

## Einführung

Bei der Messung physikalischer Größen **treten immer Fehler auf**. Eine genaue, absolut präzise Messung ist nur beim Abzählen von Datenwerten möglich (und da auch nicht immer).

**Eine physikalische Messung ohne Genauigkeitsangabe ist wertlos.**

Grundsätzlich müssen in jedem Praktikumsbericht neben den Resultaten auch Angaben zu den Fehlern vorhanden sein.

Man unterscheidet drei Fehlerarten: grobe Fehler, systematische Fehler und statistische Fehler. Grobe Fehler entstehen, wenn die Experimentatoren ihr Handwerk nicht verstehen (Experimentelle Physik ist in erster Linie Handwerk). Grobe Fehler lassen sich durch Literaturstudium, Kontrollmessungen und durch Abschätzungen vor dem Experiment über die zu erwartenden Größen vermeiden. Systematische Fehler treten bei jeder Durchführung eines Versuches in der gleichen Art und Weise auf. Statistische Fehler sind zufällig. Ihr Auftreten liegt in der Natur des Messprozesses und kann nur durch Wiederholung der Messungen minimiert werden.

Dieser kleine Text will Ihnen eine Zusammenfassung der für die Fehlerrechnung und Fehlerbehandlung benötigten Methoden geben.

### **Systematische Fehler**

Systematische Fehler sind durch die Unvollkommenheit der Messgeräte sowie durch nicht steuerbare Äußere Einflüsse bestimmt. Systematische Fehler können in vielen Fällen durch ergänzende Messungen korrigiert werden (zum Beispiel die Totzeitkorrektur bei Zählrohren). Systematische Fehler können durch eine eingehende Analyse der verwendeten Messgeräte und der Messverfahren erkannt werden. Typische systematische Fehler sind:

### Umwelteinflüsse

- Bei der ultrapräzisen Gewichtsbestimmung von Substanzen muss der Auftrieb der Luft berücksichtigt werden. Dieser Fehler kann, wenn alle relevanten Größen aufgezeichnet wurden, sehr leicht rechnerisch kompensiert werden.
- Dazu muss die Umgebungstemperatur im Laborbuch vermerkt werden.
- Die Luftfeuchtigkeit beeinflusst ebenfalls die Dichte der Luft (Warum?<sup>1</sup>)
- Viele elektrische Messgeräte haben eine von der Temperatur des Messwerkes abhängige Empfindlichkeit.
- Bei der Druckmessung mit Quecksilberbarometern muss die Kapillardepression berücksichtigt werden.
- Bei der Druckmessung mit kapazitiven Drucksensoren beeinflusst die Temperatur die Empfindlichkeit und die Nullpunktlage.
- Adsorbierte Gasschichten auf Metallspiegeln verfälschen den Reflexionskoeffizienten.
- Adsorbierte Wasserschichten beeinflussen in dramatischer Art und Weise die gemessenen Kräfte in der Rasterkraftmikroskopie.
- Wenn die Temperatur während eines Versuchs nicht konstant gehalten wird, können die Messwerte vom Anfang und vom Ende der Messzeit nicht verglichen werden. Ursachen für einen Temperaturgang können zum Beispiel Wetterumschläge oder die Körpertemperatur der Experimentatoren sein. Also: immer den Temperaturgang aufzeichnen bei Präzisionsmessungen.

<sup>1</sup> Die mittlere Dichte der Luft hängt von den enthaltenen Molekülen ab. Luft besteht aus ~80% N<sub>2</sub> (Molekülgewicht 28) und ~20% O<sub>2</sub> (Molekülgewicht 32). Wasser in der Luft ("Feuchte") besteht aus Molekülen mit dem Atomgewicht 18. Deshalb ist feuchte Luft leichter als trockene.

- Bei gewissen Experimenten muss auf den Mondstand (Gravitation) geachtet werden. Wenn die Umwelteinflüsse nicht überwacht werden, können meistens nur qualitative Ergebnisse erwartet werden.

Eine weitere Kategorie systematischer Fehler ist die Rückwirkung des Messgerätes auf das gemessene System.

- Bestimmung der Spannung oder des Stromes mit den üblichen Messgeräten (⇒ Grundpraktikumsversuch)
- Veränderung des Widerstandes durch den Messstrom.
- Temperaturmessung: die Wärmekapazität der Messsonde entzieht der Probe Wärme und verfälscht so das Messresultat.
- Falsche Impedanzen oder ungeeignete Abschlusswiderstände können elektrische Messungen zum Glücksspiel werden lassen.
- Bei Rastertunnelmikroskopen ist die Sonde (Tunnelspitze) prinzipbedingt eng an die Probe gekoppelt.

