

Neue Formalismen zum Schätzen von Messunsicherheiten – Ausgleich nach kleinsten Quadraten

Michael Grabe, Braunschweig

Manuskripteingang: 25. Juni 2001; zur Veröffentlichung angenommen: 19. April 2005

In zwei vorangehenden Artikeln wurde ein zum *ISO-Guide* alternatives Fehlermodell diskutiert. Der vorliegende Beitrag überträgt letzteres auf die Methode der kleinsten Quadrate. Dabei zeigt sich, dass die minimierte Summe der Residuenquadrate die ihr konventionell zugeschriebenen statistischen Eigenschaften verliert, das Gauß-Markoff'sche Theorem zusammenbricht und in Konsequenz dessen die Methode der kleinsten Quadrate nicht mehr in der Lage ist, so genannte „Optimal-Schätzer“ bereitzustellen. Die Arbeit stellt einen Formalismus vor, der die wahren Werte der Messgrößen zuverlässig lokalisiert. Die Kopplungen der Schätzer des Ausgleichs sind sowohl durch zufällige als auch durch systematische Fehler bedingt. Folglich tritt eine neue, durch systematische Fehler bedingte Klasse geometrischer Körper neben die begrifflich bekannten Konfidenz-Ellipsoide, die zufälligen Fehlern Ausdruck verleihen. Der Verfasser bezeichnet die neuen Körper als *Sekuritäts-Polytope*.

Schlagwörter: Revision der Fehlerfortpflanzung, Messunsicherheiten, Methode der kleinsten Quadrate, Gauß-Markoff'sches Theorem, *ISO-Guide*

A new Formalism for the Estimation of Measurement Uncertainties – Least Squares Adjustment

In two preceding papers, we have discussed an error model being alternative to that of the *ISO-Guide*. In the current paper, we transfer this model to the method of least squares. We show that the minimized sum of squared residuals loses its conventionally assumed statistical properties, the Gauß-Markoff theorem breaks down and, in consequence thereof, the method of least squares is no longer in a position to provide so-called “optimal estimators”.

The paper presents an formalism which reliably localizes the true values of the measurands. The couplings of the adjustment's estimators are due to random as well as to systematic errors. Consequently, a new class of geometric solids stemming from systematic errors has to be put alongside the conceptually known confidence ellipsoids expressing the influence of random errors. The author terms the new bodies *security polytopes*.

Keywords: Revision of error propagation, measurement uncertainties, method of least squares, Gauß-Markoff theorem, *ISO-Guide*

1 Einführung

In seinen *Abhandlungen zur Methode der kleinsten Quadrate* [1] unterschied C. F. Gauß zwischen unregelmäßigen oder zufälligen Messfehlern auf der einen Seite und konstanten oder regelmäßigen Messfehlern auf der anderen. Wie bekannt, schloß Gauß letztere mit der Begründung aus, entweder habe der Beobachter deren Ursachen abzustellen oder doch wenigstens jeden einzelnen Messwert von ihrem Einfluss zu befreien.

Konstante oder regelmäßige Messfehler bleiben zumindest während der Messzeit konstant. Sie sind nach Betrag und Vorzeichen unbekannt, jedoch mit Hilfe von Intervallen eingrenzbar. Heute werden sie als *unbekannte systematische Fehler* oder einfach nur als *systematische Fehler*¹ bezeichnet. Da sie sich weder durch apparative Maßnahmen noch auf anderen Wegen beseitigen oder umgehen lassen [2] und überdies in einer mit den zufälligen Fehlern vergleichbaren Größenordnung liegen, sind die Gauß'schen Formalismen der Fehlerrechnung notwendigerweise unvollständig. Der vorliegende Beitrag klassifiziert systematische Fehler als biaserzeugend [4–6] und analysiert ihre Auswirkungen auf die Methode der kleinsten Quadrate [7]. Wie in [7] dargestellt, legt die Existenz von Biases (oder Vorlasten) es nahe, zwischen Erwartungswerten und wahren Werten² zu unterscheiden.

Die Methode der kleinsten Quadrate „löst“ in sich widersprüchliche – eigentlich also unlösbare – lineare (oder linearisierte) Gleichungssysteme durch Approximation. Aus Sicht der Gauß'schen Fehlerrechnung haben die Varianzen der Komponenten des Lösungsvektors „optimale“ Eigenschaften, d. h. sie sind minimal. Dieser letztere Umstand ist Ausdruck des Gauß–Markoff'schen Theorems, das indessen voraussetzt, dass die Schätzer des Ausgleichs erwartungstreu, unbelastet oder unbiased sind. In dieser klassischen Interpretation generiert die Methode der kleinsten Quadrate *erwartungstreue, in den Eingangsdaten lineare Schätzer minimaler Varianz*.

Sind die Schätzer des Ausgleichs nicht erwartungstreu, so ändern sich zwangsläufig die statistischen Eigenschaften der minimierten Summe der Residuenquadrate. Gleichzeitig damit bricht das Gauß–Markoff'sche Theorem zusammen, sodass die Schätzer des Ausgleichs die ihnen gemeinhin unterstellten „optimalen Eigenschaften“ verlieren. Schließlich sind im Rahmen der Formalisierung der Kopplungen zwischen den Schätzern des Ausgleichs neben den durch zufällige Fehler bedingten *Konfidenz-Ellipsoiden* andere, seitens der systematischen Fehler definierte Körper zu erwarten.

¹ Aus messtechnischer Sicht sind Fehler immer unbekannt. Unter diesem Aspekt ist der Zusatz „unbekannt“ entbehrlich.

² Wahre Werte sind diejenigen Werte, die die Formeln der Natur definieren. Wahre Werte sind in aller Regel unbekannt.

