Hinweise zur Beurteilung von Messungen, Messergebnissen und Messunsicherheiten (ABW) („Fehlerrechnung“)

Scientific knowledge is a body of statements of varying degree of certainty — some most unsure, some nearly sure, *but none absolutely certain.*
R.P. Feynman

Literatur

- W. Walcher, *Praktikum der Physik*, B.G. Teubner Verlag
- H.J. Eichler, H.-D. Kronfeldt, J. Sahm, *Das Neue Physikalische Grundpraktikum*, Springer Verlag
- Deutsches Institut für Normung, DIN 1319
- J.R. Taylor, *Fehleranalyse*, Wiley-VCH
- M. Drosg, *Der Umgang mit Unsicherheiten*, Facultas Verlag

1 Vorbemerkungen

Dieses Skript soll die Grundlage für die Behandlung von Messunsicherheiten und die Bewertung von Messergebnissen im physikalischen Praktikum bilden. Es kann aber nur einen Einblick in diese sehr komplexe Thematik geben. Eine vollständige und ausführliche Behandlung dieses Themas würde den Rahmen dieses Skriptes und des Grundpraktikums sprengen. Die Zielsetzung im Praktikum ist es, dass Sie die Grundlage der Behandlung von Messunsicherheiten lernen und deren Anwendung üben.

1.1 Unterscheidung von *Unsicherheiten* und *Fehlern*

Es liegt in der Natur einer Messung, dass sie nicht beliebig genau sein kann. Es hat also nichts mit einem Fehlverhalten oder einer falschen Messung zu tun, dass ein Ergebnis vom „wahren Wert“ abweicht. Deshalb sollte man auch nicht von „Fehlern“ sprechen, sondern von „Unsicherheiten“oder „Abweichungen“.

Um „Fehler“ handelt es sich, wenn die Abweichungen auf dem Versagen des messenden Physikers oder seiner Apparatur zurückzuführen sind, etwa wenn man 1,50 V

statt 1,05 V abliest oder aufschreibt. Man kann diese Fehler weder vollständig ausschließen noch sinnvoll mathematisch bearbeiten. Daher ist der Begriff „Fehlerrechnung“ nicht sinnvoll.
Das Auftreten solcher Fehler und ihr Schaden wird aber begrenzt, wenn man sorgfältig arbeitet und wichtige Arbeitsschritte unabhängig wiederholt. Dieser Typ von Abweichungen sei im Vorfeld durch eine Plausibilitätsprüfung der Messwerte erledigt, und wird nicht weiter betrachtet.

Die meisten Unsicherheiten, mit denen wir es im Praktikum, oder mit denen Physiker in Forschung und Anwendung zu tun haben, beruhen auf Unvollkommenheiten unserer Messgeräte und unserem Umgang mit ihnen. Jedes Messgerät hat eine (bauartbedingte) Genauigkeit, die unsere Messung begrenzt. So lässt sich beispielsweise ein Multimeter mit vier Ziffern unmöglich mit fünf signifikanten Stellen ablesen.

Neben der Genauigkeit der verwendeten Messgeräte spielt aber auch deren Einfluss auf den gemessenen Vorgang eine Rolle. So kann das Messinstrument z.B. als zusätzlicher „Verbraucher“ elektrischer Energie auftreten, sein Einfluss muss dann rechnerisch eliminiert werden, da man sich ja eigentlich für den „ungestörten“ Zustand interessiert.

Schließlich gibt es noch *Schwankungen des Messwertes* selbst. Dies sind eigentlich keine Mess*abweichungen*, da ja der momentan vorliegende Messwert korrekt ermittelt wird. Meist ist man jedoch an der zu Grunde liegenden physikalischen Größe, oft dem Mittelwert, interessiert und nicht an der aktuellen Fluktuation.
Als wichtiges Beispiel ist hier die Zerfallskonstante des radioaktiven Zerfalls zu nennen. Die einzelnen gemessenen Zerfälle folgen perfekt statistischen Gesetzen und schwanken dementsprechend.

1.2 Schreibweise von Werten und Unsicherheiten

Die Unsicherheit einer Größe x wird durch ein vorangestelltes Δ gekennzeichnet, also hier Δx. Ein Messwert oder Ergebnis wird immer zusammen mit seiner Unsicherheit, und mit der dazugehörenden Einheit angegeben, also allgemein in der Form

$$(x \pm \Delta x)\ \text{Einheit} \tag{1}$$

Man kann die Unsicherheit als
absolute Unsicherheit angeben, d.h. die Unsicherheit wird einfach an den ermittelten Wert der Größe mit einem $\pm$-Zeichen in der gleichen Einheit angehängt, z.B. $(355,62 \pm 0,03)$ mm oder $(2,53 \pm 0,13)$ mV.

Alternativ kann man auch die
relative Unsicherheit als Anteil vom Messwert angeben. In diesem Fall schreibt man dann z.B. $355,62$ cm $\pm\, 0,08\,\%$ oder $2,53$ mV $\pm\, 5\,\%$. Bei der Fortpflanzung mehrerer Unsicherheiten bietet diese Angabe in einigen Fällen erhebliche Vorteile (s. 7.3.2).

Weiterhin ist es zweckmäßig, die Unsicherheit einer Messgröße auch **implizit** in Form von signifikanten Stellen anzugeben, z.B. 355,62 mm und nicht 355,6 mm oder 355,620 mm. Diese Form der Fehlerangabe muss immer durchgeführt werden, und mit der expliziten Angabe konsistent sein, insbesondere bei der Übernahme von Werten vom Taschenrechner. So ist z.B. die Angabe $(6543, 21 \pm 50)$ g für eine Masse falsch, richtig wäre hier (6540 ± 50) g, oder noch besser $(6, 54 \pm 0, 05)$ kg. .

Es ist auch nicht sinnvoll die Unsicherheiten selber zu genau anzugeben. Für das Praktikum soll gelten, dass Unsicherheiten auf *zwei* geltende Ziffern gerundet werden, falls die führende Ziffer eine „1“ oder eine „2“ ist. In allen anderen Fällen wird die Unsicherheit nur auf *eine* geltende Ziffer angegeben.

Bei der Angabe von Werten sollen sinnvolle ‚mathematische Vorsilben‘ verwendet werden. Die Angabe 0,35562 m ist gerade noch lesbar, besser wäre aber 35,562 cm (falls der Wert wirklich so genau ist). Die Angabe von $4, 66 \cdot 10^{-7}$ m für eine Lichtwellenlänge ist zwar formal korrekt, aber ohne Hirnakrobatik nicht mehr verarbeitbar. So wie man die Entfernung von München nach Berlin weder im Lichtjahren noch in mm angibt, sondern in km (auch weil dadurch das Problem mit den signifikanten Stellen verringert wird), sollte man Lichtwellenlängen in nm (oder μm bei Infrarotstrahlung) angeben, also 466 nm. Völlig indiskutabel sind Schreibweisen wie 4.66E-7 oder 466E-9. Für andere Zahlen und Größen gilt das natürlich analog.

2 Systematische und Statistische Unsicherheiten

Bei der Behandlung von Messabweichungen ist es zweckmäßig zwischen zwei Typen von Abweichungen[1] zu unterscheiden, die eingangs beschriebenen groben Fehler seien dabei bereits aussortiert :

Statistische (unkorrelierte) Unsicherheiten: Die Auswirkung dieser Art von Abweichungen kann vermindert werden, wenn man unter ‚gleichen Bedingungen‘ die Messung mehrfach wiederholt und den Mittelwert all dieser Messungen heranzieht. Die Abweichung ist nicht durch einen dauerhaften falschen Zustand des Messgerätes, z.B. eine falsche Eichung, hervorgerufen, sondern durch eine sich zeitliche Änderung im Messvorgang

Systematische (korrelierte) Unsicherheiten: Diese Art von Abweichung verringert sich nicht, sooft man auch die Messung wiederholt. Eine zumindest über den Zeitraum der Wiederholungen gleichbleibende abweichende Anzeige des Messgerätes liegt vor.
Auch der Einfluss der Messgeräte selbst auf die Messung fällt unter diese Unsicherheiten.

Detaillierte Kenntnisse der Messbedingungen und des Messgerätes erlauben manchmal eine Korrektur der Abweichung.

[1] Auch wenn die hier gemachte Zuordnung statistische = unkorrelierte Unsicherheit, und systematische = korrelierte Unsicherheit in dieser Einfachheit sicher nicht ganz korrekt ist, ist sie für die Anwendung im Praktikum ausreichend. Vom mathematischen Gesichtspunkt ist die Korrelation zwischen den Größen und Unsicherheiten ausschlaggebend.

Hier ein Beispiel:
Mit einem Lineal als Messgerät soll die Länge eines Aluminiumquaders gemessen werden. Zu diesem Zweck legt man das Lineal mit dem Nullpunkt der Skala an das eine Ende des Quaders und liest am anderen Ende des Quaders auf der Skala den Messwert ab.