Eine weitere Quelle systematischer Fehler sind unvollkommene Messgeräte.

- Viele Geräte haben nichtlineare Kennlinien, die die Messresultate verfälschen.
- Eichfehler und Alterungserscheinungen gehören zu den systematischen Fehlern.
- Induzierte Ladungen, zum Beispiel durch die Bewegung des Experimentators in trockener Luft, können höchstempfindliche Messungen von Strom oder Spannung stören.
- Der Eigenverbrauch von Messgeräten zählt zu den systematischen Fehlern.
- Lageabhängigkeit der Messung
- Äußere Magnetfelder
- Reibung

Messungen sind sehr viel präziser, wenn nicht ein Wert gemessen wird, sondern der Messwert mit einer Referenz (eventuell abgleichbar) verglichen wird. Hat man eine gute Theorie des Messgerätes und des Messprozesses zur Hand, können diese systematischen Fehler herausgerechnet werden.

Eine weitere, nicht zu unterschätzende Quelle systematischer Fehler sind die Experimentatoren. Dazu gehört insbesondere "Bias", mangelnde Objektivität. Die meisten falschen Messresultate sind dadurch zustande gekommen, dass der Experimentator das Resultat aus unzureichenden Daten herausgelesen hat, das er haben wollte<sup>2</sup>.

Systematische Fehler können entdeckt werden, indem

- Theorien nach ungewollten Effekten untersucht werden (Mehr als nur die erste Näherung betrachten).
- Die Messbedingungen verändert werden. (andere Messgeräte, andere Messanordnung)
- Eine Größe mit zwei grundsätzlich verschiedenen Messverfahren bestimmt wird.
- Messverfahren die auf der Kompensationsmethode beruhen verwendet werden.
- Messverfahren mit minimalen systematischen Fehlern verwenden.

---

<sup>2</sup> Der Wunsch ist der Vater des Gedankens.

## Vor einem Experiment müssen die folgenden Fragen geklärt werden:

- Können im Experiment nicht erfasste Umwelteinflüsse die Messwerte beeinträchtigen?
- Entsprechen die abgelesenen Messwerte wirklich den zu messenden physikalischen Größen?
- Sind die zur Auswertung verwendeten Gleichungen richtig oder wenigstens innerhalb der erstrebten Genauigkeit genügend gut (Theorie des Messvorganges)
- Wo besteht die Gefahr, grobe Fehler oder subjektive Beobachtungen zu machen?
- Mit einer Überschlagsrechnung sind die Einheiten, die Größenordnungen und die Vorzeichen zu bestimmen.
- Sind die verwendeten Messinstrumente in ihrer Eichung (noch) zuverlässig?

### **Statistische Fehler**

Die Forderung nach Reproduzierbarkeit ist nie streng erfüllt. Auch wenn systematische Fehler ausgeschlossen werden können oder wenn sie bekannt sind, verbleiben zufällige, statistische Fehler. Diese haben ihre Ursache zum Beispiel in

- Physik des Phänomens (Radioaktivität)
- Rauschen des Messsignals
- Unkontrollierbare kleine Schwankungen des gemessenen Signals.

Die statistischen Fehler unterliegen, wie ihr Name sagt, den Gesetzen der Stochastik. Im folgenden wollen wir diese Fehler näher untersuchen.

### **Statistische Fehler**

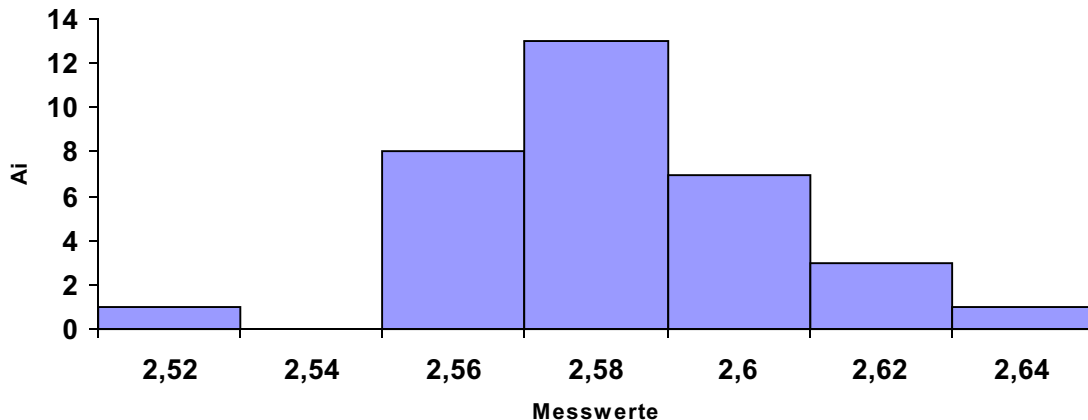
Wenn eine Messung nur einmal durchgeführt wird, so kann man über die Zuverlässigkeit keine Angaben machen. Es ist deshalb unumgänglich, Messungen zu wiederholen.