2 Formalismus des Ausgleichs

Ausgleichen nach kleinsten Quadraten heißt senkrecht projizieren. Eine Analogie möge uns helfen: Wir beleuchten einen Projektionsschirm in Richtung seiner Normalen mit parallelem Licht. Dann ist der Schattenwurf eines schräg gegen die Normale gehaltenen Stabes ein Äquivalent zur senkrechten Projektion des Ausgleichs nach kleinsten Quadraten. Nehmen wir an, das zu untersuchende physikalische Problem lasse sich mittels eines linearen Systems der Form

$$\begin{aligned} a_{11}\beta_1 + a_{12}\beta_2 + \dots + a_{1r}\beta_r &= x_1 \\ a_{21}\beta_1 + a_{22}\beta_2 + \dots + a_{2r}\beta_r &= x_2 \\ \dots & \\ a_{m1}\beta_1 + a_{m2}\beta_2 + \dots + a_{mr}\beta_r &= x_m \end{aligned}$$

beschreiben. Zweckmäßigerweise fassen wir die r Unbekannten β_k ; ($k = 1, \dots, r$) zu einem $r \times 1$ -Spaltenvektor β zusammen und die m rechten Seiten, die Messgrößen x_i ; ($i = 1, \dots, m$), zu einem $m \times 1$ -Spaltenvektor x . Schließlich sollen die Koeffizienten a_{ik} ; ($i = 1, \dots, m$; $k = 1, \dots, r$) die $m \times r$ -Matrix A definieren. Mit diesen Festsetzungen geht das System über in

$$A\beta = x. \tag{1}$$

Wir setzen voraus, A habe den Rang r und es gelte $m > r$.

Nummehr lassen wir fehlerhafte rechte Seiten zu, d. h. wir verknüpfen das physikalische Modell (1) mit den in [5; 6] definierten Fehlergleichungen

$$x_i = x_{0,i} + \varepsilon_i + f_i, \quad i = 1, \dots, m. \tag{2}$$

Die $x_{0,i}$ bezeichnen die (unbekannten) wahren Werte der physikalischen Messgrößen, die ε_i zufällige und die f_i systematische Messfehler. Erstere seien voneinander unabhängig und normalverteilt, letztere zeitkonstant und durch Intervalle der Form

$$-f_{s,i} \leq f_i \leq f_{s,i}; \quad i = 1, \dots, m \tag{3}$$

eingegrenzt. Das Fehlermodell (2) schließt den wahren Vektor der Eingangsdaten

$$x_0 = (x_{0,1} \quad x_{0,2} \quad \dots \quad x_{0,m})^T \tag{4}$$

ein³. Faktisch heißt das, wir verlangen, das System (1) sei physikalisch sinnvoll, d. h. es existiere das mathematisch wohldefinierte, wahre System

$$A\beta_0 = x_0, \tag{5}$$

³ T kennzeichne die Transponierung von Vektoren und Matrizen.

in dem $\beta_0 = (\beta_{0,1} \ \beta_{0,2} \ \dots \ \beta_{0,r})^T$ den wahren Lösungsvektor bezeichne. Auflösen von (5) liefert

$$\beta_0 = (A^T A)^{-1} A^T x_0. \tag{6}$$

Unter dem Einfluss fehlerhafter Eingangsdaten wird das System (1) numerisch inkonsistent. Bezeichnet jetzt x den Spaltenvektor der fehlerhaften Messdaten (2),

$$x = (x_1 \ x_2 \ \dots \ x_m)^T, \tag{7}$$

so gehen (1) und (5) über in

$$A\beta \approx x. \tag{8}$$

Die Methode der kleinsten Quadrate „beseitigt“ die Inkonsistenzen dieses Systems, indem sie den Vektor x der fehlerhaften Messdaten mit Hilfe des Operators

$$P = A (A^T A)^{-1} A^T \tag{9}$$

senkrecht in den Spaltenraum der Matrix A projiziert. Sei Px die im Spaltenraum der Matrix A liegende Komponente von x . Dann ist das *per Postulat* definierte System

$$A\bar{\beta} = Px \tag{10}$$

widerspruchsfrei und somit lösbar. Die Komponente Px unterscheidet sich vom Vektor x um den Residuenvektor $\bar{r} = x - Px$. Lösungsvektor des ausgeglichenen Systems ist

$$\bar{\beta} = (A^T A)^{-1} A^T x = B^T x, \quad B = A (A^T A)^{-1}. \tag{11}$$

Der Vektor $\bar{\beta}$ approximiert den wahren Lösungsvektor β_0 . In Konsequenz des senkrechten Projizierens nimmt die *Summe der Residuenquadrate* ihren kleinstmöglichen Wert

$$Q_{\min} = (x - Px)^T (x - Px) \tag{12}$$

an.

Indessen, senkrecht Projizieren beseitigt Inkonsistenzen beliebiger Art. Wäre also das wahre System (5) nicht definiert oder nicht definierbar, lägen überdies problemhäre, prinzipielle Überlegungsfehler und des Weiteren nichterkannte systematische Fehler vor, so lieferte die Methode der kleinsten Quadrate dennoch eine – unter diesen Voraussetzungen allerdings unbrauchbare – „Lösung“. Dieser Umstand wäre jedoch nicht dem Ausgleichsformalismus anzulasten, sondern der Art seiner Handhabung. Die Interpretation des Ausgleichs, ohne Bezüge zu *wahren Werten* herzustellen, bleibt metrologisch gesehen eben fragwürdig. Welche Konsequenzen es haben kann, wenn AuswerteprozEDUREN sich nicht auf wahre Werte stützen, zeigen die Beispiele (i) und (ii) des Abschnitts 4. Nach Meinung des Verfassers liegt hier der neuralgische Punkt des *ISO-Guides*.

3 Konsequenzen systematischer Fehler

Fassen wir das Fehlermodell (2) vektoriell, in

$$x = x_0 + (x - \mu) + f \tag{13}$$

bezeichne μ den Spaltenvektor der Erwartungswerte⁴ $\mu_i = E\{X_i\}; i = 1, \dots, m$ der Zufallsvariablen⁵ X_i und f den Spaltenvektor der m systematischen Fehler $f_i; i = 1, \dots, m$,

$$\mu = (\mu_1 \ \mu_2 \ \dots \ \mu_m)^T, \quad f = (f_1 \ f_2 \ \dots \ f_m)^T.$$

Mit (13) geht (11) über in

$$\bar{\beta} = B^T [x_0 + (x - \mu) + f] = \beta_0 + B^T (x - \mu) + B^T f. \tag{14}$$

Der Erwartungswert von $\bar{\beta}$,

$$\mu_{\bar{\beta}} = E\{\bar{\beta}\} = \beta_0 + B^T f, \tag{15}$$

unterscheidet sich um die Vorlast oder den Bias $B^T f$ vom wahren Lösungsvektor β_0 . Im Sinne der Aussage (14) sollte die Unsicherheit des Schätzers $\bar{\beta}$ aus der *additiven* Verknüpfung der Abschätzungen der Terme $B^T(x - \mu_x)$ und $B^T f$ hervorgehen. Wir werden hierauf in Abschnitt 4 zurückkommen.

