Einführung in die Fehlerrechnung

Einleitung

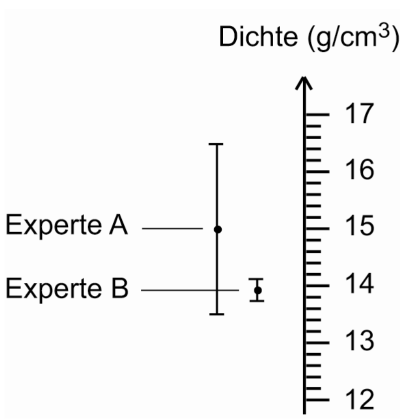
Die Fehlerrechnung hat den Zweck, Unsicherheiten in Messungen zu untersuchen und zu evaluieren. Die Erfahrung zeigt, dass jede, auch die am sorgfältigsten ausgeführte Messung, eine gewisse Unsicherheit beinhaltet. Da die ganze Struktur und alle Anwendung von Wissenschaft auf Messungen beruht, ist die Fähigkeit, Fehler abzuschätzen und sie minimal zu halten von fundamentaler Bedeutung.

Die Unvermeidbarkeit von Unsicherheiten

In den Naturwissenschaften hat das Wort "Fehler" nicht die übliche Bedeutung von "Irrtum", sondern gemeint ist die unvermeidbare Unsicherheit, die allen Messungen zu eigen ist (Sonderfall systematische Fehler siehe unten). Diese Unsicherheit kann auch nicht dadurch eliminiert werden, dass eine Messung besonders sorgfältig ausgeführt wird. Das einzige was man tun kann, ist Massnahmen zu treffen, die dafür sorgen, dass diese Unsicherheit so klein als (sinnvollerweise) möglich ist. *Erläuterung: Messung der Tischhöhe im Hörsaal*. Dieses Beispiel veranschaulicht, dass keine physikalische Grösse (Länge, Temperatur, Druck, Zeit etc.) mit absoluter Genauigkeit gemessen werden kann.

Warum muss man die Grösse der Unsicherheit kennen?

Nehmen wir an, wir geben die Messung der Dichte eines Goldringes in Auftrag, um herauszufinden ob es sich dabei um 18 Karat Gold mit einer bekannten Dichte von 15.5 g/cm^3 oder um eine billigere Legierung mit einer bekannten Dichte von 13.8 g/cm^3 handelt. Zwei Experten, die mit unterschiedlichen Methoden messen, liefern uns folgende Ergebnisse: Der Experte A misst eine Dichte von 15, wobei die Unsicherheit 1.5 g/cm^3 beträgt und Experte B misst eine Dichte von 13.9, wobei die Unsicherheit 0.2 g/cm^3 beträgt.



Der entscheidende Punkt in diesem Beispiel ist die Erkenntnis, dass die Ergebnisse der Experten ohne Angaben der Unsicherheit absolut wertlos wären! Ohne die Angaben der Unsicherheiten wären wir nämlich durch den Experten A fehlgeleitet worden in der Annahme, dass es sich um 18 Karat Gold handelt, was definitiv falsch ist, wie die Messung des Experten B beweist. Beachte übrigens, dass dies nicht heisst, dass der Experte A falsch gemessen hat! Aus ähnlichen Überlegungen ist die Angabe der Unsicherheit einer Messung zwingend notwendig, wenn es darum geht, experimentell eine neue Theorie zu verifizieren (vgl. Allgemeine Relativitätstheorie von Einstein und ihre Verifizierung durch Dyson, Eddington und Davidson).

Abschätzung der Unsicherheit

Die Evaluierung der Unsicherheit einer gemessenen Grösse kann ziemlich kompliziert sein. Zum Glück gibt es aber viele Beispiele, in denen diese Abschätzung unter Anwendung gesunden Menschenverstandes einfach möglich ist. Ganz einfach ist es beim Ablesen an einer Skala (z.B.: Länge, Zeit, Spannung, Temperatur etc.) und zwar ganz unabhängig davon, ob es sich um eine Analoge oder Digitale Anzeige handelt. *Beispiele zur Erläuterung: Länge eines Bleistiftes, Spannung ablesen an einem Voltmeter, Bestimmung der Masse mit einer Waage.* Es gibt aber auch Fälle, in denen der Messfehler viel schwieriger zu evaluieren ist. Zum Beispiel bei der Messung eines Zeitintervalls mit einer Stoppuhr. Die Hauptschwierigkeit besteht hier nicht im Ablesen der Skala, sondern in der variablen menschlichen Reaktionszeit. In solchen Fällen lässt sich die Unsicherheit aber meist einfach abschätzen, indem man die Messung mehrfach wiederholt und aufgrund der **Streuung** der Einzelergebnisse die Unsicherheit abschätzt (detaillierte Behandlung siehe unten).