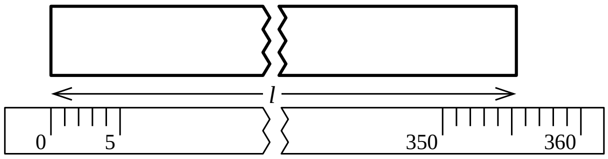

Abbildung 1: Längenmessung

Dabei gibt es folgende Unsicherheiten:

1. *Das Anlegen des Lineals an den Quader (Nullposition) erfolgt mit einer gewissen Genauigkeit die in das Ergebnis eingeht. Dieser Vorgang wird sicher zu einem anderen Ergebnis führen wenn man ihn wiederholt, ist also statistischer Natur.*
2. *Das Ablesen am anderen Ende des Lineals wird bei einer unabhängigen Wiederholung, z.B. durch eine andere Person, zu einem neuen Ergebnis führen, ist also ebenfalls statistischer Natur.*
3. *Weitere Abweichungen liegen in der Eichung des Messgerätes begründet: Wurde bei der Herstellung des Lineals die Skala in "falschem Abstand" aufgedruckt, so werden alle Messungen mit diesem einen Lineal ein spezifisches falsches Ergebnis liefern. Diese Abweichung ist von systematischer Natur. Ohne eine Wiederholung des Herstellungsprozesses wird sich an der falschen Eichung nichts ändern.*
4. *Schließlich kann die Eichung des Lineals zunächst mit nur geringer Abweichung erfolgt sein, aber das verwendete Material hat einen großen thermischen Ausdehnungskoeffizienten. Findet die Messung bei einer anderen Temperatur statt, so ergibt sich eine Abweichung, die für diese Temperatur reproduzierbar ist, also eine systematische Abweichung. Einen ähnlichen Einfluss kann die Luftfeuchtigkeit haben oder Alterungsprozesse im Trägermaterial.*

Es wird hier sichtbar, wie die Genauigkeit eines so einfachen Vorgangs von den Details abhängt. So ist es möglich die Abweichungen vom Typ 1 und 2 durch eine entsprechend hohe Zahl von Versuchen praktisch zu Null zu machen, vorausgesetzt, der gesamte Vorgang des Anlegens des Lineals wird wiederholt. Liest man nur mehrfach hintereinander den Messwert ab, so wird zwar der Fehler 2 reduziert - sofern man diszipliniert genug ist, bei jeder Ablesung wirklich eine neue Schätzung vorzunehmen - aber der Fehler 1 bleibt bestehen.

Die Abweichung vom Typ 4 ist sehr ambivalent. Hat man keine Informationen über den Zusammenhang zwischen Temperatur und Messwert, so gilt obige Beschreibung. Manchmal ist jedoch der Temperaturgang eines Messgerätes vom Hersteller vermessen. Oder im einfachsten Fall wie dem Lineal ist der Ausdehnungskoeffizient des Trägermaterials bekannt. Kennt man dann auch die Temperatur bei der Messung, kann man diese Abweichung durch eine Korrektur ersetzen und von der Liste der Abweichungen streichen.

Ist die Temperaturabhängigkeit des Messgerätes nicht bekannt, aber man hat die Möglichkeit die Temperatur zu variieren, so kann man eine Mittelung über die Temperatur durchzuführen. Damit wird aus der systematischen Abweichung eine statistische die sich durch eine höhere Zahl von Versuchen (hier bei wechselnden Temperaturen) reduzieren lässt.

3 Abschätzung der Unsicherheit bei Einzelmessungen

Ist es nicht möglich oder sehr aufwendig, eine Messung mehrfach zu wiederholen, so muss die quasistatistische Unsicherheit der entsprechenden Messgröße geschätzt werden.

Als Anhaltspunkte für die statistische Genauigkeit einer Einzelmessung kann im Praktikum gelten:

- Reaktionszeiten (bei einem „unvorhergesehenen“ Ereignis) mit einer Stoppuhr liegen bei 0,2 bis 0,3 Sekunden. Werden Zeit-Differenzen gemessen werden sind die Abweichungen kleiner, da sich die Reaktionszeit bis auf die Schwankungen herausmittelt.

  Ein Abschätzung seiner persönlichen Reaktionszeit kann man erhalten, indem eine Person plötzlich ein Lineal fallen lässt - an der Wand aufgelegt -, das die andere Person mit dem Daumen versucht am ‚Durchfallen‘ zu hindern. Aus Fallstrecke und Erdbeschleunigung ergibt sich die Reaktionszeit $t = \sqrt{2s/g}$.
- Die Position von zwei parallelen versetzten Strichen kann typisch mit einer Genauigkeit von der halben bis ganzen Strichbreite bestimmt werden.
- Die Position eines Strichs zwischen zwei Endwerten kann durch Interpolation auf typisch 20% bis 50% des Abstandes zwischen den Strichen bestimmt werden, bei kleiner Strichbreite geht es besser.

  Zusätzliche Schwierigkeiten sind durch Parallaxe-Probleme möglich. Eine Abschätzung seiner persönlichen Fähigkeit bestimmte Strichpositionen zu ermitteln kann man Mehrfachablesung des gleichen Messwerts durch andere Personen erhalten.

Beim Abschätzen der Unsicherheiten müssen Sie aber zwei Dinge beachten:

1. Schätzungen der statistischen Unsicherheiten gehen **nur** dann in die Berechnung der Gesamtabweichung ein wenn keine genaueren Informationen vorliegen, insbesondere wenn es nicht möglich oder zu aufwendig ist Mehrfachmessungen durchzuführen.
2. Die *systematischen* Unsicherheiten müssen weiterhin extra betrachtet werden.

Angewandt auf das Quader-Beispiel:

Bei der Längenmessung mit dem Lineal ergebe sich als Messwert $l = 355,4$ mm Zur Bestimmung der Genauigkeit von l kann man den Quader etwas verschieben und betrachten, wie weit man in der Lage ist, eine Verschiebung zu registrieren. Da die linke Quaderkante mit einem Strich auf dem Lineal zusammenfällt, ist es möglich Abweichungen in der Größenordnung der Strichdicke also von ± 0.1 mm festzustellen. Die Bestimmung der Abweichung an der rechten Kante ist schwieriger, da ein entsprechender Vergleichsstrich fehlt. Man muss also die Skala zwischen zwei Millimeterstrichen interpolieren. Eine Unterscheidung zwischen der Mittelposition also 1/2 und einer 1/4-Position, erscheint noch möglich. Wesentlich genauer scheint ohne einen weiteren Vergleichsstrich kaum möglich.

Für die geschätzten statistischen Abweichungen erhalten wir also: $\Delta l_{1,\mathrm{stat}} = 0,1$ mm und $\Delta l_{2,\mathrm{stat}} = 0,3$ mm.

4 Statistische Unsicherheit bei Mehrfachmessung

4.1 Mittelwert und Erwartungswert

Wiederholt man Messungen an demselben Messobjekt mit demselben Messgerät unter gleichen Bedingungen, so werden sich die einzelnen Messwerte x_i trotzdem aufgrund der unterschiedlichen statistischen Abweichung voneinander unterscheiden. Sie streuen um einen Mittelwert $\bar{x}$, der sich mit wachsender Zahl n der Messungen dem wahren Wert x_e (dem Erwartungswert der Messgröße) nähert. Wenn keine besonderen Umstände vorliegen die eine spezielle Gewichtung der Abweichung notwendig machen, verwendet man als Wert für die Messgröße das *arithmetische Mittel*:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{2}$$

4.2 Standardabweichung

Das arithmetische Mittel $\bar{x}$ hat die Eigenschaft, dass die Summe der Abweichungen $\Delta x_i = (x_i - \bar{x})$ der Einzelmessungen gerade verschwindet

$$\sum_{i=1}^{n} \Delta x_i = \sum_{i=1}^{n} (x_i - \bar{x}) = 0 \tag{3}$$

Die Summe der Quadrate der Abweichungen dagegen verschwindet nicht, solange auch nur einer der Messwerte vom Bezugswert abweicht:

$$\sum_{i=1}^{n} (\Delta x_i)^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \neq 0 \tag{4}$$

Weiterhin wird diese Summe genau für den Mittelwert $\bar{x}$ als Referenzwert minimal, d.h. ersetzt man $\bar{x}$ durch einen anderen Bezugswert $\tilde{x}$, so wird die Summe größer.

Als Maß für die Streuung der Messwerte um den Mittelwert definiert man mit der Summe der Abweichungsquadrate die *Standardabweichung* σ:

$$\sigma_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{5}$$

σ_x ist eine positive Größe die genau dann zu Null wird wenn alle Messwerte übereinstimmen.

σ_x liefert eine Schätzung der Abweichung vom wahren Wert x_e, der im allgemeinen aber nicht bekannt ist. Daher verwendet man die Abweichung vom Mittelwert $\bar{x}$. Für eine Einzelmessung ist diese Schätzung natürlich nicht durchführbar, was durch den Faktor $1/n-1$ berücksichtigt wird. Für $n = 1$ ist σ nicht definiert.

Die Gleichung der Standardabweichung kann so umgeformt werden, dass der Mittelwert zunächst nicht benötigt wird:

$$\sigma_x = \sqrt{\frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]} \tag{6}$$

Diese Notation bietet rechentechnisch Vorteile, z.B. wenn während der Messwertaufnahme die Unsicherheit direkt bestimmt werden soll, und sich der Mittelwert laufend ändert.