Sind die Eingangsdaten des linearen Systems, so wie in (7) notiert, einzelne Messwerte (liegen also für jeweils festes i keine Wiederholungsmessungen vor), so sehen wir uns – im Gegensatz zum Gauß'schen Formalismus – nicht in der Lage, die statistischen Fluktuationen der Messdaten $x_i; i = 1, \dots, m$ mittels der minimierten Summe der Residuenquadrate zu beschreiben. Im einfachsten Fall sind die Zufallsvariablen $X_i; i = 1, \dots, m$ ein und derselben Grundgesamtheit zugeordnet. Ist das der Fall, so haben die Eingangsdaten einheitlich die theoretische Varianz

$$\sigma^2 = E\{(X_i - \mu_i)^2\}; \quad i = 1, \dots, m. \tag{16}$$

Wie bekannt entwickelt der klassische Gauß'sche Ausgleichsformalismus aus der minimierten Summe der Residuenquadrate Q_{\min} den Erwartungswert

$$E\{Q_{\min}\} = (m - r) \sigma^2,$$

sodass

$$s^2 = Q_{\min} / (m - r)$$

⁴ $E\{\}$ kennzeichne das Bilden von Erwartungswerten.

⁵ Jedem Messwert x entspreche eine Zufallsvariable X ; wir sagen „ x sei eine Realisierung von X .“

als empirische Varianz der Eingangsdaten auffassbar ist, d. h. s^2 gilt als Schätzer der theoretischen Varianz (16). Naturgemäß liefert das nichtgaußsche Fehlermodell (13) etwas gänzlich anderes, nämlich

$$E\{Q_{\min}\} = \left\{ (x - A\bar{\beta})^T (x - A\bar{\beta}) \right\} = (m - r) \sigma^2 + f^T f - f^T P f. \tag{17}$$

Das aber heißt, dass sich dem ausgeglichenen System nur dann ein Schätzer s^2 der unbekanntem theoretische Varianz σ^2 entnehmen lässt, wenn der auf systematische Messfehler zurückgehende Term $f^T f - f^T P f$ verschwindet. Im Allgemeinen ist das nicht der Fall, jedoch gibt es Ausnahmen [18].

Im Allgemeinen müssen wir also den Eingangsdaten die Struktur arithmetischer Mittel zuweisen. Im Sinne der in [5; 6] entwickelten Überlegungen sollten wir dabei mit wohldefinierten Messbedingungen arbeiten, d. h. für jedes der m Mittel jeweils genau n Wiederholungsmessungen einfließen lassen,

$$\begin{cases} x_{il} = x_{0,i} + \varepsilon_{il} + f_i \\ \bar{x}_i = x_{0,i} + (\bar{x}_i - \mu_i) + f_i \end{cases}, \quad i = 1, \dots, m; \quad l = 1, \dots, n. \tag{18}$$

Das Arbeiten mit ungleich vielen Wiederholungsmessungen wäre im Prinzip natürlich auch zulässig, jedoch verstieße es gegen die Vorstellung des wohldefinierten statistischen Modells normalverteilter Zufallsgrößen und brächte so formale Nachteile mit sich [5; 6]. Denn nur wenn gleich viele Wiederholungsmessungen vorliegen, ist die empirische Varianz-Kovarianz-Matrix der Eingangsdaten

$$\begin{cases} s = (s_{ij}) \\ s_{ij} = \frac{1}{n-1} \sum_{l=1}^n (x_{il} - \bar{x}_i) (x_{jl} - \bar{x}_j); \quad i, j = 1, \dots, m \end{cases} \tag{19}$$

definiert. Wir werden zeigen, dass sich die auf zufällige Fehler zurückgehenden Anteile der Gesamtunsicherheiten der Schätzer $\bar{\beta}_k, k = 1, \dots, r$ jetzt mit Hilfe klassischer Student'scher Vertrauensbereiche zum Tragen bringen lassen. Aus Gründen der Zweckmäßigkeit führen wir für die Diagonalelemente der Varianz-Kovarianz-Matrix s noch die Notation $s_{ii} \equiv s_i^2$ ein.

Das Gauß-Markoff'sche Theorem, siehe z. B. [8–10], besagt: Unter allen in den Eingangsdaten linearen, erwartungstreuen Schätzern minimiert der nach kleinsten Quadraten konstruierte Schätzer die Diagonalelemente der theoretischen Varianz-Kovarianz-Matrix des Lösungsvektors $\bar{\beta}$.

Der Gauß'sche Ausgleich betrachtet die Wurzeln jener Diagonalelemente als „Messunsicherheiten“. Letztere gelten als „optimal“ (d. h. minimal), wenn die Gewichtsmatrix im Sinne des Gauß-Markoff'schen Theorems festgelegt wird. Aber erinnern wir uns: Wichten verschiebt die numerischen Werte der Schätzer, im Zuge damit verringern sich deren Messunsicherheiten. Naturgemäß ist nicht entscheidbar, wie sich die Schätzer bewegen – von den unbekanntem wahren Werten weg oder zu ihnen hin.⁶ Und sogar von der Bewegungsrichtung unabhängig kann es geschehen, dass die infolge des Wichtens verringerten Messunsicherheiten, anstatt die wahren Werte zu lokalisieren, „ins Leere greifen“ – es sei denn, der Ausgleichsformalismus beuge einer derartigen, metrologisch nicht akzeptablen Situation vor. Indessen sind weder der Gauß'sche Formalismus noch der ISO-Guide hierfür gerüstet.

Wenn $E\{s\}$ die theoretische Varianz-Kovarianz-Matrix der Eingangsdaten bezeichnet, wäre im Sinne des Theorems

$$G = [E\{s\}]^{-1/2}$$

die optimale Gewichtsmatrix. Jedoch wurde bereits darauf hingewiesen, dass das Gauß-Markoff'sche Theorem für nichterwartungstreue Schätzer ungültig ist. Also kann die so definierte Gewichtsmatrix nicht optimal sein. Die Antwort auf die Frage, wie denn zu wichten sei, liegt im Fehlermodell selbst. Letzteres nämlich fordert

$$\bar{\beta}_k - u_{\bar{\beta}_k} \leq \beta_{0,k} \leq \bar{\beta}_k + u_{\bar{\beta}_k}; \quad k = 1, \dots, r. \tag{20}$$

In der Tat, wie der folgende Abschnitt zeigt, genügen die Messunsicherheiten des alternativen Fehlermodells eben dieser Forderung. Dabei dürfen wir die Gewichtsmatrix sogar ad hoc festlegen, zunächst beispielsweise in der Form

$$G = \text{diag}\{g_1, g_2, \dots, g_m\}; \quad g_i = \frac{1}{u_{\bar{x}_i}}, \tag{21}$$

worin die $u_{\bar{x}_i}$ die Unsicherheiten der Eingangsdaten \bar{x}_i sind. Um zu wichten, ist das auszugleichende System von links mit G zu multiplizieren.

