtungen auf das Ergebnis einer Ausgleichung und die Suche grober Fehler*, *Allgemeine Vermessungsnachrichten*, 88(1): 17–34.
- HECK, B. (1995): *Rechenverfahren und Auswertemodelle in der Landesvermessung*, 2. Aufl., H. Wichmann Verlag, Karlsruhe.
- HECK, B. und R. JÄGER (1986): *Zur Sensitivität von Strecken- und Streckenverhältnismessungen in Deformationsnetzen*, *Zeitschrift für Vermessungswesen*, 111(10): 459–468.
- HECK, B., H. KUTTERER, K. LINDNER und M. MAYER (1996): *Application of the Spectral Analysis and Sensitivity Analysis for the Design and Analysis of GPS Networks*, In: DIETRICH, R. (Hrsg.), *The Geodetic Antarctic Project GAP95*



- *German Contributions to the SCAR 95 Epoch Approach*, 53–72, DGK Reihe B 304, Deutsche Geodätische Kommission, München.
- HEINDL, G. und E. REINHART (1976a): *Ausgleichung im Sinne minimaler Maximalfehler*, DGK Reihe A 84, Deutsche Geodätische Kommission, München.
- HEINDL, G. und E. REINHART (1976b): *Eine allgemeine Methode zur Berechnung von MINIMAX-Fehlern, Teil 1: Bei vorliegenden Messungen*, Zeitschrift für Vermessungswesen, 101(4): 126–132.
- HEINDL, G. und E. REINHART (1976c): *Eine allgemeine Methode zur Berechnung von MINIMAX-Fehlern, Teil 2: Vor der Durchführung von Messungen*, Zeitschrift für Vermessungswesen, 101(6): 238–241.
- HEINDL, G. und E. REINHART (1978): *Eine allgemeine Methode zur Berechnung von MINIMAX-Fehlern, Teil 3: Bei teils vorliegenden, teils projektierten Messungen*, Zeitschrift für Vermessungswesen, 103(4): 150–155.
- HEINE, K. (1999): *Beschreibung von Deformationsprozessen durch Volterra- und Fuzzy-Modelle sowie Neuronale Netze*, DGK Reihe C 516, Deutsche Geodätische Kommission, München.
- HEISTER, H. und R. STAIGER (Hrsg.) (2001): *Qualitätsmanagement in der geodätischen Messtechnik*, Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 42, Verlag Konrad Wittwer, Stuttgart.
- HENK, M., J. RICHTER-GEBERT und G. M. ZIEGLER (1997): *Basic Properties of convex Polytopes*, In: GOODMAN, J. E. und J. O'ROURKE (Hrsg.), *Handbook of Discrete and Computational Geometry*, 243–270, Discrete Mathematics and its Applications, CRC Press, Boca Raton New York London.
- HENNES, M. (1999): *Grundlegende Aspekte zur Bestimmung der Leistungsfähigkeit von Robottachymetern*, Allgemeine Vermessungsnachrichten, 106(11/12): 374–385.
- HENNES, M., R. DÖNICKE und H.-P. CHRIST (1999): *Zur Bestimmung der temperaturgradienteninduzierten Richtungsverschwenkung beim Tunnelvortrieb*, Vermessung, Photogrammetrie, Kulturtechnik, 97(8): 418–426.
- HENNES, M. und H. INGENSAND (2000): *Komponentenkalibrierung versus Systemkalibrierung*, In: SCHNÄDELBACH, K. und M. SCHILCHER (Hrsg.), *Ingenieurvermessung 2000: Beiträge zum XIII. Kurs für Ingenieurvermessung*, 166–177, Vermessungswesen bei Konrad Wittwer Bd. 33, Verlag Konrad Wittwer, Stuttgart.
- HENNES, M., M. JURETZKO und B. WITTE (1994): *Die prüfstreckenabhängige Modellierung von Additionskonstanten elektrooptischer Distanzmesser*, Allgemeine Vermessungsnachrichten, 101(4): 121–132.
- HENNES, M. und B. KRICKEL (2000): *Zur Entwicklung von Untersuchungsverfahren zur Bestimmung der Leistungsfähigkeit von Robottachymetern*, Flächenmanagement und Bodenordnung, 1(1): 26–33.