Standardabweichung des Einzelwertes

Die Standardabweichung σ_x beantwortet also die Frage: *Wie weit weicht typisch ein einzelner Messwert x_i vom Mittelwert $\bar{x}$ ab.* Dies ist zur Beurteilung des Messwertes von Interesse. Sie ist eine Größe die – von Schwankungen abgesehen - von der Zahl der Messwerte unabhängig ist.

Standardabweichung des Mittelwertes

Für die Betrachtung der Messunsicherheiten ist die interessierende Frage aber:
Wie weit weicht typisch der Mittelwert $\bar{x}$ aller Messungen vom wahren Wert x_e ab?

Diese Abweichung sollte mit zunehmender Zahl der Messungen immer kleiner werden, der Mittelwert nähert sich immer mehr dem wahren Wert an. Bei einer großen Anzahl von Messwerten gilt für die *Standardabweichung des Mittelwerts*:

$$\sigma_{\bar{x}} = \frac{1}{\sqrt{n}} \cdot \sigma_x \tag{7}$$

Diese Größe sinkt mit der Wurzel der Anzahl der Messwerte.

Betrachten wir wieder die Längenmessung in unserem Quader-Beispiel. Es werden nun 30 Einzelmessungen durchgeführt deren Ergebnisse in Tabelle 1 dargestellt sind. Für den Mittelwert erhalten wir $\bar{l} = 355,62$ mm. Die Standardabweichung des Einzelwerts beträgt $\sigma_x = 0,16$ mm, die des Mittelwerts $\sigma_{\bar{x}} = 0,03$ mm.

4.3 Normalverteilung

Die Standardabweichung ist ein Maß für die mittlere Abweichung der Messwerte vom wahren Wert. Es gibt natürlich Messwerte die weniger abweichen und die mehr abweichen, ihre Häufigkeit ist unterschiedlich. Geht die Zahl der Messwerte $n \to \infty$, so würde jeder einzelnen Messwert mit einer Häufigkeit h vorkommen, der Wahrscheinlichkeitsdichte. Im vorliegenden Fall ist die Dichte durch die Gaußsche Normalverteilung gegeben:

$$h(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \tag{8}$$

Nr.	l_i (mm)	Δl_i (mm)	Nr.	l_i (mm)	Δl_i (mm)	Nr.	l_i (mm)	Δl_i (mm)
1	355,6	-0,02	11	355,6	-0,02	21	355,6	-0,02
2	355,8	0,18	12	355,9	0,28	22	355,7	0,08
3	355,5	-0,12	13	356,0	0,38	23	355,7	0,08
4	355,6	-0,02	14	355,6	-0,02	24	355,5	-0,12
5	355,6	-0,02	15	355,3	-0,32	25	355,4	-0,22
6	355,9	0,28	16	355,7	0,08	26	355,5	-0,12
7	355,5	-0,12	17	355,8	0,18	27	355,7	0,08
8	355,4	-0,22	18	355,6	-0,02	28	355,6	-0,02
9	355,6	-0,02	19	355,4	-0,22	29	355,6	-0,02
10	355,7	0,08	20	355,5	-0,12	30	355,7	0,08

Tabelle 1: Messwerte bei 30-facher Längenmessung eines Werkstückes. Die Abweichung Δl_i ist die Differenz des jeweiligen Messwertes vom Mittelwert ($\Delta l_i = l_i - \bar{l}$).

Die Gauß-Funktion („Glockenkurve“) ist symmetrisch um den Mittelwert $\bar{x}$ und so normiert, dass das Integral von $h(x)$ zwischen $-\infty$ und ∞ gerade 1 ergibt.

Die Häufigkeitsverteilung der Messwerte des Quader-Beispiels aus Tabelle 1 und die dazugehörige Normalverteilung ist in Abb. 2 dargestellt.

Die Breite der Häufigkeitsverteilung wird durch die Standardabweichung σ bestimmt, je. größer diese ist, desto breiter die Kurve. Im Abstand σ vom Mittelwert aus gerechnet befindet sich der Wendepunkt der Kurve. Berechnet man die Fläche bis zu diesem Wert, so ergibt sich:

$$\frac{1}{\sqrt{2\pi}\,\sigma} \int_{\bar{x}-\sigma}^{\bar{x}+\sigma} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}\, dx = 0,683 \tag{9}$$

Die Wahrscheinlichkeit, einen Messwert im Intervall von $\bar{x} - \sigma$ bis $\bar{x} + \sigma$ zu finden beträgt (für $n \to \infty$) also 68,3% .

Andere gebräuchliche Intervalle sind:

1-fache Standardabweichung	$\bar{x} - \sigma$ bis $\bar{x} + \sigma$	(68,3%)
2-fache Standardabweichung	$\bar{x} - 2\sigma$ bis $\bar{x} + 2\sigma$	(95,5%)
3-fache Standardabweichung	$\bar{x} - 3\sigma$ bis $\bar{x} + 3\sigma$	(99,7%)

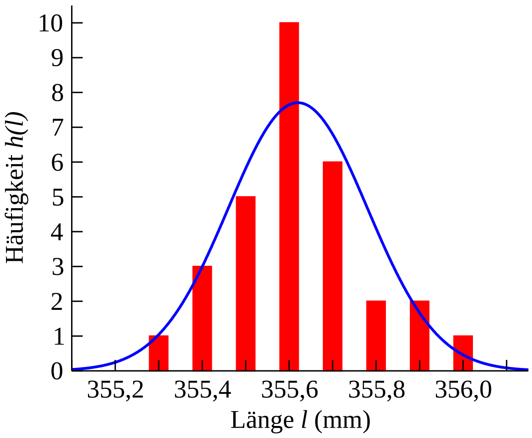

Abbildung 2: Histogramm der Längenmessung mit zugehöriger Normalverteilung

n	$1-\alpha = 68,2\%$ t	$t/\sqrt{n}$	$1-\alpha = 95\%$ t	$t/\sqrt{n}$	$1-\alpha = 99,7\%$ t	$t/\sqrt{n}$
2	1,84	1,30	12,71	8,98	235,8	166,7
3	1,32	0,76	4,30	2,48	19,21	11,09
4	1,20	0,60	3,18	1,59	9,22	4,61
5	1,15	0,51	2,78	1,24	6,62	2,96
6	1,11	0,45	2,57	1,05	5,51	2,25
8	1,08	0,38	2,37	0,84	4,53	1,60
10	1,06	0,34	2,26	0,71	4,09	1,29
20	1,03	0,23	2,09	0,48	3,45	0,77
30	1,02	0,19	2,05	0,37	3,28	0,60
50	1,01	0,14	2,01	0,28	3,16	0,45
100	1,00	0,10	1,98	0,20	3,08	0,31
200	1,00	0,07	1,97	0,14	3,04	0,21

Tabelle 2: Werte für die Studentfunktion t und $t/\sqrt{n}$ bei verschiedenen Vertrauensniveaus.

4.4 Die statistischen Unsicherheit einer Messgröße

Üblicherweise wird die einfache Standardabweichung des Mittelwerts als statistische Unsicherheit angegeben:

$$\Delta\bar{x} = \sigma_{\bar{x}} = \frac{1}{\sqrt{n}} \cdot \sigma_x \tag{10}$$

Der Bereich $[\bar{x}-\sigma_{\bar{x}}; \bar{x}-\sigma_{\bar{x}}]$ heißt *Konfidenzbereich auf dem Vertrauensniveau 68,3%.*

Gibt man die Unsicherheit auf anderen Vertrauensniveaus an, z.B. 95,5% oder 99,7%, also $\Delta\bar{x} = 2\,\sigma_{\bar{x}}$ oder $\Delta\bar{x} = 3\,\sigma_{\bar{x}}$, dann ist dies ausdrücklich anzumerken.

4.5 Korrektur für wenige Messwerte

Liegen nur wenige Messwerte vor, muss statt der Normalverteilung die Student-t-Verteilung[2] verwendet werden. Den Unterschied beider Verteilungen berücksichtigen wir durch einen Korrekturfaktor t, der aus der Tabelle 2 entnommen werden kann.

Für die Unsicherheit des Mittelwerts gilt dann:

$$\Delta\bar{x} = \frac{t}{\sqrt{n}}\,\sigma_x = t \cdot \sigma_{\bar{x}} \tag{11}$$

Wie zu erwarten ergibt sich z.B. für $n \to \infty$ und 68,3% Vertrauensniveau t = 1 oder für $n \to \infty$ und 99,7% Vertrauensniveau t = 3, d.h. die Korrektur verschwindet. Für kleine Zahlen und hohe Vertrauensniveaus hingegen wird die Korrektur erheblich.

[2] William Sealey Gosset veröffentlichte seine Arbeit unter dem Pseudonym *Student*, da sein damaliger Arbeitgeber eine Veröffentlichung nicht genehmigte.