Systematische Fehler

Es kann auch vorkommen (aber sollte nach Möglichkeit stets vermieden werden), dass bei Messungen ein systematischer Fehler auftritt. Es handelt sich dabei um einen wirklichen Fehler im Sinne eines "Irrtum", **der sich auch nicht durch wiederholtes Messen evaluieren (oder gar verkleinern) lässt**. Dies ist dann der Fall, wenn zum Beispiel die Uhr, die für eine Messung verwendet wurde, 5% zu schnell läuft oder ein Messgerät falsch kalibriert ist. Solche Fehler sind oft schwierig zu erkennen. Man kann solche Fehler aber ausschliessen, wenn man mit einem 2. Messgerät eine Kontrollmessung durchführt.

Begriffe

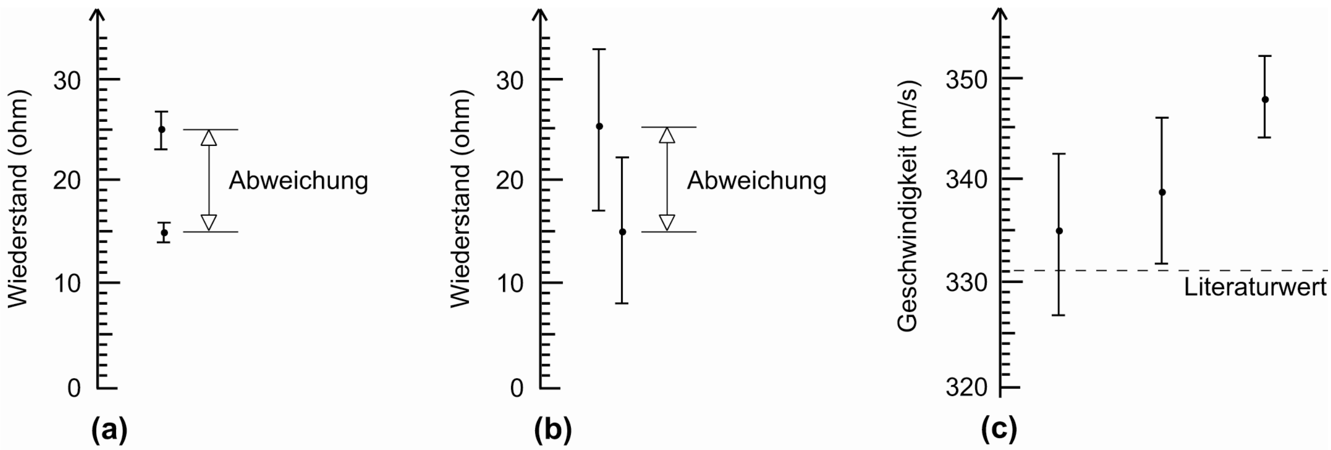
In diesem Kapitel geht es darum, die grundlegenden Begriffe einzuführen, die in der Fehlerrechnung von Bedeutung sind. Es ist üblich, das **Ergebnis einer Messung** wie folgt anzugeben:

Gemessener Wert $\quad x = x_{\text{best}} \pm \Delta x$

Wobei x_{best} als **"best estimate"** und Δx als der sogenannte **"Fehler"** (Unsicherheit, Fehlerbereich) bezeichnet werden. Wie man diese Grössen ermittelt, ist das eigentliche Thema der Fehlerrechnung und wird in den folgenden Kapiteln beschrieben. Beachte, dass die Anzahl der **signifikanten Stellen** im best estimate und im Fehler die gleiche sein sollte, also z. B.:
$\mathbf{g = 9.82 \pm 0.02\ m/s^2}$ und nicht $g = 9.82 \pm 0.022765\ m/s^2$ oder $g = 9.821754 \pm 0.02 m/s^2$.