### **Stichproben**

Die erhaltenen Messresultate sind eine Stichprobe aus der Menge aller möglichen Messresultate. Jede Stichprobe hat eine gewisse, noch zu berechnende Wahrscheinlichkeit, dass sie die Grundgesamtheit repräsentiert. Stichproben werden üblicherweise als Histogramme dargestellt. Die gemessenen Daten  $x$  werden in Klassen mit

$$(x_i - \Delta x_i) < x < x_i$$

eingeteilt.



**Abbildung 1 Histogramm von Messwerten.**

### Mittelwerte

Aus diesem Histogramm kann man den Mittelwert

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

berechnen. Hier sind die  $A_i$  die Häufigkeiten der in eine bestimmte Klasse fallenden Messwerte. In unserem Falle ist dies 2,583030303. Man kann diese Gleichung umschreiben, indem man die relativen Häufigkeiten

$$a_i = \frac{A_i}{A}$$

verwendet. Der Mittelwert ist dann

$$\bar{x} = \sum_{i=1}^n a_i x_i$$

Der so gefundene Mittelwert ist im allgemeinen nicht mit dem wahren Wert der Messgröße identisch.

Wenn wir die Klassenbreite verringern und gleichzeitig die Anzahl Messpunkte erhöhen, so erhält man die differentielle Verteilungsfunktion  $f(x)$  der zugrunde liegenden Gesamtheit. Diese Funktion  $f(x)$  ist normiert, da wir die relativen Häufigkeiten zu ihrer Ableitung verwendet hatten.

$$1 = \int_{-\infty}^{+\infty} f(x) dx$$

Wir müssen unterscheiden zwischen den aus den experimentellen Daten empirisch gefundenen Verteilungsfunktionen und den durch theoretische Modelle berechneten Verteilungen. Eine wichtige Aufgabe einer Datenanalyse kann sein, zu zeigen, dass eine empirische Verteilungsfunktion mit einer theoretisch gefundenen verträglich ist.

### Varianzen und Standardabweichungen

Eine Verteilungsfunktion ist nicht nur durch ihren Mittelwert, sondern auch durch die Lage- und Dispersionsgrößen gegeben. Wir haben gesehen, dass der arithmetische Mittelwert (Mittelwertebenen) eine solche Größe ist. Die Lagegrößen müssen die folgenden Postulate erfüllen:

- Die Schätzung soll im Falle einer unendlich großen Stichprobe den Wert der Grundgesamtheit annehmen.
- Die beste Schätzung ist die mit der kleinsten Streuung, d.h. mit der kleinsten Fehlerabweichung.

Als Lagegrößen kommen in Frage:

- Der arithmetische Mittelwerte(Mittelwerteoben)  $\bar{x} = 2,583030303$ .

$$G(x) = \sqrt[A]{\prod_{i=1}^n x_i^{a_i}} = 2,58292652$$

- Das geometrische Mittel

$$\frac{1}{R(x)} = \frac{1}{A} \sum_{i=1}^N A_i \frac{1}{x_i} = 0,3871733; \quad R(x) = 2,58282275$$

- Das reziproke Mittel
- Der Median ist der Wert, bei dem gleich viele Werte links und rechts davon liegen. Hier ist  $M(x)=2,58$

Der Median ist besonders dann zu verwenden, wenn die Stichprobe eine große Streuung aufweist. Wenn aus anderen Daten bekannt ist, dass nicht alle Messwerte die gleiche Güte haben, kann man die  $A_i$  auch als Gewicht benutzen. Hier hatten wir die Anzahl Messwerte pro Klasse als Güte des Messwertes genommen.