Für unser Quader-Beispiel ergeben sich also folgende Werte:
$n = 30$, $\bar{x} = 355,62$ *mm*, $\sigma_x = 0,16$ *mm. Für ein Vertrauensniveau von 68,3% ist der Korrekturfaktor* $t = 1,02$, *und damit* $\Delta\bar{x} = 0,03$ *mm.*
Mit einer Wahrscheinlichkeit von 68,3% liegt also die wahre Quaderlänge l *zwischen 355,59 mm und 355,65 mm.*

Will man sicher gehen und setzt das Vertrauensniveau auf 99,7%, so ist $t = 3,28$ *und man erhält* $\Delta\bar{x} = 0,10$ *mm*

Es sei hier noch betont, dass keineswegs jede zufällige Abweichung einer Gaußschen Normalverteilung folgen muss. Ob das im konkreten Fall zutrifft kann durch geeignete Testverfahren (z.B. den χ^2-Test) ermittelt werden, für die jedoch auf die Literatur verwiesen wird.

5 Systematische Unsicherheiten

In den beiden letzten Kapiteln wurden statistische Abweichungen behandelt, einerseits die Abweichungen der Einzelmessung, die lediglich geschätzt werden können, und andererseits Abweichungen bei Mehrfachmessungen soweit sie einer Gaußschen Normalverteilung gehorchen.

Ein mindestens ebenso wichtiger Teil sind die systematischen Unsicherheiten. Grundsätzlich gehört jedoch viel Erfahrung zum Erkennen diese Abweichungen

Es gibt kein Kriterium, festzustellen, ob man alle Unsicherheiten erkannt hat, aber durch systematische Analyse können die meisten Unsicherheiten abgeschätzt werden.

Da grundsätzlich Messwerte nur einen Aussagewert haben wenn ihre Genauigkeit und der Gültigkeitsbereich bekannt sind, macht diese prinzipiell fehlende Ausschlussmöglichkeit unerkannter Messabweichungen physikalische Messungen zu Aussagen, die von der Erfahrung des Messenden und Auswertenden abhängen. Es ist daher gerade bei wenig experimenteller Erfahrung wesentlich, verstärkt das Augenmerk auf das Erkennen systematischer Abweichungen zu richten.

Im Folgenden sollen einige Anhaltspunkte zum Erkennen systematischer Unsicherheiten besprochen werden.

5.1 Unsicherheiten von Messgeräten

Messgeräte haben eine endliche, oft durch die Bauart begrenzte Genauigkeit. In vielen Fällen liegen Angaben des Messgeräte-Herstellers vor, aus denen direkt die Genauigkeit des Messgerätes hervorgeht. Dabei sind einige wichtige Begriffe zu beachten:

Die Auflösung: Dies ist die offensichtlichste Eigenschaft eines Messgerätes. Wenn auf einer digitalen Anzeige, z.B. nur 4 Stellen angegeben sind, dann ist unmöglich mehr Stellen abzulesen. Diese kleinste Einheit der letzten Stelle eines digitalen Messgerätes wird auch als ‚1 Digit' bezeichnet. Die Auflösung eines digitalen Messgerätes beträgt also im Normalfall ± 1 Digit. Ein analoges Messgerät hat natürlich auch eine Auflösung, sie ergibt sich aus dem Abstand zwischen zwei Zeigerpositionen die mit dem Auge noch gut unterschieden werden können.

Die Reproduzierbarkeit: Wenn man mehrfach eine Messung unter absolut gleichen Bedingungen wiederholt, so ist die Reproduzierbarkeit der maximale Unterschied dieser Messungen. Es ist sozusagen der (pseudo-)statistische Anteil der systematischen Abweichung. Der statistische Beitrag des Messenden ist hier abgetrennt.

Diese ersten beiden Unsicherheiten sind allerdings eher wie statistische Abweichungen zu behandeln. Der Vollständigkeit halber stehen Sie trotzdem hier bei den Unsicherheiten von Messgeräten

Meist werden Messgeräte an zwei Punkten kalibriert, bei Null und bei Vollausschlag. Daraus resultieren entsprechende Unsicherheiten:

Der Skalierungsfehler: Damit bezeichnet man die Abweichung eines Messgerätes vom Sollwert bei Vollausschlag. Hier gehen Fehler des Eichnormals ein, oder ungenau definierte Bauelemente. Dieser Typ trägt mit wachsendem Messwert proportional zur Gesamtabweichung bei. Entsprechend wird er meist durch eine Prozentangabe beschrieben, also eine relative Abweichung.

Der Nullpunktsfehler: Da am Nullpunkt definitionsgemäß der Messwert 0 vorliegen soll, spielt das Eichnormal keine Rolle mehr. Übrig bleibt z.B. die Temperatur- oder Alterungsdrift der Bauteile. Bei den meisten modernen Messgeräten ist diese Abweichung deutlich kleiner als die Abweichung bei Vollausschlag. Vom Typ her ist dies eine absolute Abweichung, die in vielen Fällen korrigiert werden kann.

Die Nichtlinearität: Meist berechnet man die systematische Abweichung eines Messgerätes aus

$$\text{Nullpunktsfehler} + \text{Messwert} \cdot \text{Skalierungsfehler}.$$

Dies ist äquivalent zu einer linearen Näherung. Das Kalibrieren am Nullpunkt und am Skalenendpunkt bietet natürlich noch nicht die Gewähr für ein lineares Verhalten dazwischen. Die daraus entstehenden Unsicherheiten werden als Nichtlinearitäten bezeichnet.

Die Fehlerklasse: Diese Art der Angabe von Unsicherheiten ist typisch für analoge Messgeräte. Klasse 1 entspricht 1 % vom Vollausschlag im entsprechenden Messbereich als Unsicherheit. Klasse 1 bei 30 V Vollausschlag bedeutet also eine absolute Ungenauigkeit von 0,3 V in gesamten Messbereich, unabhängig vom Messwert. Oft sind diese Angaben noch an eine Gebrauchslage des Gerätes gekoppelt.

5.2 Einfluss durch Umgebungsparameter

Durch den Einfluss der Umgebung können sowohl die Messapparatur als auch das zu messende System selbst beeinflusst werden. Die daraus resultierenden Unsicherheiten sind am schwierigsten greifbar, da die Umgebungsparameter zum Teil nur sehr schwer zu ermitteln und zu beeinflussen sind.

Beispiele für solche Umgebungsparameter sind die Temperatur, die Luftfeuchtigkeit, elektrische und magnetische Streufelder (Erdmagnetfeld, Felder von Stromkabeln), Streulicht, ...
Aber auch die Alterung von Geräten und Bauteilen gehören zu dieser Rubrik.

5.3 Beeinflussung des Messwerts durch die Messung

Das System, dessen Eigenschaften man bestimmen will, kann durch die Messprozedur verändert werden. Mann misst dann nicht die Eigenschaften des Systems, sondern die des Systems mit Messvorrichtung.

Diese Beeinflussung kann in den meisten Fällen beliebig klein gemacht werden, oder durch Korrekturen berücksichtigt werden.

Ein Beispiel hierfür ist die Bestimmung von Strom und Spannung in einer elektrischen Schaltung. Der endliche Widerstand eines normalen Spannungsmessgerätes führt dann zu einem nichtverschwindenden Beitrag in der Strommessung. Andersherum führt der nichtwerschwindende Widerstand im Strommessgerät zu einem verfälschten Wert bei der Spannungsmessung. Es ist so nicht möglich Strom und Spannung an einem elektrischen Bauteil gleichzeitig genau zu bestimmen.

Besondere Bedeutung hat dies auch in der Quantenmechanik, wo die Messung einer Größe immer das System beeinflusst.

6 Vorgehen zum Aufspüren von Fehlerquellen

In den vorangegangenen Kapiteln wurden diverse mögliche Fehlerquellen diskutiert, aber auch gesagt, dass es keine verlässliche Methode gibt um systematische Abweichungen zu detektieren. Durch die folgende Vorgehensweise kann man aber schon einen Großteil der Unsicherheiten erfassen:

Systematische Unsicherheiten:

- Nähern Sie die Unsicherheit durch einen festen Anteil (Nullpunkt) und einen skalierenden Anteil (Skala). Das entspricht einer linearen Näherung
- Eliminieren Sie die Nullpunktsabweichung durch eine Kalibrierprozedur. Zeigt z.B. ein Messgerät konstant 0,1 V an, so ziehen Sie diesen Wert vom Messergebnis ab.

- Legen Sie für den Skalenfehler einen Wert fest, aus Herstellerangaben, aus der Praktikumsanleitung, durch Fragen des Betreuers, durch Schätzung.

Statistische Unsicherheiten:

- Reduzieren Sie den statistischen Anteil der Unsicherheit durch eine ausreichende Zahl von Versuchen.
- Bilden Sie bei nicht abhängigen Größen Mittelwerte, verwenden Sie bei korrelierten Größen eine Regressionsgerade zur Verknüpfung der Einzelmessungen.
- Bestimmen Sie die statistische Unsicherheit aus der Standardabweichung des Mittelwerts, falls Sie mehrere Messungen zur Verfügung haben.
- Berechnen Sie die Unsicherheiten von Steigung und Achsabschnitt bei linearer Regression.
- Schätzen Sie die statistische Unsicherheit, falls das Erheben vieler Einzelmessungen nicht möglich ist oder zu viel Aufwand bedeutet.