Vergleich von Messwerten

Ein weiterer wichtiger Begriff ist die **Abweichung**. Die Abweichung bezeichnet die Differenz zwischen zwei gemessenen Werten der gleichen Grösse. Wenn also z. B. ein Student A für einen Widerstand 15 ± 1 Ohm erhält und Student B einen Wert von 25 ± 2 Ohm, so beträgt die Abweichung der beiden Messungen $25 - 15 = 10$ Ohm. Die Veranschaulichung in der Figur (a) zeigt, dass diese Abweichung **signifikant** ist, da es keinen Wert für den Wiederstand gibt, der mit beiden Messungen kompatibel ist. Daher muss man davon ausgehen, dass mindestens eine Messung nicht korrekt ist. In solchen Situationen sollte man überprüfen, was falsch gelaufen ist. Im zweiten Teil der Figur (b) ist ersichtlich, dass zwei Messwerte mit einer Abweichung von 10 Ohm durchaus auch **miteinander verträglich** sein können (**übereinstimmen**), wenn die zwei Fehlerbereiche überlappen.



Ganz ähnlich ist die Situation, wenn ein gemessener Wert mit einem **Literaturwert** verglichen wird. Es macht keinen Sinn ein Experiment auszuführen, ohne einen bestimmten Zweck zu verfolgen und vernünftige Schlussfolgerungen zu ziehen. Literaturwerte sind typischerweise Ergebnisse von sehr aufwändigen Experimenten, die zur Bestimmung einer Naturkonstante oder anderer wichtiger Grössen mit grösst möglicher Genauigkeit ausgeführt wurden. Der Literaturwert für die Gaskonstante beträgt zum Beispiel 8.31451 $\pm$ 0.00007 $J/(mol \cdot K)$. Liegt nun dieser Literaturwert innerhalb des Fehlerbereichs der eigenen Messung, sind diese Werte **miteinander verträglich**, was heisst, dass man korrekt gemessen hat (siehe Figur c). Liegt der Literaturwert deutlich ausserhalb des Fehlerbereichs, so besteht eine **signifikante Abweichung** und das ermittelte Resultat ist somit fehlerbehaftet (z. B. mit einem systematischen Fehler).

Statistische Analyse zufällig gestreuter Messwerte

Mittelwert und Standardabweichung

Nehmen wir an, wir wiederholen die Messung einer Grösse x 5 mal und erhalten folgende Messergebnisse $x_1, ..., x_5$:

71, 72, 72, 73, 71

Es ist intuitiv klar (und es lässt sich mit statistischen Mitteln beweisen), dass der "best estimate" dieser Messung dem sogenannten **Mittelwert** $\bar{x}$ entspricht:

$$x_{\text{best}} = \bar{x} = \frac{71 + 72 + 72 + 73 + 71}{5} = 71.8$$

Oder allgemein formuliert:

$$\bar{x} = \frac{\sum_i x_i}{N}$$

Der nächste wichtige Begriff ist die sog. **Standardabweichung** σ_x. Die Standardabweichung der Messungen $x_1, ..., x_N$ ist eine Abschätzung der durchschnittlichen Unsicherheit der einzelnen Messungen und wird wie folgt bestimmt:

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum_i (x_i - \bar{x})^2}$$

Ausformuliert bedeutet dies, dass zur Ermittlung der Standardabweichung die Abweichungen aller Einzelmessungen vom Mittelwert ermittelt werden, dann werden diese quadriert und aufsummiert. Nachdem dieser Wert gemittelt wird, zieht man die Wurzel und erhält die Standardabweichung. Somit bleibt noch die Unsicherheit des Mittelwertes abzuschätzen. Es lässt sich zeigen, dass die Unsicherheit der Zielgrösse (also des Mittelwertes) erhalten wird indem man die Standardabweichung σ_x durch $\sqrt{N}$ dividiert. Die Resultierende Grösse nennt man **Standardabweichung des Mittelwertes**:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}$$

Somit lautet die korrekte Angabe einer gemessenen Grösse:

$x = x_{\text{best}} \pm \Delta x \quad \text{wobei} \quad \Delta x = \sigma_{\bar{x}}$

Fehlerfortpflanzung

Die meisten physikalischen Grössen lassen sich nicht durch eine einfache, direkte Messung ermitteln. Häufig erfolgt die Bestimmung der Grösse in zwei ganz unterschiedlichen Schritten. Als erstes werden mehrere Grössen gemessen, aus welchen dann im zweiten Schritt die eigentliche Zielgrösse berechnet wird (in anderen Worten: die Zielgrösse G ist eine Funktion von $x, y, ..., z$). In diesem Fall erfordert die Abschätzung des Fehlers ebenfalls zwei Schritte: **Die Abschätzung der Unsicherheit aller ermittelten Grössen sowie die Untersuchung, wie sich diese Unsicherheiten bei der Berechnung der Zielgrösse "fortpflanzen"**. Daraus ergibt sich die Unsicherheit des Endergebnisses.
Zur Bestimmung des Fehlers einer Grösse G, die aus den voneinander unabhängigen Grössen $x, ..., z$ bestimmt wurde, sind die folgenden Fälle zu unterscheiden.