Führen wir uns noch vor Augen, dass das wahre System (5) unempfindlich gegenüber jeder Art von Wichtung ist, dass es also stets den wahren Lösungsvektor β_0 reproduziert,

$$GA\beta_0 = Gx_0; \quad \beta_0 = [(GA)^T(GA)]^{-1} (GA)^T Gx_0.$$

Das heißt dann, dass wir die aktuellen Werte der Unsicherheiten $u_{\bar{\beta}_k}$, in gewissen Grenzen jedenfalls, durch

⁶ Bewegungen dieser Art lassen sich lediglich im Rahmen von Simulationsrechnungen verfolgen.

Wahl der Gewichte g_i „steuern“ können. D. h., wir dürfen den Ausgleich zyklisch wiederholen und die Unsicherheiten $u_{\bar{\beta}_k}$ durch Variation der g_i nach *trial and error* minimieren. Diese die Messunsicherheiten minimierende Prozedur ist so etwas wie ein Ersatz des Gauß–Markoff’schen Theorems.

4 Unsicherheiten der Komponenten des Lösungsvektors

Multiplizieren des auszugleichenden Systems

$$A\beta \approx \bar{x} \tag{22}$$

von links mit der Gewichtsmatrix (21) liefert

$$GA\beta \approx G\bar{x}. \tag{23}$$

Um die Schreibweise zu vereinfachen, ersetzen wir die Produkte GA und $G\bar{x}$ wieder durch A und \bar{x} , d. h. wir unterstellen gedanklich, das System (22) sei bereits gewichtet worden.

Das Fehlermodell (13) zerlegt den Lösungsvektor $\bar{\beta}$ in drei Komponenten,

$$\begin{cases} \bar{\beta} = B^T [x_0 + (\bar{x} - \mu) + f] \\ \bar{\beta} = \beta_0 + B^T (\bar{x} - \mu) + B^T f \end{cases} \tag{24}$$

Die Unsicherheiten der Komponenten $\bar{\beta}_k, k = 1, \dots, r$ des Lösungsvektors $\bar{\beta}$ sind anhand der Terme $B^T(\bar{x} - \mu)$ und $f_{\bar{\beta}} = B^T f$ zu gewinnen. Um den auf zufällige Fehler zurückgehenden Anteil der Gesamtunsicherheit zu schätzen, entwickeln wir aus den Messdaten (18) die empirische Varianz-Kovarianz-Matrix des Lösungsvektors $\bar{\beta}$, was wir können, da voraussetzungsgemäß wohldefinierte Messbedingungen zugrunde liegen. Wir bezeichnen die Spaltenvektoren der Matrix B mit $b_k; k = 1, \dots, r$ und ihre Elemente mit b_{ik} , d. h.

$$B = (b_{ik}) = (b_1 \quad b_2 \quad \dots \quad b_r).$$

Ändern der Summationsreihenfolge führt auf

$$\begin{aligned} \bar{\beta}_k &= \sum_{i=1}^m b_{ik}\bar{x}_i = \sum_{i=1}^m b_{ik} \left[\frac{1}{n} \sum_{l=1}^n x_{il} \right] \\ &= \frac{1}{n} \sum_{l=1}^n \left[\sum_{i=1}^m b_{ik}x_{il} \right] = \frac{1}{n} \sum_{l=1}^n \bar{\beta}_{kl} \end{aligned} \tag{25}$$

mit

$$\bar{\beta}_{kl} = \sum_{i=1}^m b_{ik}x_{il}. \tag{26}$$

Mit (25) und (26) lassen sich die Elemente

$$s_{\bar{\beta}_k\bar{\beta}_{k'}} = \frac{1}{n-1} \sum_{l=1}^n (\bar{\beta}_{kl} - \bar{\beta}_k) (\bar{\beta}_{k'l} - \bar{\beta}_{k'}) = b_k^T s b_{k'}$$

der empirischen Varianz-Kovarianz-Matrix des Lösungsvektors darstellen,

$$s_{\bar{\beta}} = B^T s B = (s_{\bar{\beta}_k\bar{\beta}_{k'}}); \quad k, k' = 1, \dots, r. \tag{27}$$

Zweckmäßigerweise notieren wir die Diagonalelemente auch hier wieder in der Form $s_{\bar{\beta}_k\bar{\beta}_k} \equiv s_{\bar{\beta}_k}^2$. Die Matrix s ist die in (19) definierte empirische Varianz-Kovarianz-Matrix der Eingangsdaten. Wie (25) zeigt, sind die $\bar{\beta}_k$ Summen normalverteilter und voneinander unabhängiger Schätzer β_{kl} . Damit aber sind wir in der Lage, die Erwartungswerte $\mu_{\bar{\beta}_k} = E\{\bar{\beta}_k\}$ der Schätzer β_k mittels Student’scher Vertrauensintervalle einzugrenzen,

$$\bar{\beta}_k - \frac{t_P (n-1)}{\sqrt{n}} s_{\bar{\beta}_k} \leq \mu_{\bar{\beta}_k} \leq \bar{\beta}_k + \frac{t_P (n-1)}{\sqrt{n}} s_{\bar{\beta}_k}; \tag{28}$$

$(k = 1, \dots, r).$

Die t -Faktoren beziehen sich auf das Vertrauensniveau P .

Die worst-case Abschätzungen der Komponenten des Vektors der fortgepflanzten systematischen Fehler

$$f_{\bar{\beta}} = B^T f$$

liefern

$$f_{s,\bar{\beta}_k} = \sum_{i=1}^m |b_{ik}| f_{s,i}. \tag{29}$$

Im Hinblick auf die Zerlegung (24) sichert die additive Verknüpfung der Aussagen (28) und (29) die Forderung (20), [5; 6]. Nach allem sind die Resultate des Ausgleichs durch

$$\bar{\beta}_k \pm u_{\bar{\beta}_k}; \quad u_{\bar{\beta}_k} = \frac{t_P (n-1)}{\sqrt{n}} s_{\bar{\beta}_k} + \sum_{i=1}^m |b_{ik}| f_{s,i}; \tag{30}$$

$k = 1, \dots, r$

gegeben.

Zwei Beispiele mögen der Veranschaulichung des Formalismus dienen.