7 Fehlerfortpflanzung

Bereits in dem ersten sehr simplen Beispiel in dem die Länge eines Quaders gemessen wurde zeigte sich, dass sich Messabweichungen aus einzelnen Beiträgen zusammensetzen. Man ist jedoch meist an der Gesamtgenauigkeit einer Größe interessiert die sich mittels einer Formel aus einzelnen Messgrößen ergibt und es stellt sich die Frage nach der Verknüpfung der einzelnen Abweichungen.

7.1 Größtfehlerbetrachtung

Der einfachste Ansatz ist:

> Alle Abweichung aller beteiligten Größen tragen voll zur Gesamtabweichung bei und beeinflussen sich nicht gegenseitig.

Rechnerisch umgesetzt wird dies indem im ersten Schritt für jede einzelne Größe x_i eine Gesamtabweichung Δx_i aus allen bekannten Abweichung gebildet wird.

Im nächsten Schritt wird in der Formel aus der sich die resultierende Größe ergibt jede beteiligte Größe x_i so eingesetzt, dass das Ergebnis maximal wird. Dieser Vorgang wird wiederholt mit dem Ziel das Ergebnis zu minimieren.

Erweitern wir unser Quader-Beispiel:
Statt einer Kantenlänge soll die Dichte bestimmt werden. Dazu misst man die Ausdehnung in den drei Dimensionen und die Masse.

$$\rho = \frac{m}{x \cdot y \cdot z} \tag{12}$$

Um die maximale Dichte zu erhalten muss man berechnen:

$$\bar{\rho}_+ = \frac{\overline{m} + \Delta m}{(\bar{x} - \Delta x) \cdot (\bar{y} - \Delta y) \cdot (\bar{z} - \Delta z)} \tag{13}$$

Die minimale Dichte ergibt sich aus

$$\bar{\rho}_- = \frac{\overline{m} - \Delta m}{(\bar{x} + \Delta x) \cdot (\bar{y} + \Delta y) \cdot (\bar{z} + \Delta z)} \tag{14}$$

Für die Abweichungen der Einzel-Größen wird dabei jeweils die Aufsummierung der Abweichungen eingesetzt (die Nummerierung entspricht der Aufzählung zu Beginn) :

$$\Delta\bar{x} = \Delta x_{1,\mathrm{stat}} + \Delta x_{2,\mathrm{stat}} + \Delta x_{1,\mathrm{syst}} + \Delta x_{2,\mathrm{syst}} \tag{15}$$

Dieses Verfahren berücksichtigt weder bei der Aufsummierung noch bei der Fortpflanzung der Abweichung den statistischen Zusammenhang der einzelnen Beiträge wie er z.B. in der Gaußschen Normalverteilung zum Ausdruck kommt. Es liefert daher im Allgemeinen zu große Werte.

Darüber hinaus ergeben sich technische Probleme bei der Berechnung des größten oder kleinsten Wertes wenn z.B. die entsprechende Größe in einer Summe sowohl im Zähler als auch im Nenner steht. Es ist dann unklar ob ein Minus oder ein Plus-Zeichen den kleinsten bzw. größten Wert liefert. Dieses Verfahren ist daher nur für Notfälle geeignet, liefert jedoch oft eine brauchbare Abschätzung.

7.2 Fortpflanzung mittels partieller Ableitungen

Wir betrachten eine gesuchte Größe g, die nicht direkt gemessen werden kann, sich aber aus mehreren direkt gemessenen Größen x, y, z, … berechnen lässt:

$$g = g(x, y, z, \ldots) \tag{16}$$

Wir gehen davon aus, dass die Messserien für die Größen[3] x, y und z normalverteilt sind. Dann kann der Mittelwert $\bar{g}$ für die gesuchte Größe g bestimmt werden, indem die einzelnen Mittelwerte in die Funktion g einsetzt werden:

$$\bar{g} = g(\bar{x}, \bar{y}, \bar{z}) \tag{17}$$

Sind die Messgrößen x, y und z voneinander statistisch unabhängig, so kann die Gesamtabweichung aus folgender Formel bestimmt werden:

$$\Delta\bar{g} = \sqrt{\Delta\bar{x}^2 \left[\frac{\partial g}{\partial x}\right]^2_{\bar{x},\bar{y},\bar{z}} + \Delta\bar{y}^2 \left[\frac{\partial g}{\partial y}\right]^2_{\bar{x},\bar{y},\bar{z}} + \Delta\bar{z}^2 \left[\frac{\partial g}{\partial z}\right]^2_{\bar{x},\bar{y},\bar{z}}} \tag{18}$$

Obige Formel ist unter der Bezeichnung *Gaußsche Fehlerfortpflanzung* für den Spezialfall der Standardabweichung bekannt.

[3] Wegen der besseren Übersichtlichkeit beschränken wir uns hier auf die drei Größen x, y und z

Die Ausdrücke $\partial g/\partial x$, $\partial g/\partial y$ und $\partial g/\partial z$ stehen für die partiellen Ableitungen der Funktion $g(x, y, z)$. Deren Quadrate, wobei die jeweiligen Werte aller Variabler eingesetzt werden, bilden die *Wichtungsfaktoren* für die einzelnen Messunsicherheiten. Diese beschreiben den Einfluss der *Einzel*abweichung auf die *Gesamt*abweichung.

Betrachten wir wieder unser Quader-Beispiel bei dem wir die Dichte bestimmen wollen: Die partielle Ableitung nach der Masse m lautet:

$$\frac{\partial \rho}{\partial m} = \frac{1}{x\,y\,z} \tag{19}$$

Die partielle Ableitung nach x lautet (die anderen Dimensionen entsprechend):

$$\frac{\partial \rho}{\partial x} = \frac{-1}{x}\,\frac{m}{x\,y\,z} \tag{20}$$

Damit ergibt sich für die Unsicherheit der Dichte:

$$\Delta\bar{\rho} = \sqrt{\Delta\bar{x}^2\left(\frac{-1}{x}\,\frac{\overline{m}}{\bar{x}\bar{y}\bar{z}}\right)^2 + \Delta\bar{y}^2\left(\frac{-1}{y}\,\frac{\overline{m}}{\bar{x}\bar{y}\bar{z}}\right)^2 + \Delta\bar{z}^2\left(\frac{-1}{z}\,\frac{\overline{m}}{\bar{x}\bar{y}\bar{z}}\right)^2 + \Delta\overline{m}^2\left(\frac{1}{\bar{x}\bar{y}\bar{z}}\right)^2} \tag{21}$$

In dieser Gleichung müssen jetzt durch Einsetzen der Mittelwerte in die partiellen Ableitungen die Wichtungsfaktoren bestimmt werden. Der Rest ist Fußarbeit.

Wichtige Konsequenzen dieses Ansatzes:

- Zur Minimierung des Messaufwandes sollte man bei Verbesserungen an der Stelle ansetzen an der das Produkt aus Abweichung und Wichtungsfaktor am größten ist.

- Die Wichtungen sind nicht dimensionslos!
  Dies erscheint zunächst nebensächlich, hat jedoch massive Konsequenzen:

  Als Beispiel sei der Einfluss der Winkelabweichung auf eine Größe gegeben,wobei in der Gleichung die Winkelfunktion $\sin(\alpha)$ auftrete. Leitet man diese Funktion nach α ab, so ergibt sich $\cos(\alpha)$. Der Messwert habe eine Unsicherheit von $\Delta\alpha = 0,1°$. Setzt man jetzt scheinbar völlig konsistent bei der Berechnung des Wichtungsfaktors und der Abweichung die Werte jeweils in Grad ein, so ist das Ergebnis trotzdem um den Faktor 57 falsch!

  Der Grund, die Bildung der Ableitung erfolgte in rad! Zum Rechnen in Grad ist es notwendig, bereits beim Ableiten $\sin(\alpha)$ durch $\sin(2\pi/360 \cdot r)$ zu ersetzen, wobei r der Winkel α im Winkelmaß ist. Beim Nachdifferenzieren tritt dann der Korrekturfaktor $2\pi/360$ auf.

Obige Fortpflanzung der Abweichungen liefert gegenüber der Größtfehlerbetrachtung in zweierlei Hinsicht eine Verbesserung:

1. Man spart das primitive und manchmal nicht durchführbare Einsetzen in die Gleichung durch Berechnung einer Wichtung durch partielles Ableiten. Dies ist in allen Anwendungsfällen sinnvoll.

2. Man ersetzt das oft überbewertende ‚lineare Addieren' durch ein ‚quadratisches Addieren', das die Gaußsche Normalverteilung berücksichtigt.

   Falls keine statistische Normalverteilung vorliegt, muss man die Wichtung mit einer linearen Addition kombinieren.

7.3 Spezialfälle der Fortpflanzung der Abweichungen

In der Praxis kann sich die Ermittlung der partiellen Ableitungen recht aufwendig gestalten. Vereinfachungen ergeben sich wenn der Zusammenhang zwischen der gesuchten Größe g und den direkt gemessenen Größen durch bestimmte einfache Funktionen gegeben ist:

7.3.1 Summen oder die Differenzen

Errechnet sich die gesuchte Größe nur aus Summen und Differenzen von Messgrößen

$$g = \pm x \pm y \pm z \tag{22}$$

dann sind die partiellen Ableitungen $+1$ oder -1, also gilt:

$$\Delta\bar{g} = \sqrt{\Delta\bar{x}^2 + \Delta\bar{y}^2 + \Delta\bar{z}^2} \tag{23}$$

Alle Wichtungsfaktoren sind hier 1. Die Gesamtabweichung ist die quadratische Summe der absoluten Einzelabweichungen.