1. Einzelmessungen: Wurden die Grössen $x, ..., z$ durch eine Einzelmessung bestimmt, ist deren Unsicherheit von systematischer Natur und daher durch die Ableseungenauigkeit bestimmt. Die Auswirkung dieser Unsicherheiten auf die Grösse G lässt sich dann durch eine Reihenentwicklung erster Ordnung abschätzen:

$$|\Delta G_{sys}| = \left| \frac{\partial G}{\partial x} \cdot \Delta x_{sys} \right| + ... + \left| \frac{\partial G}{\partial z} \cdot \Delta z_{sys} \right|$$

2. Mehrfachmessungen: Bestimmt man die Grössen $x, ..., z$ durch Mehrfachmessungen, lässt sich deren statistischer Fehler durch die im vorherigen Abschnitt besprochene Standardabweichung bestimmen. Zur Berechnung des Fehlers der Grösse G bedient man sich dann der sogenannten **Gauss'schen Fehlerfortpflanzungsformel**:

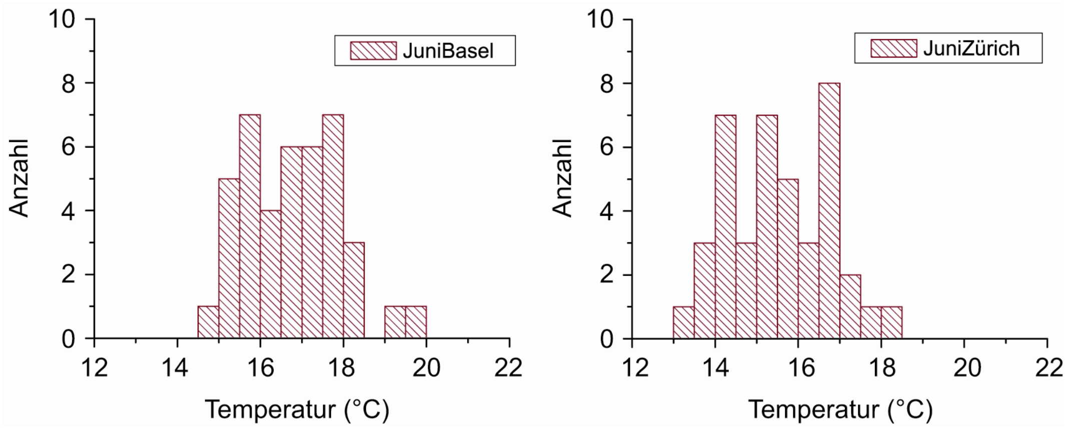
$$|\Delta G_{stat}| = \sqrt{\left(\frac{\partial G}{\partial x} \cdot \Delta x_{stat} \right)^2 + ... + \left(\frac{\partial G}{\partial z} \cdot \Delta z_{stat} \right)^2}$$

   Mit Hilfe dieser Formel lassen sich die Fehler auch dann ermitteln, wenn die Berechnung der Zielgrösse beliebig aufwändig ist. Für einfache Fälle wie simple Produkte, Quotienten oder Summen und Differenzen ist die Lösung jeweils sehr einfach.