- HOFMANN-WELLENHOF, B., H. LICHTENEGGER und J. COLLINS (2001): *Global Positioning System - Theory and Practice*, 5. Aufl., Springer, Wien New York.
- HUBER, P. (1981): *Robust Statistics*, John Wiley & Sons.
- ILLNER, I. (1983): *Freie Netze und S-Transformation*, Allgemeine Vermessungsnachrichten, 90(5): 157–170.
- ILLNER, M. (1986): *Anlage und Optimierung von Verdichtungsnetzen*, DGK Reihe C 317, Deutsche Geodätische Kommission, München.
- INGENSAND, H. (2001): *Systematische Einflüsse auf praktische Messungen mit dem Tachymeter und Digitalnivellier*, In: HEISTER, H. und R. STAIGER (Hrsg.), *Qualitätsmanagement in der geodätischen Messtechnik*, 120–137, Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 42, Verlag Konrad Wittwer, Stuttgart.
- JÄGER, R. (1985): *Zur Anwendung von Streckenverhältnisbeobachtungen in Überwachungsnetzen und auf Eichstrecken*, Allgemeine Vermessungsnachrichten, 92(2): 53–65.
- JÄGER, R. (1988): *Analyse und Optimierung geodätischer Netze nach spektralen Kriterien und mechanischen Analogien*, DGK Reihe C 342, Deutsche Geodätische Kommission, München.
- JÄGER, R. und R. BILL (1985): *Genauigkeits- und Zuverlässigkeitsmaße in Verdichtungsnetzen*, In: *Beurteilung geodätischer Netze*, 70–98, Sonderheft der Mitttelungen des Deutscher Verein für Vermessungswesen- Landesverein Baden-Württemberg, 33. Jg.
- JÄGER, R. und H. KALTENBACH (1990): *Spectral Analysis and Optimization of Geodetic Networks based on Eigenvalues and Eigenfunctions*, Manuscripta Geodaetica, 15: 302–311.
- JAULIN, L., M. KIEFFER, O. DIDRIT und E. WALTER (2001): *Applied Interval Analysis*, Springer, London Berlin Heidelberg.
- JOECKEL, R. und M. STÖBER (1999): *Elektronische Entfernung- und Richtungsmessung*, 4. Aufl., Vermessungswesen bei Konrad Wittwer, Bd. 18, Verlag Konrad Wittwer, Stuttgart.
- JOHN, F. (1948): *Extremum problems with inequalities as subsidiary conditions*, In: *Studies and essays - Courant anniversary volume*, 187–204, Interscience Publisher, London New York Sydney, vgl.: J. Moser (1985): *Fritz John - Collected Papers*, Bd. 2, 543–560, Birkhäuser, Boston.
- JUHNKE, F. (1990): *Volumenminimale Ellipsoidüberdeckungen*, Algebra und Geometrie, 30: 143–153.
- KAHMEN, H. (1993): *Vermessungskunde*, 18. Aufl., Walter de Gruyter, Berlin New York.
- KALTENBACH, H. (1992): *Optimierung geodätischer Netze mit spektralen Zielfunktionen*, DGK Reihe C 393, Deutsche Geodätische Kommission, München.
- KALTENBACH, H. und G. SCHMITT (1988): *A New Approach for criterion matrices based on graph theory*, Manuscripta Geodaetica, 13: 296–305.

- KERN AG (1986): *Bedienungsanleitung Präzisionsdistanzmesser Mekometer ME 5000*.
- KIEFFER, M., L. JAULIN, E. WALTER und D. MEIZEL (2000): *Robust Autonomous Robot Localization Using Interval Analysis*, *Reliable Computing*, 6: 337–362.
- KLATTE, R. (1993): *C-XSC: a C++ class library for extended scientific computing*, Springer, New York Berlin Heidelberg.
- KLATTE, R., U. KULISCH, M. NEAGA, D. RATZ und C. ULLRICH (1991): *PASCAL-XSC- Sprachbeschreibung mit Beispielen*, Springer, New York Berlin Heidelberg.
- KLOBUCHAR, J. (1996): *Ionospheric effects on GPS*, In: PARKINSON, B. W. und J. J. SPILKER (Hrsg.), *Global Positioning System: Theory and Applications*, 485–515, *Progress in Astronautics and Aeronautics*, Bd.163, American Society of Aeronautics and Astronautics.