Ein wichtiges Beispiel:
Auf einer Skala liest man an zwei Positionen den Messwert ab und ist an der Differenz interessiert:

$$g = \bar{x}_2 - \bar{x}_1 \tag{24}$$

Beide Messwerte haben die gleiche Genauigkeit $\Delta\bar{x}_1 = \Delta\bar{x}_1 = \Delta\bar{x}$. Dann gilt

$$\Delta\bar{g} = \sqrt{\Delta\bar{x}^2 + \Delta\bar{y}^2} = \sqrt{2} \cdot \Delta\bar{x} \tag{25}$$

7.3.2 Produkte oder der Quotienten

Errechnet sich die gesuchte Größe nur aus Produkten oder Quotienten von Messgrößen

$$g = \frac{x \cdot y}{z} \tag{26}$$

so sind Wichtungsfaktoren $\frac{g}{x}$, $\frac{g}{y}$, $\frac{g}{z}$. Es gilt dann:

$$\frac{\Delta\bar{g}}{\bar{g}} = \sqrt{\left(\frac{\Delta\bar{x}}{\bar{x}}\right)^2 + \left(\frac{\Delta\bar{y}}{\bar{y}}\right)^2 + \left(\frac{\Delta\bar{z}}{\bar{z}}\right)^2} \tag{27}$$

Die relative Gesamtabweichung ist die quadratische Summe der relativen Einzelabweichungen.

Dies lässt sich erweitern auf den Fall, dass die Messgrößen in unterschiedlichen Potenzen vorkommen:

$$g = \frac{x^k \cdot y^l}{z^m} \tag{28}$$

Dann gilt:

$$\frac{\Delta\bar{g}}{\bar{g}} = \sqrt{k^2 \cdot \left(\frac{\Delta\bar{x}}{\bar{x}}\right)^2 + l^2 \cdot \left(\frac{\Delta\bar{y}}{\bar{y}}\right)^2 + m^2 \cdot \left(\frac{\Delta\bar{z}}{\bar{z}}\right)^2} \tag{29}$$

Außerdem sollte man beachten:
Soweit irgend möglich ist es zweckmäßig die Fortpflanzung der Abweichungen in Einzelschritte zu zerlegen, die sich durch Addition der absoluten oder der relativen Unsicherheiten jeweils einfach durchführen lassen.

Man erspart sich dadurch das oft mühsame und fehleranfällige partiellen Ableiten.

Noch einmal Dichte in unserem Quader-Beispiel:

Wir nehmen an, dass jede Kante des Quaders durch Ablesen eines beliebig aufgelegten Maßstabs gemessen wird. Dann kann man dem absoluten Messfehler einer Kante wie in 7.3.1 aus den beiden Einzelmessungen bei Position 1 und 2 mit gleicher Messabweichung zusammensetzen: $\Delta\bar{x} = \sqrt{2}\,\Delta\bar{x}_1$. *Die relative Messabweichung einer Kante ergibt sich aus den beiden Messungen zu:*

$$\frac{\Delta\bar{x}}{\bar{x}} = \sqrt{2}\,\frac{\Delta\bar{x}_1}{\bar{x}_2 - \bar{x}_1} \tag{30}$$

Die anderen Kanten entsprechend. Damit ergibt sich die Gesamtabweichung der Dichte:

$$\frac{\Delta\bar{\rho}}{\bar{\rho}} = \sqrt{2 \cdot \left(\frac{\Delta\bar{x}_1}{\bar{x}_2 - \bar{x}_1}\right)^2 + 2 \cdot \left(\frac{\Delta\bar{y}_1}{\bar{y}_2 - \bar{y}_1}\right)^2 + 2 \cdot \left(\frac{\Delta\bar{z}_1}{\bar{z}_2 - \bar{z}_1}\right)^2 + \left(\frac{\Delta\overline{m}}{\overline{m}}\right)^2} \tag{31}$$

Nehmen wir an die weiteren Kanten der Quader wären jeweils nur 1/2 bzw. 1/3 so lang, der relative Fehler also entsprechend doppelt oder dreifach so groß, z.B.:

$$\frac{\Delta\bar{x}}{\bar{x}} = \frac{0{,}03\text{ mm}}{355{,}62\text{ mm}} = 0{,}008\%,\ \frac{\Delta\bar{y}}{\bar{y}} = 0{,}016\%,\ \frac{\Delta\bar{z}}{\bar{z}} = 0{,}024\% \tag{32}$$

Dann ergibt sich für den Anteil der Kantenmessung aus obigem Beispiel eine Unsicherheit für das Volumen:

$$\frac{\Delta\bar{V}}{\bar{V}} = \sqrt{2 \cdot 0{,}008\%^2 + 2 \cdot 0{,}016\%^2 + 2 \cdot 0{,}024\%^2} = 0{,}042\% \tag{33}$$

Dazu kommt jetzt die Unsicherheit der Massenbestimmung. Unser Goldbarren hat ein Volumen von ca. $7{,}496\text{ dm}^3$. *Die Dichte von Gold ist* $19{,}320\text{ kg/dm}^3$ *(die von Wolfram* $19{,}350\text{ kg/dm}^3$*) . Der Klotz hat also stolze 144,82 kg Masse. Damit scheidet eine Analysenwaage eindeutig aus. Die Frage lautet: „Wie genau muss unsere Waage sein um zwischen Gold und Wolfram unterscheiden zu können?“*

Wir benötigen eine Genauigkeit der Dichte von $\frac{\Delta\bar{\rho}}{\bar{\rho}} = 1 - \frac{19{,}320}{19{,}350} = 0{,}0016$, *wenn wir annehmen, der ganze Quader bestünde aus Gold oder Wolfram. Also*

$$\frac{\Delta\overline{m}}{\overline{m}} = \sqrt{\left(\frac{\Delta\bar{\rho}}{\bar{\rho}}\right)^2 - \left(\frac{\Delta\bar{V}}{\bar{V}}\right)^2} = \sqrt{0{,}0016^2 - 0{,}0004^2} = 0{,}15\% \tag{34}$$

Eine Waage die deutlich genauer als 0,15% misst bei 150 kg Traglast? Das scheint machbar, es entspricht 100 g Absolut-Genauigkeit. Eine normale Personenwaage, selbst eine geeichte, *reicht allerdings dafür schon nicht mehr. Nimmt man hingegen an, der Fälscher hätte nur 10% des Goldes durch Wolfram ersetzt, so müsste die Waage auf 10 g genau sein. Gelänge das nicht, würde der Verlust von 15 kg Gold mit etwa 225.000.- € zu Buche schlagen (Der Goldpreis liegt bei typisch 15 €/g, der von Wolfram bei 18 €/kg[4]).*

Aus dem obigen Beispiel kann man erkennen, dass es Nichts bringt eine der Größen sehr genau zu messen und die anderen nicht. Wir hatten jeweils 30 Messungen aller Kanten durchgeführt und das bei weniger als $\pm 0,5$ K Temperaturdifferenz, eine sehr aufwendige Aufgabe. Wenn wir dem keine entsprechend genaue Waage an die Seite stellen können war die ganze Mühe umsonst.

8 Berücksichtigung der statistischen und systematischen Unsicherheiten

Im letzten Abschnitt wurde die Fortpflanzung der statistischen Unsicherheiten beschrieben. Hier bleibt nun noch die Frage, wie man nun zusätzliche systematischen Abweichungen behandelt.

Prinzipiell ist es notwendig, zunächst die statistischen und die systematischen Anteile einzeln zu bestimmen, und dann das weitere Vorgehen danach auszurichten.

8.1 Fortpflanzung systematischer Abweichungen

Systematische Abweichungen entziehen sich naturgemäß einer statistischen Behandlung. Man kann insbesondere nicht davon ausgehen, dass es unwahrscheinlich ist, dass alle Abweichungen der Beteiligten Einzelgrößen seien gleichzeitig extremal.

Wieder unser Quader-Beispiel:

Eine Veränderung der Umgebungstemperatur bei der Messung zieht eine Abweichung aller Längenmessungen gemeinsam nach sich, da der Maßstab natürlich bei jeder Messung ‚falsch' ist. Man muss also in diesem Fall davon ausgehen, dass die Abweichungen nicht quadratisch addiert *werden dürfen, sondern* linear addiert *werden müssen. Der Einfluss jeder Einzelabweichung muss voll auf die Gesamtabweichung durchschlagen.*

Im Fall einer systematischen Abweichung der Längenmessung durch einen nicht gleichmäßig geteilten Maßstab erscheint es jedoch sinnvoll anzunehmen, dass je nach Position die Abweichung unterschiedlich ist. Entsprechend kann man auch davon ausgehen, dass nicht gerade alle Abweichungen zum gleichen Zeitpunkt in die gleiche Richtung gehen. Hier ist also eine quadratische Addition sinnvoll.