(i) *Gewichtetes Mittel von Mitteln.* Gesucht sei die Unsicherheit $u_{\bar{\beta}}$ des gewichteten Mittels

$$\bar{\beta} = \sum_{i=1}^m w_i \bar{x}_i$$

der m Mittel

$$\bar{x}_i = \frac{1}{n} \sum_{l=1}^n x_{il}; \quad u_{\bar{x}_i} = \frac{t_P (n-1)}{\sqrt{n}} s_{x_i} + f_{s,i}; \tag{31}$$

$(i = 1, \dots, m).$

Wir setzen voraus, jedem der Mittel \bar{x} liege ein und derselbe wahre Wert x_0 zugrunde.

Wäre das nicht der Fall, d. h. besäßen die Mittel unterschiedliche wahre Werte, so hätten wir ganz offensichtlich ein Interpretationsproblem. Denn, wie bekannt, ist arithmetisches Mitteln nichts anderes als der Spezialfall $r = 1$ des Ausgleichs nach kleinsten Quadraten. Im Falle unterschiedlicher wahrer Werte, wäre (5) nicht definiert und das wiederum hieße, man versuchte ein physikalisch nicht definiertes Problem auszugleichen. Davon unabhängig sollen Messunsicherheiten Messfehlern Ausdruck verleihen, nicht aber Diskrepanzen zwischen wahren Werten.

Wir fassen die \bar{x}_i zu einem Spaltenvektor $\bar{\mathbf{x}}$ zusammen. Aus formalen Gründen nehmen wir einen $m \times 1$ -Spaltenvektor $\mathbf{a} = (1 \ 1 \ \dots \ 1)^T$ hinzu, sodass

$$\mathbf{a}\beta \approx \bar{\mathbf{x}}. \tag{32}$$

Multiplizieren von links mit \mathbf{G} und Lösen des linearen Systems nach kleinsten Quadraten liefert

$$\bar{\beta} = \sum_{i=1}^m w_i \bar{x}_i; \quad w_i = g_i^2 / \sum_{i=1}^m g_i^2. \tag{33}$$

Zum Schätzen der Unsicherheiten greifen wir auf das Fehlermodell (18) zurück, sodass

$$\bar{\beta} = \beta_0 + \sum_{i=1}^m w_i (\bar{x}_i - \mu_i) + \sum_{i=1}^m w_i f_i. \tag{34}$$

Bezeichne $\mu_{\bar{\beta}}$ den Erwartungswert von $\bar{\beta}$,

$$\mu_{\bar{\beta}} = E \{ \bar{\beta} \} = \beta_0 + \sum_{i=1}^m w_i f_i. \tag{35}$$

Um den auf zufällige Fehler zurückgehenden Term abschätzen zu können, machen wir wieder von der Vollständigkeit des Satzes der Eingangsdaten Gebrauch. Durch Vertauschen der Summationsreihenfolge in

$$\begin{aligned} \bar{\beta} &= \sum_{i=1}^m w_i \left[\frac{1}{n} \sum_{l=1}^n x_{il} \right] = \frac{1}{n} \sum_{l=1}^n \left[\sum_{i=1}^m w_i x_{il} \right] = \frac{1}{n} \sum_{l=1}^n \bar{\beta}_l; \\ \bar{\beta}_l &= \sum_{i=1}^m w_i x_{il} \end{aligned} \tag{36}$$

finden wir, wenn

$$\mathbf{w} = (w_1 \ w_2 \ \dots \ w_m)^T \tag{37}$$

den Spaltenvektor der Gewichte w_i symbolisiert,

$$s_{\bar{\beta}}^2 = \frac{1}{n-1} \sum_{l=1}^n (\bar{\beta}_l - \bar{\beta})^2 = \sum_{i,j} w_i w_j s_{ij} = \mathbf{w}^T \mathbf{s} \mathbf{w}, \tag{38}$$

mit s nach (19). Die empirische Varianz $s_{\bar{\beta}}$ liefert ein Vertrauensintervall für den Erwartungswert $\mu_{\bar{\beta}}$ des Schätzers $\bar{\beta}$. Ist dann

$$f_{s,\bar{\beta}} = \sum_{i=1}^m w_i f_{s,i}$$

der Worst-case-Schätzer des auf systematische Fehler zurückgehenden Unsicherheitsanteils, so ist

$$\bar{\beta} \pm u_{\bar{\beta}}; \quad u_{\bar{\beta}} = \frac{t_P (n-1)}{\sqrt{n}} s_{\bar{\beta}} + \sum_{i=1}^m w_i f_{s,i} \tag{39}$$

das Resultat des Mitteln von Mitteln. Kritisiert werden könnte ein zu großer systematischer Unsicherheitsanteil. Insbesondere könnte es so aussehen, als nähme $f_{s,\bar{\beta}}$ mit wachsendem m unangemessen zu. Dass dem nicht so ist, verfolgen wir anhand zweier Fälle:

Sei $m = 3$ und unterstellen wir $u_{\bar{x}_1} = u_{\bar{x}_2} = u_{\bar{x}_3} = a$. Dann gilt

$$f_{s,\bar{\beta}} = \frac{1}{3} (f_{s,1} + f_{s,2} + f_{s,3}). \tag{40}$$

Für $u_{\bar{x}_1} = a, u_{\bar{x}_2} = 2a, u_{\bar{x}_3} = 3a$ folgt

$$f_{s,\bar{\beta}} = \frac{1}{49} (36f_{s,1} + 9f_{s,2} + 4f_{s,3}). \tag{41}$$

Demnach bewegen sich die systematischen Fehleranteile durchaus in akzeptabler Größenordnung.

(ii) *Gewichteter Ausgleich*

Solange dem Experimentator die wahren Werte der Messgrößen unbekannt sind, wie in aller Regel der Fall, bleibt offen, ob seine Messunsicherheiten jene wahren Werte denn auch tatsächlich lokalisieren – einfach deshalb, weil sie unbekannt sind. Leider ist diese metrologisch alles entscheidende Frage prinzipiell nicht beantwortbar. Dessen ungeachtet sind wir sehr wohl in der Lage, unsere Schätzverfahren mit Hilfe simulierter Daten zu testen.

Simulationsrechnungen gehen von a priori bekannten wahren Werten aus, überlagern ihnen nach Belieben zufällige und systematische Fehler und testen, ob die aus den gegebenen Eigenschaften der simulierten Daten folgenden Messunsicherheiten die a priori bekannten wahren Werte lokalisieren.