Es ist einsichtig, dass man zusätzlich versucht, die Breite einer Verteilung zu charakterisieren. Diese Größen heißen Dispersionsgrößen. Man verwendet:

$$d = \frac{1}{A} \sum_{i=1}^n A_i |x_i - \bar{x}| = 0,01737374$$

- Den durchschnittlichen Fehler:

$$\sigma^2 = \frac{1}{A} \sum_{i=1}^n A_i (x_i - \bar{x})^2 = 0,00020761$$

- Die Varianz: . Wenn man die Varianz bezüglich eines andern Wertes B bildet, so gilt mit

$$\sigma^2(B) = \frac{1}{A} \sum_{i=1}^n A_i (x_i - B)^2$$

$$\sigma^2(0) = \frac{1}{A} \sum_{i=1}^n A_i (x_i)^2 = \overline{x^2}$$

$$\sigma^2(\bar{x}) = \frac{1}{A} \sum_{i=1}^n A_i (x_i - \bar{x})^2$$

$$\text{also } \sigma^2(B) = \sigma^2(\bar{x}) + (B - \bar{x})^2$$

Da dies alles positiv definite Größen sind, ist die Varianz minimal, wenn sie bezüglich des arithmetischen Mittelwertes  $\bar{x}$  berechnet wird.

Wenn wir eine kontinuierliche Verteilung haben und  $p(x)$  die Gewichtsfunktion ist, gilt:

$$\bar{x} = \frac{\int_{-\infty}^{\infty} x \cdot p(x) \cdot f(x) \cdot dx}{\int_{-\infty}^{\infty} p(x) \cdot f(x) \cdot dx}$$

- Arithmetischer Mittelwert:

$$\sigma^2 = \frac{\int_{-\infty}^{\infty} (x - \bar{x})^2 \cdot p(x) \cdot f(x) \cdot dx}{\int_{-\infty}^{\infty} p(x) \cdot f(x) \cdot dx}$$

- Varianz.

- allgemein gilt: Der mit der Verteilungs- und Gewichtsfunktion bewertete Mittelwert der

$$\overline{h(x)} = \frac{\int_{-\infty}^{\infty} h(x) \cdot p(x) \cdot f(x) \cdot dx}{\int_{-\infty}^{\infty} p(x) \cdot f(x) \cdot dx}$$

Funktion  $h(x)$  ist

Shepard gibt an, dass ein besserer Wert für die Varianz erhalten wird bei in Klassen eingeteilten Messgrößen, wenn man die folgende Formel verwendet.

$$\sigma^2 \cong \frac{1}{A} \sum_{i=1}^n A_i (x_i - \bar{x})^2 - \frac{h^2}{12}$$

wobei  $h$  die Klassenbreite ist. In unserem Falle wäre  $\sigma^2 \cong 0,00017428$ .

### Mittlerer Fehler und Varianz

Da der Mittelwert der Grundgesamtheit,  $\mu$ , im allgemeinen nicht bekannt ist, wird die berechnete Varianz nicht die Varianz der Grundgesamtheit sein. Wir versuchen nun den besten Schätzwert für die Varianz zu berechnen. Nehmen wir an, wir würden  $\mu$  kennen. Dann gilt

$$\hat{\sigma}^2 = \frac{1}{A} \sum_{i=1}^n A_i (x_i - \mu)^2 = \frac{1}{A} \sum_{i=1}^n A_i (x_i - m + m - \mu)^2$$

Im folgenden setzen wir alle  $A_i=1$  und  $\mu=0$ . Dann ist  $A=n$ . Durch ausmultiplizieren erhalten wir

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 + \frac{1}{n} \sum_{i=1}^n (m - \mu)^2 + 2 \frac{1}{n} \sum_{i=1}^n (x_i - m)(m - \mu) \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 + \frac{1}{n} \sum_{i=1}^n m^2 + 2 \frac{m}{n} \sum_{i=1}^n (x_i - m) \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 + m^2 \end{aligned}$$

mit  $m = \frac{1}{n} \sum_{i=1}^n x_i$  erhält man

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 + \frac{1}{n^2} \left( \sum_{i=1}^n x_i \right)^2$$

Mit

$$\frac{1}{n^2} \left( \sum_{i=1}^n x_i \right)^2 = \frac{1}{n^2} \sum_{i=1}^n x_i^2 + \frac{1}{n^2} \sum_{i \neq j} x_i x_j \cong \frac{1}{n} \overline{x^2} = \frac{\hat{\sigma}^2}{n}$$

wird

$$\sigma^2 \cong \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - m)^2}{n-1} \equiv s^2$$

Die Größe  $s$  ist der **mittlere Fehler einer Einzelmessung**. Der Übergang von  $n$  nach  $n-1$  ist zu verstehen als der Verlust eines Freiheitsgrades. Da wir den Mittelwert der Grundgesamtheit  $\mu$  nicht kennen, muss die Stichprobe zur Bestimmung von  $m$  herhalten. Dies ergibt eine neue Beziehung zwischen den Datensätzen, reduziert also die Anzahl Freiheitsgrade.