- KOCH, K. R. (1997): *Parameterschätzung und Hypothesentests*, 3. Aufl., Ferd. Dümmler Verlag, Bonn.
- KOCH, K.-R. (2000): *Einführung in die Bayes-Statistik*, Springer, Berlin Heidelberg New York.
- KREINOVICH, V. (1995): *Data Processing Beyond Traditional Statistics: Applications of Interval Computations. A Brief Introduction*, Supplement to the international journal of Reliable Computing, 13–21.
- KREYSZIG, E. (1979): *Statistische Methoden und ihre Anwendungen*, 7. Aufl., Vanderhoek & Ruprecht, Göttingen.
- KUANG, S. (1993): *Quality Control of GPS for Tunneling: Precision and Reliability Aspects*, *Zeitschrift für Vermessungswesen*, 118(7): 329–345.
- KUANG, S. (1996): *Geodetic Network Analysis and Optimal Design*, Ann Arbor Press, Inc., Chelsea, Michigan.
- KÜHN, W. (1998): *Zonotope Dynamics in Numerical Quality Control*, In: HEGE, H.-C. und K. POLTHIER (Hrsg.), *Mathematical Visualization, Algorithms, Applications and Numerics*, 125–134, Springer, Berlin Heidelberg New York.
- KULISCH, U. (1976): *Grundlagen des numerischen Rechnens: mathematische Begründung der Rechnerarithmetik*, Reihe Informatik, Bd. 19, Bibliographisches Institut, Mannheim Wien Zürich.
- KULISCH, U. (Hrsg.) (1989): *Wissenschaftliches Rechnen mit Ergebnisverifikation - eine Einführung*, Ferd. Vieweg & Sohn, Braunschweig Wiesbaden.
- KULISCH, U. und W. MIRKANKER (1983): *A New Approach to Scientific Computations*, Academic Press, Boston San Diego New York.
- KURZHANSKI, A. und I. VÁLYI (1997): *Ellipsoidal Calculus for Estimation and Control*, Birkhäuser Verlag, Boston Basel Berlin.
- KUTTERER, H. (1994): *Intervallmathematische Behandlung endlicher Unschärfen linearer Ausgleichungsmodelle*, DGK Reihe C 423, Deutsche Geodätische Kommission, München.
- KUTTERER, H. (1998): *Quality aspects of a GPS reference network in Antarctica - a simulation study*, *Journal of Geodesy*, 72(2): 51–63.
- KUTTERER, H. (2002a): *Joint treatment of random variability and imprecision in GPS data analysis*, *Journal of Global Positioning Systems*, 1(2): 96–105.
- KUTTERER, H. (2002b): *Zum Umgang mit Ungewissheit in der Geodäsie - Bausteine für eine neue Fehlertheorie*, DGK Reihe C 553, Deutsche Geodätische Kommission, München.
- KUTTERER, H. (2003): *Statistical hypothesis tests in case of imprecise data*, In: SANSÒ, F. (Hrsg.), *Proceedings of the 5. Hotine-Marussi-Symposium*, International Association of Geodesy Symposia, in press, Springer, Berlin Heidelberg New York.
- LEICA GEOSYSTEMS (1998): *TPS1100 Professional Series- Gebrauchsanweisung, Version 1.2*.
- LEICA GEOSYSTEMS (2000): *Gebrauchsanweisung TPS System - Version 2.2*.
- LEICHTWEISS, K. (1959): *Über die affine Exzentrizität konvexer Körper*, *Archiv der Mathematik*, 10: 187–199.
- LINHART, J. (1986): *Extremaleigenschaften der regulären 3-Zonotope*, *Studia Scientiarum Mathematicarum, Hungarica*, 21: 181–188.
- LINHART, J. (1988): *Über die Varianz der Breite von Zonotopen*, *Beiträge zur Algebra und Geometrie*, 27: 55–62.
- LOHNER, R. (1989): *Einschließungen bei Anfangs- und Randwertaufgaben gewöhnlicher Differentialgleichungen*, In: KULISCH, U. (Hrsg.), *Wissenschaftliches Rechnen mit Ergebnisverifikation - eine Einführung*, 183–208, Ferd. Vieweg & Sohn, Braunschweig Wiesbaden.
- LUTES, J. (2001): *Automated dam displacement monitoring using a robotic total station*, Technical Report, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Bd. 214.