Die folgenden Bilder zeigen qualitativ die Resultate eines Ausgleichs für drei Schätzer $\bar{\beta}_1, \bar{\beta}_2, \bar{\beta}_3$. Den beiden oberen Bildern liegt der hier diskutierte Formalismus zugrunde, im unteren Bild sind die systematischen Fehler im Sinne des *ISO-Guides* „randomisiert“, d. h. den zufälligen Fehlern formal gleichgestellt worden. Dabei wurde absichtlich zu einem Datensatz gegriffen, anhand dessen der *ISO-Guide* die wahren Werte nicht zu lokalisieren vermag. Auf die Ei-

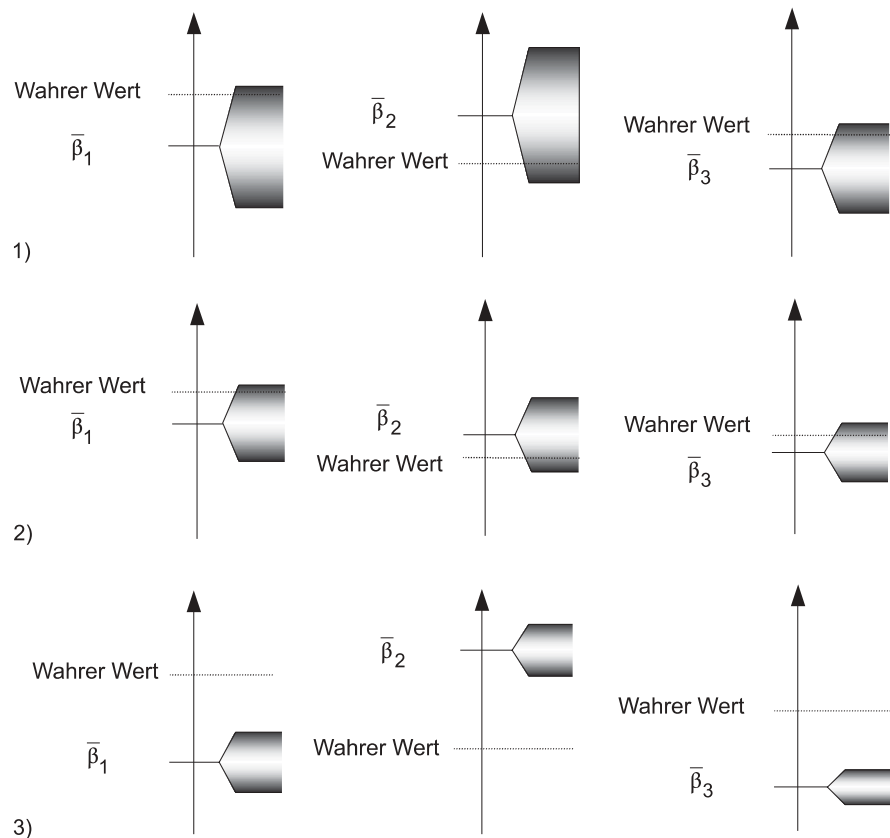


Bild 1: 1) Ungewichteter Ausgleich eines in sich widersprüchlichen linearen Systems, dem ein in wahren Werten definiertes, physikalisch sinnvolles System zugrunde liegt. Die Unsicherheiten des alternativen Fehlermodells lokalisieren die wahren Werte der Schätzer.

2) Gewichteter Ausgleich desselben System. Das Wichten bewirkt zweierlei: Es verschiebt die Schätzer des Ausgleichs und reduziert deren Unsicherheiten. Dennoch lokalisieren die Unsicherheiten die wahren Werte der Schätzer.

3) Gewichteter Ausgleich desselben Systems anhand des im ISO-Guide empfohlenen Fehlermodells. Situationsbedingt lokalisieren die Unsicherheiten entweder die wahren Werte oder greifen, wie hier dargestellt, „ins Leere“.

Figure 1: 1) Unweighted least squares adjustment of an inconsistent linear system which is readable in terms of true values and so identifies a physically meaningful system. The uncertainties of the alternative error model localize the true values of the estimators.

2) Weighted least squares adjustment of the same system. Weights have two types of consequences: they shift the estimators of the adjustment and they reduce the associated uncertainties. Still the uncertainties localize the true values of the estimators.

3) Weighted least squares adjustment of the same system based on the ISO-Guide's error model. The uncertainties may localize the true values or, as shown here, "grasp at nothing".

genschaft des *ISO-Guides*, knappe Messunsicherheiten zu liefern, wurde bereits in [5; 6] hingewiesen. Problematisch dabei ist, dass der Experimentator weder weiß noch jemals wissen kann, wann seine Unsicherheiten zu knapp ausfallen, d. h. ob sie wahre Werte lokalisieren. Demgegenüber stellt das alternative Fehlermodell die knappsten möglichen Messunsicherheiten bereit, die die wahren Werte „quasi sicher“ lokalisieren [5; 6].

Im Sinne des Gauß'schen Ausgleichsformalismus sollte die minimierte Summe der Residuenquadrate \bar{Q}_{\min} des gewichteten Ausgleichs den Freiheitsgrad $m - r$ des linearen Systems approximieren,

$$\bar{Q}_{\min} \approx m - r. \quad (42)$$

Demgegenüber hat jahrelange Erfahrung immer wieder $\bar{Q} \gg m - r$ zu Tage gefördert [15], was nahelegt, die Randomisierung systematischer Fehler zu überdenken. Das alternative Fehlermodell interpretiert systematische Fehler stattdessen als Vorlasten oder Biases, die Messapparaturen gewissermaßen „präkonditionieren“.