## Mittlerer Fehler des Mittelwertes

Bei verschiedenen Stichproben schwanken der Mittelwert und die Varianz. Wenn wir für die Berechnung des Mittelwertes die Schreibweise

$$\langle y \rangle = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

verwenden, und für den Erwartungswert

$$E(g(x)) = \sum_{i=1}^n p(x_i) g(x_i)$$

verwenden, dann gilt bei gleicher Grundgesamtheit für alle  $x_i$ , dass sie den Erwartungswert  $E(x_i) = \mu$  und  $E(x_i - \mu)^2 = \sigma^2$  ist. Wenn wir den Mittelwert der Messwerte,  $m$  einsetzen, erhalten wir auch  $E(m) = \mu$ . Für die Varianz gilt auch

$$\begin{aligned} \sigma_m^2 &= E\left[\frac{1}{n} \sum_{i=1}^n x_i - \mu\right]^2 = E\left[\frac{1}{n} \sum_{i=1}^n (x_i - \mu)\right]^2 \\ &= \frac{1}{n^2} \sum_{i=1}^n E[(x_i - \mu)^2] + \frac{1}{n^2} \sum_{i \neq j} E[(x_i - \mu)(x_j - \mu)] \end{aligned}$$

Wenn die Messdaten statistisch unabhängig sind, so ist der zweite Term=0 und wir erhalten

$$\sigma_m^2 = \frac{1}{n^2} n E[(x_i - \mu)^2] = \frac{\sigma^2}{n}$$

der mittlere Fehler  $s_m$  des Mittelwertes ist also um den Faktor  $\sqrt{n}$  kleiner als der mittlere Fehler der Einzelmessung

$$s_m = \frac{s}{\sqrt{n}}$$

Daraus lernt man, dass, um ein Resultat doppelt so genau zu erhalten, viermal mehr Messungen durchgeführt werden müssen.

## Momente der Verteilung

Das  $k$ -te Moment einer Verteilung ist definiert durch den Erwartungswert

$$\mu_k(x, B) = E[(x_i - B)^k]$$

Wenn wir  $B=0$  setzen, so erhalten wir

$$\begin{aligned} \mu'_1 &= E(x) = \bar{x} \\ \mu'_2 &= \sigma^2(0) = \overline{x^2} \end{aligned}$$

Bei den höheren Momenten benutzt man eine Normierung, damit diese dimensionslos werden.

Gebäuchlich außer dem Mittelwert und der Varianz sind:

- Die Schiefe (skewness):  $\gamma_1 = \frac{\mu_3}{\sigma^3}$ . Wenn die Schiefe positiv ist, heißt das, dass größere Abweichungen auf der positiven Seite liegen. Symmetrische Verteilungen haben die Schiefe 0.
- Die Überhöhung (peakedness):  $\gamma_2 = \frac{\mu_4}{\sigma^4}$ . Die Größe  $\gamma_2 - 3$  heißt Exzess, da sie die Abweichung von der Gaußverteilung angibt.

## Verteilungen

Es gibt einige Modellverteilungen, die in der Physik sehr gebräuchlich sind. Dies sind die Binominalverteilung, die bei Würfelexperimenten oder würfelartigen Experimenten zugrunde liegt, die Poisson-Verteilung und, im Grenzfall sehr großer Grundgesamtheiten, die Normalverteilung.

### Binominalverteilung

Die Binominalverteilung beschreibt Experimente, bei denen in jedem einzelnen Experiment mit der Wahrscheinlichkeit  $p$  ein Ereignis eintritt und mit der Wahrscheinlichkeit  $q=1-p$  ein zweites Ereignis eintritt (oder das erste nicht eintritt). Ein Würfelspiel genügt, zum Beispiel, diesen Gesetzen. Die Wahrscheinlichkeit, dass  $k$ -mal das erste und  $n-k$  mal das zweite Ereignis eintritt ist

$$p^k q^{n-k} = p^k (1-p)^{n-k}$$

Wenn die Reihenfolge dieser Ereignisse irrelevant ist, dann hat man die Binominalverteilung

$$B(k, n, p) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Die Verteilung ist normiert, da gilt

$$\sum_{k=0}^n B(k, n, p) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$$

Der Mittelwert ist

$$\langle k \rangle = \sum_{k=0}^n k \cdot B(k, n, p) = \sum_{k=1}^n \frac{n \cdot (n-1)!}{(k-1)! [n-1-k+1]!} p^k q^{n-1-k+1} = np$$