- MATTHIAS, H. (1961): *Umfassende Behandlung der Theodolitachsenfehler auf vektorieller Grundlage unter spezieller Berücksichtigung der Taumelfehler der Kippachse*, *Berichte des IGP*, Bd.10, Verlag Leemann, Zürich.
- MAYER, G. (1989): *Grundbegriffe der Intervallrechnung*, In: KULISCH, U. (Hrsg.), *Wissenschaftliches Rechnen mit Ergebnisverifikation - eine Einführung*, 101–117, Ferd. Vieweg & Sohn, Braunschweig Wiesbaden.
- MEIER-HIRMER, B. und R. MEIER-HIRMER (1997): *Untersuchungen zur intrinsischen Genauigkeit der Digitalnivelliere NA2000 / NA3003 und DiNi20 / DiNi10*, *Vermessungswesen und Raumordnung*, 59(5+6): 286–295.
- V. MIERLO, J. (1981): *Second order design: precision and reliability aspects*, *Allgemeine Vermessungsnachrichten*, 88(3): 95–101.
- V. MIERLO, J. (1982): *Difficulties in Defining the Quality of Geodetic Networks*, In: BORRE, K. und W. WELSCH (Hrsg.), *Proceedings Survey Control Networks*, 259–274, Schriftenreihe der Hochschule der Bundeswehr München, Bd.7.

- MILANESE, M., J. NORTON, H. PIET-LAHANIER und E. WALTER (Hrsg.) (1996): *Bounding Approaches to System Identification*, Plenum Press, New York London.
- MOORE, R. E. (1969): *Intervallanalyse*, R. Oldenbourg Verlag, München Wien.
- MORITZ, H. (1961): *Zur Reduktion elektronisch gemessener Strecken und beobachteter Winkel wegen Refraktion*, Zeitschrift für Vermessungswesen, 86(7): 246–252.
- MORITZ, H. (1962): *Zur Geometrie der Refraktion*, Österreichische Zeitschrift für Vermessungswesen, 50(1): 3–13.
- MÖSER, M. (2000): *Grundlagen*, 3. Aufl., Handbuch der Ingenieurgeodäsie, Bd. 1, Herbert Wichmann Verlag, Heidelberg.
- NEUMAIER, A. (1990): *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge New York Port Chester.
- NEUMAIER, A. (2002): *Taylor forms, use and limits*, Reliable Computing, 9: 43–79.
- NIEMEIER, W. (2002): *Ausgleichsrechnung*, Walter de Gruyter, Berlin New York.
- NOBLE, B. und J. DANIEL (1988): *Applied Linear Algebra*, 3. Aufl., Prentice-Hall, London.
- NUDING, E. R. (1975): *Intervallrechnung und Wirklichkeit*, In: NICKEL, K. (Hrsg.), *Interval Mathematics*, 263–267, Lecture Notes in Computer Sciences, Bd. 29, Springer, Berlin Heidelberg New York.
- PAPOULIS, A. (1991): *Probability, Random Variables, and Stochastic Processes*, 3. Aufl., McGraw-Hill Book Company, New York San Francisco St. Louis.
- PARKINSON, B. W. (1996): *GPS-Error analysis*, In: PARKINSON, B. W. und J. J. SPILKER (Hrsg.), *Global Positioning System: Theory and Applications*, 469–483, Progress in Astronautics and Aeronautics, Bd.163, American Society of Aeronautics and Astronautics.
- PELZER, H. (1980): *Fehlerlehre und Statistik*, In: PELZER, H. (Hrsg.), *Geodätische Netze in Landes- und Ingenieurvermessung*, 25–78, Vermessungswesen bei Konrad Wittwer, Bd. 5, Verlag Konrad Wittwer, Stuttgart.
- REINHART, E. (1975): *Exakte Abschätzung von Maximalfehlern aus vorgegebenen Toleranzen der Beobachtungsgrößen*, DGK Reihe C 211, Deutsche Geodätische Kommission, München.
- REMBE, C., E. P. HOFER und B. TIBKEN (1999): *Model based Identification as a new tool to extract physical parameters of microactuators from measurements with error bounds*, In: *Technical Proceedings of the 1999 International Conference on Modeling and Simulation of Microsystems*, 276–279.