5 Unsicherheitsräume

Der Messprozess kann verlangen, mittels einiger oder aller Unsicherheiten $u_{\bar{\beta}_k}$ der Schätzer $\bar{\beta}_k$; $k = 1, \dots, r$ so genannte *Unsicherheitsräume* aufzuspannen. Naturgemäß ist die Geometrie solcher Räume durch die Art der Kopplungen zwischen den Schätzern, d. h. durch das Fehlermodell selbst bestimmt. Wie

$$\bar{\beta} = \mathbf{B}^T \bar{x} = \beta_0 + \mathbf{B}^T (\bar{x} - \mu) + \mathbf{B}^T f \tag{43}$$

zeigt, fließen die m Eingangsdaten \bar{x}_i ; ($i = 1, \dots, m$) gleichermaßen in jede der r Komponenten

$$\bar{\beta}_k = \beta_{0,k} + \sum_{i=1}^m b_{ik} (\bar{x}_i - \mu_i) + \sum_{i=1}^m b_{ik} f_i; \quad k = 1, \dots, r \tag{44}$$

des Lösungsvektors ein. Die Gauß'sche Fehlerrechnung kennt ausschließlich zufällige Messfehler. Unter diesem Aspekt drückt sie die Kopplungen zwischen den Schätzern des Ausgleichs durch Konfidenz-Ellipsoide aus. Letzteres, so der Brauch, werden dem χ^2 -verteilten Exponenten der mehrdimensionalen normalen Verteilungsdichte entnommen. Allerdings stützen sich die so definierten Konfidenz-Ellipsoide auf experimentell unzugängliche theoretische Varianz-Kovarianz-Matrizen, was naturgemäß erhebliche Interpretationsschwierigkeiten birgt. Das Arbeiten mit der wohldefinierte Messbedingungen verlangenden Hotelling'schen Dichte ist günstiger, da letztere sich auf experimentell zugängliche empirische Varianz-Kovarianz-Matrizen beschränkt. Mit

$$\mu_{\bar{\beta}} = \beta_0 + \mathbf{B}^T f$$

geht (43) über in

$$\bar{\beta} = \mu_{\bar{\beta}} + \mathbf{B}^T (\bar{x} - \mu). \tag{45}$$

Die Hotelling'sche Dichte, [14]

$$p_T(t; m, n) = \frac{2\Gamma(n/2)}{(n-1)^{m/2} \Gamma[(n-m)/2] \Gamma[m/2]} \frac{t^{m-1}}{\left[1 + \frac{t^2}{n-1}\right]^{n/2}}; \tag{46}$$

$t > 0$; $n > m$,

beschreibt die statistischen Eigenschaften des Differenzvektors $(\bar{x} - \mu)$ mittels des ausschließlich durch empirische Größen definierten Ellipsoids

$$t^2(m, n) = n (\bar{x} - \mu)^T s^{-1} (\bar{x} - \mu). \tag{47}$$

In (47) bezeichnet s die empirische Varianz-Kovarianz-Matrix der Eingangsdaten (19)⁷ und Γ die Gamma-Funktion. Für normalverteilte Messwerte sind die in (25) und (26) definierten $\bar{\beta}_{kl}$ und $\bar{\beta}_k$ normalverteilt. Da nach (27) die empirische Varianz-Kovarianz-Matrix

$$s_{\bar{\beta}} = \mathbf{B}^T s \mathbf{B} = (s_{\bar{\beta}_k \bar{\beta}_{k'}})$$

des Lösungsvektors bekannt ist, ist⁸

$$n (\bar{\beta} - \mu_{\bar{\beta}})^T (\mathbf{B}^T s \mathbf{B})^{-1} (\bar{\beta} - \mu_{\bar{\beta}}) = t^2(r, n) \tag{48}$$

das die statistischen Eigenschaften des Differenzvektors $\bar{\beta} - \mu_{\bar{\beta}}$ charakterisierende Hotellings'sche Ellipsoid. Der Vektor $\mu_{\bar{\beta}}$ ist unbekannt. Dementsprechend zentrieren wir das Ellipsoid im Punkte $\bar{\beta}$ und ersetzen $\mu_{\bar{\beta}}$ formal durch den Hilfsvektor β . Das so entstehende Konfidenz-Ellipsoid

$$(\beta - \bar{\beta})^T (\mathbf{B}^T s \mathbf{B})^{-1} (\beta - \bar{\beta}) = \frac{t_P^2(r, n)}{n} \tag{49}$$

überdeckt den durch den Vektor $\mu_{\bar{\beta}}$ definierten Raumpunkt mit der Wahrscheinlichkeit

$$P\{T \leq t_P(r, n)\} = \int_0^{t_P(r, n)} p_T(t; r, n) dt. \tag{50}$$

Schließlich ist zu klären, welche geometrischen Körper die durch systematische Fehler bedingten Kopplungen zwischen den Schätzern des Ausgleichs erzeugen [16]. Nach (43) haben wir von

$$f_{\bar{\beta}} = \mathbf{B}^T f$$

d. h.

$$f_{\bar{\beta}_k} = \sum_{i=1}^m b_{ik} f_i; \quad k = 1, \dots, r \tag{51}$$

auszugehen. Die singuläre Abbildung (51) transformiert den m -dimensionalen Hyperquader

$$-f_{s,i} \leq f_i \leq f_{s,i}; \quad i = 1, \dots, m \tag{52}$$

in ein r -dimensionales, konvexes und inversionssymmetrisches Polytop. Für $r = 2$ sind das Polygone, für $r = 3$ Polyeder und für $r > 3$ unanschauliche Polytope.

⁷ Aussage (47) setzt voraus, dass s nichtsingulär ist, [11].

⁸ Der Verfasser dankt Herrn Dr. W. Wöger, PTB Braunschweig, für seinen Hinweis auf die korrekte Interpretation der Aussage (48).

So definierte Polyeder bezeichnet der Verfasser als *Sekuritäts-Polyeder* – in dieser Nomenklatur treten sie als sprachliche Pendant der *Konfidenz-Ellipsoide* auf.

Für $r = 2$ sind *Konfidenz-Ellipsen* und *Sekuritäts-Polygone* zu resultierenden flächigen Unsicherheitsbereichen zusammzusetzen, für $r = 3$ *Konfidenz-Ellipsoide* und *Sekuritäts-Polyeder* zu dreidimensionalen Unsicherheitsräumen. Jedes dieser Unsicherheitsgebiete soll, das verlangen wir, den durch den wahren Lösungsvektor β_0 markierten Punkt lokalisieren [17; 18].

6 Schluss

Die diskutierten Formalismen stützen sich auf Annahmen, die zum gegenwärtigen Zeitpunkt weder in der Experimentalphysik noch in der Messtechnik üblich sind. Fassen wir zusammen:

- systematische Messfehler sollten den Status *zeitkonstanter* Vorlasten oder Biases haben,
- jedes der in die Auswerteprozeduren einfließenden arithmetischen Mittel sollte über *die gleiche Anzahl* von Wiederholungsmessungen verfügen, schließlich
- sollten sich Messunsicherheiten ausschließlich auf empirische, d. h. experimentell realisierbare Schätzer stützen. Prinzipiell unzugängliche, theoretische Verteilungsdichten konstituierende Erwartungswerte sollten nicht in Messunsicherheiten einfließen.