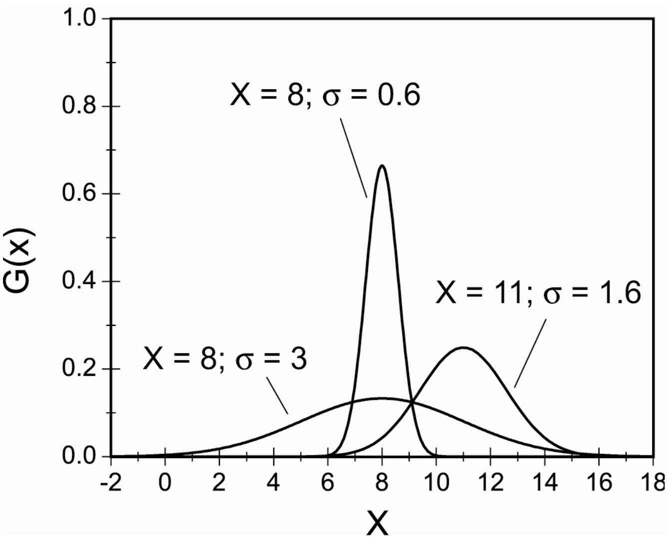
Die Normalverteilung

Histogramm

Es ist wichtig festzuhalten, dass eine seriöse statistische Analyse eines Experiments viele Einzelmessungen voraussetzt (mehr als 100!!). Aus diesem Grund benötigt man vernünftige Methoden zur Veranschaulichung einer grossen Anzahl gemessener Werte. Üblich ist die Verwendung von **Histogrammen**. Für die Erstellung eines Histogrammes werden die Messdaten sogenannten "Töpfen" (englisch: bins) mit einer definierten Breite zugeordnet. Anschliessend wird in einem Balkendiagramm aufgezeichnet, wie viele Einzelmessungen pro Topf gezählt wurden. Als Beispiel dienen hier die durchschnittlichen Juni Temperaturen gemessen in den letzten 40 Jahren in Basel und Zürich:



In allen Experimenten, in denen die Streuung der Einzelmessungen wirklich zufällig ist, beginnt das Histogramm mit zunehmender Anzahl Messungen eine bestimmte Form anzunehmen. Im Grenzfall von unendlich vielen Messungen und einer Topfbreite die gegen Null strebt, nimmt das Histogramm die Form einer symmetrischen Glockenkurve an. Diese Verteilungsfunktion wird als **Normalverteilung** oder **Gaussverteilung** bezeichnet und sieht wie folgt aus:



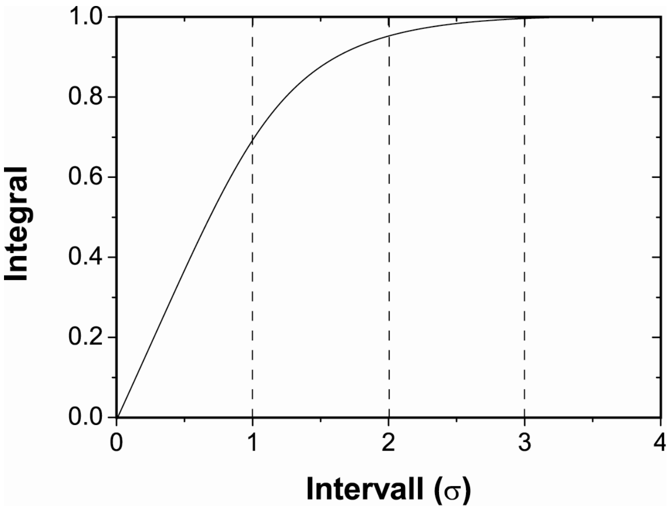
und die zugehörige Formel lautet:

$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-X)^2}{2\sigma^2}}$$

Dabei entspricht X dem Erwartungswert (Mittelwert) und σ der Standardabweichung. Zudem ist die Funktion normiert, sodass der Flächeninhalt unterhalb der Kurve (das Integral) immer 1 entspricht.

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-X)^2}{2\sigma^2}} dx = 1$$

Die Breite der Gaussfunktion gibt also Auskunft darüber, wie genau eine Messung ist. Zudem lässt sich mit Hilfe der Gaussverteilung ermitteln, wie gross die Wahrscheinlichkeit ist, in einer einzelnen Messung einen Wert innerhalb eines spezifischen Intervalls zu erhalten. Besonders interessant ist dabei die Frage, wie gross die Wahrscheinlichkeit ist, ein Ergebnis im Bereich $X \pm \sigma$ zu erzielen. Dies kann ganz einfach durch Integration über das erwünschte Intervall erreicht werden.



Aus dieser Figur ist ersichtlich, dass eine einzelne Messung mit einer Wahrscheinlichkeit von 68% innerhalb von $\pm\sigma$ fällt. Die Wahrscheinlichkeit, dass eine Messung innerhalb $\pm 2\sigma$ oder $\pm 3\sigma$ liegt beträgt 95.4% respektive 99.7%. Anders ausgedrückt: liegt z.B. der Literaturwert ausserhalb des $\pm 3\sigma$ Bereiches einer Messreihe, sind Literaturwert und Messergebnis nicht miteinander verträglich. Die beobachtete Abweichung kann dann nicht durch die statistische Streuung erklärt werden und es muss ein systematischer Fehler vorliegen.

Schlussbemerkungen

Eine ausführliche Besprechung aller aufgeführten Themen sowie zusätzliche Informationen zu den folgenden Themen:

- Gewichtete Mittelwerte
- Fitten von Daten
- Kovarianz und Korrelation
- Binomial Verteilung
- Poisson Verteilung
- etc.

finden sich im Buch **An Introduction to Error Analysis, J.R. Taylor, University Science Books, 2$^{\text{nd}}$ edition 1997.** Dieses Buch liegt im Praktikum auf.