- RIESMEIER, K. (1983): *Test von Ungleichungshypothesen in linearen Modellen mit Bayes-Verfahren*, DGK Reihe C 292, Deutsche Geodätische Kommission, München.
- ROTHACHER, M., W. GURTNER, S. SCHAER, R. WEBER und H. HASE (1996): *Azimuth- and Elevation-Dependent Phase Center Corrections for Geodetic GPS Antennas Estimated from GPS Calibration Campaigns*, In: BEUTLER, G. (Hrsg.), *GPS trends in precise terrestrial, airborne and spaceborne applications*, 335–339, International Association of Geodesy Symposia, Bd. 115, Springer, Berlin Heidelberg New York.
- ROTHACHER, M. und S. SCHÄR (1995): *GPS-Auswertetechnik*, In: HECK, B. und M. ILLNER (Hrsg.), *GPS-Leistungsbilanz 94*, 107–121, Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 18, Verlag Konrad Wittwer, Stuttgart.
- RÜEGER, J.-M. (1996): *Electronic Distance Measurement - An Introduction*, 4. Aufl., Springer, New York Berlin Heidelberg.
- RÜEGER, J.-M. (1999): *Report of the Ad-Hoc Working Party on Refractive Indices of Light, Infrared and Radio Waves in the Atmosphere*, International Association of Geodesy, IAG Special Commission SC3 - Fundamental Constants, presented at the 22nd General Assembly of IUGG, 18-30 July 1999, Birmingham.
- RÜEGER, J.-M. und F. K. BRUNNER (2000): *On System Calibration and Type Testing of Digital Level*, Zeitschrift für Vermessungswesen, 125(4): 120–130.
- RUMP, S. (1999): *INTLAB- INTerval LABoratory*, In: CSENDES, T. (Hrsg.), *Developments in Reliable Computing*, 77–104, Kluwer Academic Publishers, Dordrecht.
- RUMP, S. (2002): *INTLAB- INTerval LABoratory, a Matlab toolbox for verified computations Version 3.1*, <http://www.ti3.tu-harburg.de/~rump/intlab/index.html>.
- RUSPINI, E. H., P. P. BONISSONE und W. REDRYCZ (Hrsg.) (1998): *Handbook of Fuzzy Computation*, Institute of Physics Publishing, Bristol.
- SAASTAMOINEN, J. (1973): *Contribution to the Theory of Atmospheric Refraction - Part II Refraction Corrections in Satellite Geodesy*, Bulletin Géodésique, 107: 13–34.
- SAWARAGI, Y., H. NAKAYAMA und T. TANINO (1985): *Theory of Multiobjective Optimization*, Mathematics in Sciences and Engineering, Bd. 176, Academic Press, Boston San Diego New York.
- SCHAFFRIN, B. (1985): *Aspects of Network Design*, In: GRAFAREND, E. W. und F. SANSÒ (Hrsg.), *Optimization and Design of Geodetic Networks*, 548–597, Springer, Berlin Heidelberg New York.
- SCHLEMMER, H. (1996): *Grundlagen der Sensorik - eine Instrumentenkunde für Vermessungsingenieure*, Herbert Wichmann Verlag, Heidelberg.
- SCHMIDBAUER, H.-O. (1989): *Berechnung minimaler Wertebereiche von Punktlagen in räumlichen Netzen aus vorgegebenen Toleranzen der Beobachtungsgrößen*, DGK Reihe C 355, Deutsche Geodätische Kommission, München.
- SCHMITT, G. (1973): *Speichertechnische und numerische Probleme bei der Auflösung großer geodätischer Normalgleichungssysteme*, DGK Reihe C 195, Deutsche Geodätische Kommission, München.

- SCHMITT, G. (1977): *Some Considerations Using Interval Analysis in Adjustment Computations*, In: *Weitere Beiträge aus der BRD zur Vorlage bei der XVI Generalversammlung der IUGG, Grenoble 1975*, 87–97, DGK Reihe B 221, Deutsche Geodätische Kommission, München.
- SCHMITT, G. (1982): *Optimization of Geodetic Networks*, *Reviews of Geophysics and Space Physics*, 30(4): 877–884.
- SCHMITT, G. (1985a): *Review of Network Designs: Criteria, Risk Function, Design Order*, In: GRAFAREND, E. W. und F. SANSÒ (Hrsg.), *Optimization and Design of Geodetic Networks*, 6–10, Springer, Berlin Heidelberg New York.