Die Varianz schließlich berechnet sich zu

$$\begin{aligned} \sigma^2 &= \sum_{k=0}^n k^2 B(k, n, p) - \bar{k}^2 \\ &= \sum_{k=0}^n k(k-1) \frac{n!}{k!(n-k)!} p^k q^{n-k} + \sum_{k=0}^n kB(k, n, p) - \bar{k}^2 \end{aligned}$$

und

$$\begin{aligned} \sigma^2 &= n(n-1)p^2 \sum_{k=2}^n \frac{(n-2)!}{(k-2)!(n-k)!} p^{k-2} q^{(n-2)-(k-2)} + \bar{k} - \bar{k}^2 \\ &= n(n-1)p^2 + np - n^2 p^2 = np(1-p) = npq \end{aligned}$$

Die weiteren Momente sind (ohne Rechnung)

$$\gamma_1 = \frac{1-2p}{\sqrt{npq}} = \frac{1-2p}{\sigma}$$

und

$$\gamma_3 - 3 = \frac{1-6pq}{npq} = \frac{1-6pq}{\sigma^2}$$

Für die Binominalverteilung existiert eine Rekursionsformel:

$$\frac{B(k, n, p)}{B(k-1, n, p)} = \frac{n-k+1}{k} \frac{p}{q}$$

Zum Abschluss noch eine kurze Bemerkung: jede, auch asymmetrische Binominalverteilung geht für festes  $p$  bei großen  $n$  in die Normalverteilung über.



## Normalverteilung

Die Normalverteilung kann als Grenzfall der Binominalverteilung angesehen werden, wenn die Anzahl Versuche gegen unendlich geht und  $p=q=0,5$  ist. Die Wahrscheinlichkeit für das Auftreten einer Abweichung ist

$$\Delta = x - \mu = \frac{1}{2} E(2N - k) - \frac{1}{2} E(k) = E(N - k)$$

wobei  $k$  negative und  $2N-k$  positive Abweichungen auftreten. Wir können die Binominalverteilung verwenden:

$$B\left(k, 2N, \frac{1}{2}\right) = \binom{2N}{k} 2^{-2N}$$

Wir berechnen den Mittelwert. Mit  $\bar{k} = 2Np = N$  wird

$$\bar{\Delta} = E(N - \bar{k}) = 0 \quad \text{oder} \quad \bar{x} = \mu$$

Die Varianz wird dann

$$\sigma_{\Delta}^2 = \overline{\Delta^2} = E^2(N - k)^2 = E^2\left(\frac{N}{2}\right)$$

Für große  $N$  ist es sinnvoll, die Verteilung auf die Varianz 1 zu standardisieren

$$t_k = \frac{\Delta}{\sigma_{\Delta}} = (N - k) \sqrt{\frac{2}{N}}$$

Diese standardisierte Variable  $t_k$  hat den Mittelwert 0 und die Varianz 1. Verwenden nun

Histogramme, und skalieren die Achse mit  $\sqrt{\frac{2}{N}}$ . Wir erhalten

$$G(t_k) = \sqrt{\frac{2}{N}} B\left(k, 2N, \frac{1}{2}\right) = \binom{2N}{k} 2^{-2N} \sqrt{\frac{2}{N}}$$

Mit der Stirlingschen Formel

$$n! \cong n^n e^{-n} \sqrt{2\pi n}$$

wird

$$\lim_{N \rightarrow \infty} (G(0)) = \frac{(2N)! \sqrt{N} 2^{-2N}}{(N!)^2 \sqrt{2}} = \frac{1}{\sqrt{2\pi}}$$

Mit Hilfe der Rekursionsformel kann gezeigt werden, dass

$$\frac{B(k) - B(k-1)}{B(k-1)} = \frac{n-k+1}{k} \frac{p}{q} - 1 = \frac{G(t_k) - G(t_{k-1})}{G(t_{k-1})} = \frac{\Delta G}{G}$$

Für  $p=q=1/2$  wird

$$\frac{\Delta G}{G} = \frac{2(N-k)+1}{N-(N-k)}$$

$$\frac{\Delta G}{\Delta t_k} \cdot \frac{1}{G} = -\sqrt{\frac{N}{2}} \frac{2\sqrt{\frac{N}{2}} \cdot t_k + 1}{N - \sqrt{\frac{N}{2}} \cdot t_k}$$

Für  $N \rightarrow \infty$  ergibt sich

$$\lim_{N \rightarrow \infty} \frac{\Delta G}{\Delta t_k} \cdot \frac{1}{G} = \frac{dG}{dt} \cdot \frac{1}{G} = \frac{d}{dt} \ln G = -t$$

Daraus folgt

$$G(t) = Ae^{-t^2/2}$$

Mit einer Normierung erhält man die standardisierte Gaußverteilung

$$G(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

oder die allgemeine Gaußverteilung

$$G(x, \sigma, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Die Gaußverteilung ist symmetrisch bezüglich des Mittelwertes  $\mu$ , also sind alle ungradzahligen Momente 0. Die gradzahligen Momente haben den Wert

$$\mu_{2\nu} = 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2\nu - 1) \sigma^{2\nu}$$

Damit wird zum Beispiel die Überhöhung  $\gamma_4=3$ .