Es hängt also vom Einzelfall ab welche Methode der Fortpflanzung bei einer systematischen Abweichung angewandt werden muss. Die lineare Addition bei potentiell statistisch gekoppelten Vorgängen oder die quadratische Addition bei definitiv statistisch unabhängigen Vorgängen.

[4] Stand: August 2007

Lineare Addition

$$\Delta\bar{g} = \left| \Delta\bar{x} \left[\frac{\partial g}{\partial x} \right]_{\bar{x},\bar{y},\bar{z}} \right| + \left| \Delta\bar{y} \left[\frac{\partial g}{\partial y} \right]_{\bar{x},\bar{y},\bar{z}} \right| + \left| \Delta\bar{z} \left[\frac{\partial g}{\partial z} \right]_{\bar{x},\bar{y},\bar{z}} \right| \tag{35}$$

Quadratische Addition

$$\Delta\bar{g} = \sqrt{ \Delta\bar{x}^2 \left[\frac{\partial g}{\partial x} \right]^2_{\bar{x},\bar{y},\bar{z}} + \Delta\bar{y}^2 \left[\frac{\partial g}{\partial y} \right]^2_{\bar{x},\bar{y},\bar{z}} + \Delta\bar{z}^2 \left[\frac{\partial g}{\partial z} \right]^2_{\bar{x},\bar{y},\bar{z}} } \tag{36}$$

Im Einzelfall kann es schwierig sein, zu entscheiden, welche Methode die richtige ist.

8.2 Verknüpfung systematischer und statistischer Unsicherheiten

Die gleiche Problematik gilt auch für die Verknüpfung von systematischer und statistischer Abweichung. Die Frage

Ist die statistische Abweichung von der systematischen statistisch unabhängig?

ist letztlich nicht beantwortbar. Entsprechend werden in der Literatur auch beide Meinungen vertreten.

Im Praktikum soll aus Gründen der Vereinfachung folgender Algorithmus gelten, der in Abbildung 3 skizziert ist:

1. Bestimmung der statistischen Abweichungen aus allen Einzelgrößen: mittels *quadratischer Addition*
2. Bestimmung der systematischen Abweichungen aus allen Einzelgrößen: mittels *linearer Addition*
3. Verknüpfung der statistischen und systematischen Abweichung: mittels *linearer Addition*

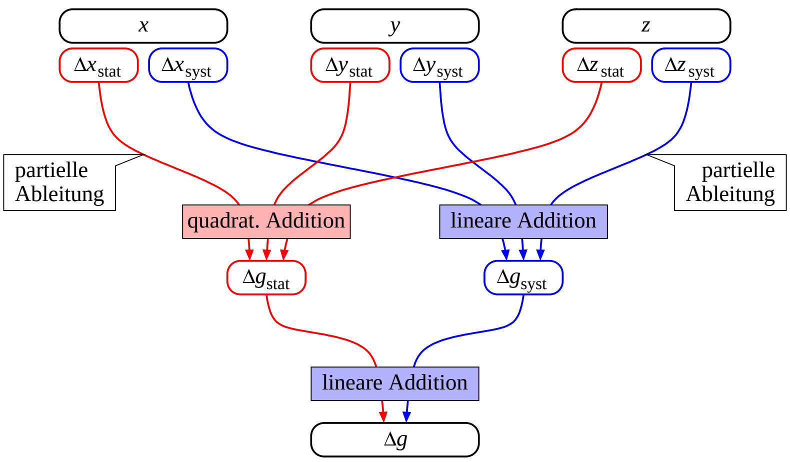

Abbildung 3: Algorithmus zur Ermittlung der Gesamtunsicherheit aus statistischen und systematischen Unsicherheiten der Einzelgrößem x, y und z.

9 Verknüpfung mehrerer Werte mit unterschiedlichen Abweichungen, der gewichtete Mittelwert

Hat man für eine Messgröße mehrere Werte zur Verfügung so scheint dies zunächst unproblematisch, es ergibt sich die Möglichkeit der Mittelwertbildung, die –egal welche Abweichung dominiert– sicher ein genaueres Ergebnis zu liefern scheint als jeder der Einzelwerte. Die einfache Mittelwertbildung setzt jedoch voraus, dass jeder der Messwerte die gleiche Genauigkeit aufweist.

Angenommen eine Person misst eine Länge einmal und eine weitere Person führt eine Messreihe mit 100 Messungen durch. Bildet man in diesem Fall den einfachen Mittelwert, so weicht das Ergebnis mit sehr hoher Wahrscheinlichkeit weiter vom wahren Wert ab als der Wert der 100-fachen Mehrfachmessung. Es gilt also die Beiträge bei der Mittelwertbildung geeignet zu wichten. Dazu dient der *gewichtete Mittelwert*:

$$\bar{x} = \frac{\sum\limits_{i=1}^{n} w_i \cdot x_i}{\sum\limits_{i=1}^{n} w_i} \tag{37}$$

wobei die Wichtungsfaktoren w_i die Kehrwerte der Abweichungsquadrate sind:

$$w_i = \frac{1}{\Delta {x_i}^2} \tag{38}$$

Für die Genauigkeit des gewichteten Mittelwerts gilt:

$$\Delta\bar{x} = \sqrt{\frac{1}{\sum\limits_{i=1}^{n} w_i}} \tag{39}$$

Sind alle Wichtungsfaktoren w_i gleich, geht der gewichtete Mittelwert in den ‚normalen' über, und dessen Abweichung sinkt wieder mit der Wurzel der Zahl der Werte.

Der gewichtete Mittelwert gilt strenggenommen nur nur für normalverteilte Größen. Er wird aber mangels Alternativen auch auf andere statistische oder systematische Abweichungen angewandt. Vor einer Anwendung ist aber zu prüfen, ob die unterschiedlichen Messungen mit dem zugehörigen Konfidenzbereich überlappen. Falls dies nicht der Fall ist, liegt eine zusätzliche, nicht erkannte Abweichung vor (z.B. grobe Fehler).

10 Korrelierte Messwerte

Alle bisherigen Betrachtung gehen von Messgrößen aus, die jeweils unabhängig voneinander aufgenommen werden. Die Verknüpfung geschieht im nächsten Schritt bei der Auswertung. Alternativ kann der vermutete formelmäßige Zusammenhang jedoch auch in die Messung integriert werden. Bei der Bestimmung der Dichte eines

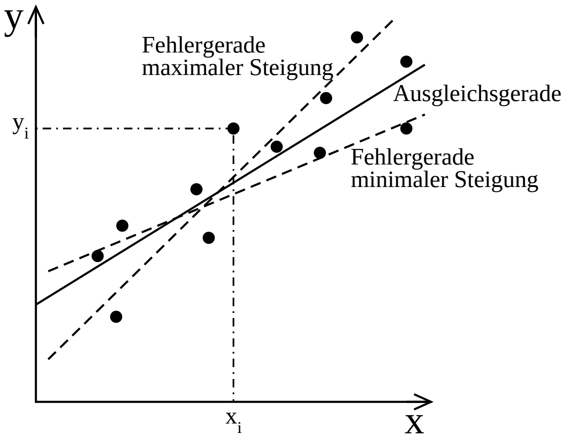

Abbildung 4:
Konstruktion einer Ausgleichsgeraden (Erläuterung im Text).

Körpers bedeutet dies z.B., dass man nicht nur bei einem Volumen misst, sondern die Massen von verschiedenen Volumina bestimmt und jeweils den resultierenden Quotienten vergleicht. Der Vorteil dieser Methode gegenüber einer punktuellen Messung liegt in einer Überprüfung des postulierten formelmäßigen Zusammenhangs. Bestimmte Typen von Fehlern äußern sich bei dieser Methode auch in spezifischen Abweichungen, wodurch Aussagen über die Ursachen möglich werden. Ein weiterer Vorteil der Methode ist die Möglichkeit der grafischen Darstellung der Gesetzmäßigkeit die erheblich anschaulicher ist als reine Zahlen.

10.1 Graphische Auswertung, Ausgleichsgerade

Die einfachste Beziehung zwischen zwei Messgrößen ist die lineare Abhängigkeit

$$y(x) = a_0 + a_1 \cdot x \tag{40}$$

mit den beiden freien Parametern a_0 und a_1.

Nimmt man Messwerte auf, von denen man eine solchen Gesetzmäßigkeit erwartet, trägt man sie graphisch auf. Die Position der Messwertepaare kennzeichnet man mit einem Symbol. Es erhöht die Übersichtlichkeit wenn man unterschiedliche Symbole für unterschiedliche Messreihen verwendet. Abbildung 4 zeigt eine solche Messung. Man erkennt, dass die Messwerte voneinander abhängig sind, sie also keine unstrukturierte Punktewolke bilden.

Vermutet man einen linearen Zusammenhang, so passt man zunächst grafisch eine *Ausgleichsgerade* an die Messwerte an, in Abbildung 4 ist dies die durchgezogene Linie. Als grober Anhaltspunkt für ihre Lage sollte gelten, dass im Mittel auf beiden Seiten der Geraden eine ähnliche Zahl von Punkten vorliegt und diese Punkte natürlich möglichst nahe an der Geraden liegen. Die Steigung und der Schnittpunkt der Geraden mit der vertikalen Achse sind dann die Parameter $a_1$und a_0.