Die Formalismen zum Schätzen von Messunsicherheiten sollten widerspruchsfrei, robust und zuverlässig sein. Bereits in [5; 6] ist gezeigt worden, dass der hier diskutierte Formalismus über eben diese Eigenschaften verfügt. Hieran ändert sich nichts im Rahmen des Ausgleichs nach kleinsten Quadraten. Letzterer überträgt ja lediglich die dort diskutierte Fehlerfortpflanzung auf in den Eingangsdaten lineare Schätzer. Allerdings bedürfen die statistischen Eigenschaften der minimierten Summe der Residuenquadrate und das Gauß–Markoff'sche Theorem der Neuinterpretation.

Im Rahmen des *ISO-Guides* bereitet die Implementierung der „Traceability“⁹ formale und experimentelle Probleme. Traceability heißt, dass die Messunsicherheit jedes Normals jeder Hierarchie-Stufe den jeweils zugehörigen wahren Werte lokalisiert. Nur unter dieser Voraussetzung nämlich kann es gelingen, physikalische Einheiten und Standards metrologisch zuverlässig weiterzugeben. Im Sinne des alternativen Fehlermodells ist Traceability allerdings ein im Prinzip entbehrlicher Begriff, stellt doch das Fehlermodell a priori die Lokalisierung wahrer Werte sicher.

Es ist wohlthuend, der Metrologie einen angemessenen, realitätsnahen Platz zuzuweisen. Zwischen Theorie

und Experiment stehend, ist ihre Aussageschärfe, physikalisch bedingt, nicht mit der der Mathematik vergleichbar. Dessen ungeachtet fällt ihr die Aufgabe zu, die wahren Werte physikalischer Größen mit Hilfe von Intervallen eingrenzen zu müssen. Denn nur so lassen sich Aussagen hoher Evidenz von solchen niedriger Evidenz trennen. Das Differenzieren von Aussagen aber, die mit der Natur verträglich sind, von solchen die es nicht sind, definiert „Physik an sich“.

Nicht immer werden die den Formalismen der Fehlerrechnung zugrunde liegenden formalen Voraussetzungen erfüllt sein, weder seitens des Experimentes noch seitens des Experimentators. Auch dürften die zufälligen Fehler der Messgrößen eher approximativ als tatsächlich normalverteilt sein. Das Abschätzen der Grenzen der systematischen Fehler schließlich verlangt vom Experimentator, die maximal mögliche Differenz zwischen Erwartungswert und wahren Wert der Messgröße zu ermitteln – indessen, jede der genannten Größen ist unbekannt. Diese Schwierigkeiten sollten dennoch kein Anlass sein, die Notwendigkeit und den Stellenwert einer selbstkonsistenten Fehlerrechnung in Frage zu stellen oder herunterzuspielen. Dies um so weniger, als das Bemühen des Experimentators, den Baukasten der Fehlerrechnung im Sinne eines „Denkens in geordneten Kategorien“ zu implementieren im Allgemeinen jedenfalls erhebliche apparative Verbesserungen mit sich bringt. Möge ihn dabei Louis de Broglie's feinsinniger Aphorismus

Wahrheit ist ein Abglanz von Schönheit

leiten.

Literatur

- [1] Gauß, C.F., Abhandlungen zur Methode der kleinsten Quadrate, Physica, Würzburg 1964.
- [2] Eisenhart, C., The Reliability of Measured Values – Part I, Fundamental Concepts, Photogrammetric Engineering **18** (1952) 543–561.
- [3] ISO-Guide, Guide to the expression of uncertainty in measurement, 1993. 1, Rue de Varambè, Boîte Postale 56, CH-1211 Geneva 20, Switzerland.
- [4] Grabe, M., Principles of „Metrological Statistics“, Metrologia, **23** (1986/87) 213–219.
- [5] Grabe, M., Gedanken zur Revision der Gauß'schen Fehlerrechnung, tm Technisches Messen, **6** (2000) 283–288.
- [6] Grabe, M., Neue Formalismen zum Schätzen von Messunsicherheiten – Ein Beitrag zum Verknüpfen und Fortpflanzen von Messfehlern, tm Technisches Messen, **3** (2002) 142–150.

⁹ Rückführbarkeit oder Rückverfolgbarkeit.

- [7] Grabe, M., On the Assignment of Uncertainties within the Method of Least Squares, Second International Conference on Precision Measurement and Fundamental Constants, Poster Paper, Washington D.C. 8–12. June 1981.
- [8] Eadie, W.T. et al., Statistical Methods in Experimental Physics, North-Holland, Amsterdam 1971.
- [9] Rao, C.R., Linear Statistical Inference and its Applications, John Wiley & Sons, New York 1973.
- [10] Seber, G.A.F., Linear Regression, John Wiley & Sons, New York 1977.
- [11] Cramér, H., Mathematical Methods of Statistics, Princeton University Press, Princeton 1966.
- [12] Fisz, M., Probability theory and mathematical statistics, John Wiley & Sons, New York 1967.
- [13] Welch, B.L., The Generalization of Student's Problem when Several Different Population Variances are Involved, *Biometrika* **37** (1950) 149-157.
- [14] Hotelling, H., The Generalization of Student's Ratio, *Annals of Mathematical Statistics* **2** (1931) 360–378.
- [15] Cohen, E.R., Taylor, B.R., The 1986 Adjustment of the Fundamental Physical Constants, *Codata Bulletin*, 63 (1968).
- [16] Grabe, M., Anpassung eines Kreises nach kleinsten Quadraten, Jahresbericht der Physikalisch-Technischen Bundesanstalt 1989, Seiten 210–211.
- [17] Grabe, M., Schätzen von Messunsicherheiten in Wissenschaft und Technik, BOD GmbH, ISBN 3-8311-1309-2, November 2000, nicht mehr beziehbar.
- [18] Grabe, M., Measurement Uncertainties in Science and Technology, Springer-Verlag, März 2005.



Dr. Michael Grabe hat in Braunschweig und Stuttgart Physik studiert. Nach der Promotion und nachfolgender Tätigkeit als Hochschulassistent wurde er wissenschaftlicher Mitarbeiter einer in Berlin und Braunschweig ansässigen Großforschungseinrichtung. Seine Arbeitsgebiete umfassten Probleme

der Bauartprüfung und -zulassung von Waagen, der dynamischen Längenmesstechnik, der Auswertung von Messungen und der Numerischen Physik.

Adresse: Am Hasselteich 5, 38104 Braunschweig,
Tel.: 0531 371642, E-Mail: michael.grabe@gmx.net,
<http://www.uncertainty.de>