Viele Experimente ergeben eine gaußförmige Verteilung der Messwerte um einen Mittelwert. Zur Abschätzung von Fehlergrenzen kann das Integral der Gaußverteilung, die Fehlerfunktion verwendet werden.

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_{-\infty}^z e^{-t^2} dt$$

$$z = \frac{x - \mu}{\sigma} = \frac{\sigma}{\sigma} = 1$$

Für den speziellen Wert  $\frac{x - \mu}{\sigma} = \frac{\sigma}{\sigma} = 1$  wird  $\text{erf}(1)=0,683$ . Das heißt: bei normalverteilten statistischen Abweichungen ist die Wahrscheinlichkeit = 68,3 %, dass eine Messung einen Fehler innerhalb  $\pm\sigma$  liefert.

## Poisson-Verteilung

Bei radioaktiven Atomen ist die Anzahl in einer bestimmten Zeit zerfallender Kerne proportional zur Gesamtzahl der Kerne. Es gilt also

$$\frac{dn}{dt} = -n\lambda$$

Daraus folgt das Zerfallsgesetz

$$n(t) = n_0 e^{-\lambda t}$$

Anstelle der Zerfallskonstante wird meistens die Halbwertszeit  $T_{1/2} = \ln 2 / \lambda$  angegeben. Unter den folgenden Annahmen kann man die dazugehörige Verteilungsfunktion ableiten.

- Die Zahl der radioaktiven Kerne sei sehr hoch (meistens sehr gut erfüllt).
- Die Anzahl der Zerfälle werde in jeweils konstanten Intervallen  $\Delta T$  bestimmt.
- Die Zerfallswahrscheinlichkeit  $n\lambda\Delta t$  sei in jedem Messintervall gleich, was gleichbedeutend ist mit  $\Delta t \ll T_{1/2}$ .

Viele Stöße ergeben nun eine Folge von Stoßzeiten ( $k_1, k_2, k_3, \dots$ ). Wir berechnen nun den Erwartungswert  $E(k)=\mu$ . Mit der Zerfallswahrscheinlichkeit  $p$  (proportional zu  $\lambda$  und  $\Delta t$ ) wird  $E(k) = \bar{k} = \mu = np$

Die Wahrscheinlichkeit für eine Zahl  $k$  von Ereignissen im Messintervall  $\Delta t$  kann aus der Binominalverteilung durch einen Grenzübergang nach unendlich ( $n \rightarrow \infty$ ,  $p \rightarrow 0$ ,  $q \rightarrow 1$ ,  $np = \mu = \text{const}$ ) berechnet werden. Wir verwenden die Rekursionsformel für die Binominalverteilung

$$B(k) = B(k-1) \left[ \frac{np}{kq} - \frac{(k-1)p}{kq} \right]$$

Daraus entsteht die Rekursionsformel für die Poisson-Verteilung

$$P(k, \mu) = P(k-1, \mu) \frac{\mu}{k}$$

Mit der Normierungsbedingung  $\sum_{k=0}^{\infty} P(k, \mu) = 1$  bekommt man

$$\sum_{k=0}^{\infty} P(k=0, \mu) \frac{\mu^k}{k!} = P(k=0, \mu) \sum_{k=0}^{\infty} \frac{\mu^k}{k!} = P(k=0, \mu) e^{\mu} = 1$$

Und daraus die Poisson-Verteilung

$$P(k, \mu) = \frac{\mu^k}{k!} e^{-\mu}$$

Die Poisson-Verteilung hat die folgenden Eigenschaften (erhalten aus dem Vergleich mit der Binominalverteilung):

- Mittelwert  $\bar{k} = E(k) = np = \mu$
- Varianz  $\sigma^2 = npq = \mu$  (np ist konstant, q geht gegen 1!!)
- Die Poisson-Verteilung ist normiert.
- Die Poisson-Verteilung hat nur einen Parameter

$$\frac{\sigma}{\mu} = \frac{1}{\sqrt{\mu}}$$

- Der relative mittlere Fehler ist
- Die Poisson-Verteilung findet man immer dann, wenn ein sehr unwahrscheinliches Ereignis bei einer großen Zahl Versuchen betrachtet wird. Neben Atomkernen sind auch die Ankunftszeiten von Photonen und Elektronen bei sehr geringem Fluss Poisson-verteilt.