Liegen für die Messwerte-Paare Informationen über die Genauigkeit vor, so zeichnet man diese als ‚Fehlerbalken' ein, der den Konfidenzbereich des Messwertes in der zugeordneten ‚Richtung' kennzeichnet. Diese Fehlerbalken sind ein wesentliches

Hilfsmittel beim Einzeichnen der Ausgleichsgeraden. Für ein hohes Konfidenzniveau, z.B. die 3-fache Standardabweichung, muss gelten:

Die Ausgleichs-Gerade schneidet alle Fehlerbalken mit Standardabweichung 3σ*.*

Gelingt es nicht eine solche Gerade zu finden, so kann dies heißen, dass

- kein linearer Zusammenhang vorliegt.
  Das Auftragen der Position bei einer gleichförmig beschleunigten Bewegung gegen die Zeit führt z.B. zu einer Parabel. Entsprechen darf es bei ausreichend hoher Messgenauigkeit nicht gelingen an die Messwerte eine Gerade anzupassen.
- Ausreißer, also grobe Fehler, vorliegen.
  Dies erkennt man daran, dass einzelne Messwerte erheblich von der Geraden abweichen, und alle anderen Messwerte gut mit einer Geraden übereinstimmen. Ausreißer belässt man zwar in der Graphik, kennzeichnet sie aber und berücksichtigt sie nicht bei der Bestimmung der Geraden.
- nicht erkannte systematische Abweichungen einfließen.
- die Standardabweichung falsch ermittelt wurde.
  Evtl. liegt keine Normalverteilung der statistischen Abweichungen vor, oder Sie haben sich einfach verrechnet.

Eine grobe Abschätzung für die Unsicherheiten der Steigung a_1 und des Achsenabschnitts a_0 erhält man, durch zwei begrenzende (Fehler-)Geraden mit minimaler und maximaler Steigung, die jeweils etwa zwei Drittel der Punkteschar umfassen. Diese sind in Abbildung 4 als gestrichelte Linien eingezeichnet.

10.2 Lineare Regression

Neben der graphischen Auswertung, die unabdingbar ist um Ausreißer zu erkennen, ist eine rechnerische Auswertung nötig.

Zur Ermittlung der *Regressionsgeraden* unterscheidet man zunächst zwischen einer *abhängigen* und einer *unabhängigen* Variablen. Die unabhängige Variable (z.B. x) wird als genau angenommen, die abhängige (y) als mit Abweichungen behaftet. Man ermittelt das Quadrat des Abstand jedes Messpunkts zu einer Geraden in Richtung der abhängigen Variable (also $(y_i - y_{\text{Gerade}}(x_i))^2$) und summiert diese auf. Die Regressionsgerade ist diejenige, bei der diese Summe minimal wird. Das Prinzip heißt *Methode der kleinsten Quadrate.*

Für einen Datensatz gibt es immer zwei Regressionsgeraden, je nachdem welche Variable man als unabhängig betrachtet. Diese sind nur dann gleich, wenn alle Punkte exakt auf einer Geraden liegen.

Zur Berechnung der Regressionsgeraden (abhängige Variable y) bestimmt man zuerst die Koeffizientendeterminante

$$D = n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 \tag{41}$$

Achsenabschnitt a_0 und Steigung a_1 erhält man dann als

$$a_0 = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\, x_i}{D} \qquad \text{und} \qquad a_1 = \frac{n \sum_{i=1}^{n} x_i\, y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{D} \tag{42}$$

Die Standardabweichung ist auch hier über die quadratische Summe der Abweichungen der Messwerte y_i zum Schätzwert definiert, der sich hier aber aus dem y-Wert der Geraden am der Stelle x_i ergibt

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^{n} \left(y_i - (a_0 + a_1\, x_i)\right)^2}{n-2}} \tag{43}$$

Da durch die verwendete Gerade ein Freiheitsgrad eingeschränkt wird, wird der Normierungsfaktor zu $1/(n-2)$. Die Unsicherheit der beiden Parameter berechnet sich zu

$$\Delta a_0 = \sigma_y \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{D}} \qquad \text{und} \qquad \Delta a_1 = \sigma_y \sqrt{\frac{n}{D}} \tag{44}$$

10.3 Korrelationskoeffizient

Hat man zu Beginn der Auswertung eine Punktwolke vorliegen, so ist die wesentliche Fragestellung:

Liegt ein linearer Zusammenhang vor?

Neben der visuellen Betrachtung gibt darüber der *Korrelationskoeffizient* (bedingt) Auskunft. Er ist das geometrische Mittel der Steigung der Regressionsgeraden zur abhängigen Variablen y und zur abhängigen Variablen x.

$$r^2 = a_1 \cdot \tilde{a}_1 \tag{45}$$

und kann Werte zwischen 0 und 1 annehmen.

Es bedeutet:

$r^2 = 0,0$	gar keine Korrelation, perfekte Unordnung
$r^2 = 0,3$	sehr geringe Korrelation
$r^2 = 0,7$	deutliche Korrelation, aber starke Schwankungen
$r^2 = 0,8$	gute Korrelation, sofern *nur statistische* Abweichungen
$r^2 = 0,9$	schlechte Korrelation, falls *keine statistische* Abweichungen
$r^2 = 1,0$	perfekte Korrelation, alle Punkte auf einer Geraden

Ein Korrelationskoeffizient von $r^2 = 0,8$ kann der Beweis eines linearen Zusammenhangs sein, wenn die Abweichungen von der Gerade rein statistischer Natur sind, aber auch der Beweis einer Nichtlinearität, wenn er durch nicht statistische Abweichungen zustande kommt.

Ohne visuelle Bewertung des Datensatzes ist der Korrelationskoeffizient nutzlos!

10.4 Lineare Regression mit Wichtung

Im Abschnitt 10.2 wurde angenommen, dass jeder Messwert die gleiche Genauigkeit aufweist. Daher war es sinnvoll, die Summe der quadratischen Abweichungen aller Messpunkte zu nehmen. Sind aber die Punkte auf dem Datenfeld unterschiedlich genau, so ist es sinnvoll, die genaueren Punkte bei der Berechnung stärker zu berücksichtigen.

Wie beim gewichteten Mittelwert wird dem mit einem Wichtungsfaktor w_i Rechnung getragen

$$w_i = \frac{1}{\Delta {y_i}^2} \tag{46}$$

Man erhält dann die Koeffizientendeterminante

$$D = \sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i\, x_i^2 - \left(\sum_{i=1}^{n} w_i\, x_i \right)^2 \tag{47}$$

den Achsenabschnitt

$$a_0 = \frac{\sum\limits_{i=1}^{n} w_i\, x_i^2 \sum\limits_{i=1}^{n} w_i\, y_i - \sum\limits_{i=1}^{n} w_i\, x_i \sum\limits_{i=1}^{n} w_i\, y_i\, x_i}{D} \tag{48}$$

und die Steigung

$$a_1 = \frac{\sum\limits_{i=1}^{n} w_i \sum\limits_{i=1}^{n} w_i\, x_i\, y_i - \sum\limits_{i=1}^{n} w_i\, x_i \sum\limits_{i=1}^{n} w_i\, y_i}{D} \tag{49}$$

Für eine *konstante absolute Unsicherheit* sind alle Wichtungsfaktoren gleich, und man erhält die ‚normale' Regression.

Für eine *konstante relative Unsicherheit* werden die Wichtungsfaktoren $w_i = 1/{y_i}^2$.

Hier sei noch darauf hingewiesen, dass es zur Vermeidung unnötiger Rechenarbeit sinnvoll ist, den systematischen Fehler aus der Berechnung der Geraden herauszulassen, und als Faktor mittels Fortpflanzung nach der Berechnung der Geraden zu berücksichtigen.

10.5 Nichtlineare Regression

Natürlich sind nicht alle Zusammenhänge in der Natur linear. Häufig vorkommende Relationen sind

exponentiell: z.B.: Laden eines Kondensators, radioaktiver Zerfall

quadratisch: z.B. Beschleunigungsvorgänge

Die meisten nichtlinearen Zusammenhänge lassen sich auf eine lineare Regression zurückführen, indem man eine geeignete Variablentransformation vornimmt. Da die Unsicherheiten dabei genauso transformiert werden wie die Größen selbst, muss man dann (fast) immer eine Gewichtete Regression anwenden.

Betrachten wie hier das Beispiel eines exponentiellen Abfalls:

$$y = y_0 \, e^{-\lambda \, t} \tag{50}$$

Durch Logarithmieren ergibt sich eine Gerade für $\ln y$

$$\ln y = \ln y_0 - \lambda \, t \tag{51}$$

Nimmt man nun eine konstante absolute Unsicherheit $\Delta y = \text{const.}$ *für die ursprünglichen Messwerte an, so erhält man nach dem Logarithmieren* $\Delta \ln y \propto 1/y$, *und entsprechend für die Wichtungsfaktoren* $w_i = y_i^2$.