- SCHMITT, G. (1985b): *Second Order Design*, In: GRAFAREND, E. W. und F. SANSÒ (Hrsg.), *Optimization and Design of Geodetic Networks*, 74–121, Springer, Berlin Heidelberg New York.
- SCHMITT, G. (1985c): *Third Order Design*, In: GRAFAREND, E. W. und F. SANSÒ (Hrsg.), *Optimization and Design of Geodetic Networks*, 122–131, Springer, Berlin Heidelberg New York.
- SCHMITT, G. (1997): *Spectral Analysis and Optimization of two-dimensional Networks*, *Geomatics Research Australasia*, 67: 47–64.
- SCHÖN, S. und H. KUTTERER (2001a): *Intervallmathematische Behandlung von Unsicherheiten in geodätischen Messungen*, *Vermessung, Photogrammetrie, Kulturtechnik*, 99(3): 132–135.
- SCHÖN, S. und H. KUTTERER (2001b): *Network Optimization with respect to Systematic Errors*, In: ADAM, J. und K.-P. SCHWARZ (Hrsg.), *Vistas for Geodesy in the New Millennium*, 329–334, International Association of Geodesy Symposia, Bd. 125, Springer, New York Berlin Heidelberg.
- SCHÖN, S. und H. KUTTERER (2001c): *Optimal Design of Geodetic Monitoring Networks by means of Interval Mathematics*, In: WHITACKER, C. (Hrsg.), *10<sup>th</sup> FIG International Symposium on Deformation Measurements*, Orange California, 41–46.
- SCHÖN, S., H. KUTTERER, M. MAYER und B. HECK (2001): *Study on the Transfer of the ITRF Datum to a GPS Network in Antarctica*, In: SIDERIS, M. G. (Hrsg.), *Gravity, Geoid and Geodynamics 2000*, 29–34, International Association of Geodesy Symposia, Bd. 123, Springer, New York Berlin Heidelberg.
- SCHWARZ, K.-P. (2000): *Presidential address, International Association of Geodesy XII IUGG General Assembly, Birmingham, UK, July 18-31, 1999*, *Journal of Geodesy*, 74(1): 41–68.
- SCHWARZ, W. (Hrsg.) (1993): *Vermessungsverfahren im Maschinen- und Anlagenbau*, Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 13, Verlag Konrad Wittwer, Stuttgart.
- SCHWIEGER, V. (1999): *Ein Elementarfehlermodell für GPS-Überwachungsmessungen, Konstruktion und Bedeutung interepochaler Korrelationen*, *Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover*, Nr. 231.
- SHEPHARD, G. C. (1974): *Combinatorial Properties of Associated Zonotopes*, *Canadian journal of mathematics*, 26(2): 302–321.
- SILLARD, P. und C. BOUCHER (2001): *A Review of Algebraic Constraints in Terrestrial Reference Frame Datum Definition*, *Journal of Geodesy*, 75: 63–73.
- STAIGER, R. (1998): *Zur Überprüfung moderner Vermessungsinstrumente*, *Allgemeine Vermessungsnachrichten*, 105(11-12): 365–372.
- STRAHLBERG, C. (1997): *Eine vektorielle Darstellung des Einflusses von Ziel- und Kippachsfehler auf die Winkelmessung*, *Zeitschrift für Vermessungswesen*, 122(5): 225–234.
- TIPLER, P. A. (1994): *Physik*, Spektrum Akademischer Verlag, Heidelberg Berlin Oxford.
- TORGE, W. (2001): *Geodesy*, 3. Aufl., Walter de Gruyter, Berlin New York.
- VIERTL, R. (1996): *Statistical Methods for Non-Precise Data*, CRC Press, Boca Raton New York London.
- WANG, J. und Y. CHEN (1999): *Outlier Detection and Reliability Measures for Singular Adjustment Models*, *Geomatics Research Australasia*, 71: 57–72.
- WELSCH, W., O. HEUNECKE und H. KUHLMANN (2000): *Auswertung geodätischer Überwachungsmessungen*, 3. Aufl., *Handbuch der Ingenieurgeodäsie*, Bd. 3, Herbert Wichmann Verlag, Heidelberg.
- WICKI, F. (1998): *Robuste Schätzverfahren für die Parameterschätzung in geodätischen Netzen*, *IGP-Mitteilungen*, Bd. 67, Eidgenössische Technische Hochschule, Zürich.
- WILHELM, W. (1994): *Die Seitenrefraktion - Ein unbeliebtes Thema? Oder ein Thema nur für Insider*, *Vermessung, Photogrammetrie, Kulturtechnik*, 92(2): 75–82.
- WITTE, B. (1990): *Möglichkeiten und Grenzen des trigonometrischen Nivellements*, *Vermessungswesen und Raumordnung*, 52: 86–96.
- XU, P. (1993): *Multi-objective second order optimization and criterion matrix design*, *Bollettino di Geodesia e Scienze Affini*, 52(4): 305–323.
- XU, P. und E. GRAFAREND (1995): *A multi-objective second-order optimal design for deforming networks*, *Geophysical Journal International*, 120: 577–589.
- YONG-QI, C. und J.-L. WANG (1996): *Reliability measures for correlated observations*, *Zeitschrift für Vermessungswesen*, 121(4): 211–218.
- ZEISKE, K. (1999): *TPS 1100 Professional Series - Eine neue Tachymetergeneration von Leica*, *Vermessungswesen und Raumordnung*, 61(2): 82–90.
- ZIEGLER, G. M. (1995): *Lectures on Polytopes*, *Graduate Texts in Mathematics*, Bd. 152, Springer, New York Berlin Heidelberg.

## Lebenslauf

6. Mai 1974	geboren in Karlsruhe
1980 - 1984	Grundschule Hochstetten
1984 - 1993	Gymnasium Neureut, Karlsruhe–Neureut
13. Mai 1993	Abitur
1993 - 1994	Zivildienst
1994 - 1999	Studium des Vermessungswesens an der Universität Karlsruhe (TH) und an der Ecole Nationale des Sciences Géographiques (ENSG/IGN) in Marne-la-Vallée (Frankreich)
22. Dezember 1999	Zeugnis über die Diplomhauptprüfung
Februar 2000 – Juni 2002	wissenschaftlicher Angestellter am Geodätischen Institut der Universität Karlsruhe (TH)
seit Juli 2002	wissenschaftlicher Angestellter am Deutschen Geodätischen Forschungsinstitut (DGFI), München

## Dank

Mein Dank gilt Herrn Prof. Dr.-Ing. Dr.-Ing. E.h. G. Schmitt für die Übernahme des Hauptreferats, die damit verbundene Unterstützung und die gewährten Freiheiten. Bei Herrn Prof. Dr. rer. nat. G. Alefeld bedanke ich mich für die Übernahme des Korreferats und hilfreiche Anmerkungen zur Thematik der Intervallmathematik. Mein herzlicher Dank gilt Herrn Priv.-Doz. Dr.-Ing. H. Kutterer für seine engagierte Betreuung und wichtigen Impulse bei der Bearbeitung einzelner Themenbereiche sowie für seine Bereitschaft, die Arbeit zu begutachten.

Herrn Prof. Dr.-Ing. B. Heck möchte ich für die wertvolle Unterstützung während meiner Karlsruher Zeit danken sowie Herrn Hon.-Prof. Dr.-Ing. H. Drewes für die Möglichkeit, dass ich in der zweiten Projekthälfte ans DGFI wechseln konnte. Weiterhin danke ich den Kollegen am Deutschen Geodätischen Forschungsinstitut (DGFI), den ehemaligen Kollegen am Geodätischen Institut der Universität Karlsruhe und alle Anderen, die mit vielen Diskussionen das Entstehen der Arbeit begleitet haben.

Mein ganz besonderer Dank gilt meiner Frau Gaël für ihre Ermunterung, ihr Verständnis und ihre Rücksichtnahme während der Fertigstellung der Arbeit.

Die vorliegende Arbeit präsentiert Ergebnisse des Forschungsvorhabens „Optimaler Entwurf geodätischer Überwachungsnetze unter Berücksichtigung strenger Toleranzen“, das im Rahmen des Bündelprojektes „Aufbau eines Informationssystems zur geodätischen Deformationsanalyse unter Einbeziehung heterogener Daten“ von der Deutschen Forschungsgemeinschaft (DFG) gefördert wurde.