### Lorentz-Verteilung

Die Lorentz-Verteilung tritt in optischen Spektren auf. Sie ist da die universelle Verteilungsfunktion. Sie wird in allgemeiner Form so geschrieben:

$$L(x, \mu, \Gamma) = \frac{1}{\pi} \frac{\Gamma/2}{(x - \mu)^2 + (\Gamma/2)^2}$$

Die Lorentz-Verteilung besitzt die folgenden Eigenschaften:

- Sie ist zu 1 normiert.
- Der Maximalwert ist  $2/\pi\Gamma$  bei  $x=\mu$
- Sie ist symmetrisch, alle ungeraden Momente sind also null.
- Die höheren geraden Momente, insbesondere die Varianz, sind nicht definiert, da die entsprechenden Integrale divergieren.
- Die charakteristische Dispersionsgrösse ist die Halbwertsbreite  $\Gamma$ .

$$L(\mu \pm \Gamma/2, \mu, \Gamma) = \frac{1}{2} L(\mu, \mu, \Gamma)$$

### Statistische Tests

Die in physikalischen Experimenten erhaltenen Daten stellen Stichproben aus einer Grundgesamtheit dar. Wir müssen nun die folgenden Fragen lösen:

- Stammen zwei Messreihen (Stichproben) aus der gleichen, unbekanntem Grundgesamtheit?
- Passt eine Messreihe zu früheren, ausgedehnten Stichproben (Hat die Probe sich verändert?)
- Passt eine Stichprobe zu einer Modellverteilung oder Hypothese?

Ein einfacher Test kann anhand der Standardabweichungen durchgeführt werden. So soll untersucht werden, ob bei einer Gesamtwurfwahl von 315672 die Zahl von 106602 Würfeln der Zahlen 5 oder 6 zur Annahme eines homogenen Würfels ( $p=1/3$ ) passt. Im Versuch ist die relative Häufigkeit  $0,3377=106602/315672$ . Weiter ist die Standardabweichung des Mittelwertes

$$s_m = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{x^2 - \bar{x}^2}{n}} = 0,00084$$

Man erwartet für einen homogenen Würfel  $\mu=105224$  mit

$$\sigma_m = \sqrt{npq} = 265$$

Die beobachtete Abweichung ist jedoch 1378, also 5,2 mal größer als  $\sigma$ . Dieser einfache Test kann gut zu einer ersten Abschätzung der Güte einer Messung dienen.

Ein weiterer einfacher Test benutzt den Vergleich der höheren Momente. Für diesen **Dispersionsindex** gilt

$$(3\gamma_1 - 2\gamma_2 + 6) \begin{cases} > 0 & \text{für Binomial- und Poissonverteilung} \\ < 0 & \text{für stark überhöhte Verteilungen} \\ = 0 & \text{für die Normalverteilung} \end{cases}$$

### t-Test

Der t-Test gibt eine Angabe über die Konsistenz zweier Mittelwerte. Er erlaubt eine Aussage, ob eine eigene Messung mit der eines andern (aber auch eine frühere eigene Messung) konsistent ist. Damit kann getestet werden, ob eine Apparatur sich mit der Zeit verändert.

Man könnte zur Annahme gelangen, dass wenn zwei Mittelwerte  $\bar{x}_1$  und  $\bar{x}_2$  sich um weniger als eine Standardabweichung unterscheiden, dass sie dann zu einer identischen Grundgesamtheit gehören. Diese Beurteilung ist willkürlich.

Besser ist es, den erwarteten Fehler der Differenz  $y = \bar{x}_1 - \bar{x}_2$  zu berechnen. Mit Hilfe des Fehlerfortpflanzungsgesetzes ergibt sich

$$s_{(\bar{x}_1 - \bar{x}_2)}^2 = s_{1m}^2 + s_{2m}^2 = \frac{\sum_{i=1}^{n_1} (x_{1i} - \bar{x}_1)^2 + \sum_{j=1}^{n_2} (x_{1j} - \bar{x}_1)^2}{n_1 + n_2 - 2} \frac{n_1 + n_2}{n_1 \cdot n_2}$$

dabei sind  $\bar{x}_1$  und  $\bar{x}_2$  die Mittelwerte zweier Messreihen,  $n_1$  und  $n_2$  die Anzahl Messungen mit

den Fehlern  $s_1$  und  $s_2$ . Für  $s_{im}^2$  setzt man am besten  $\frac{s_{12}^2}{n_i}$  ein, wobei die Varianz durch das Zusammenlegen aller Messreihen berechnet wurde.

Für die t-Verteilung und den t-Test betrachtet man normalverteilte Zufallsvariablen  $y$ . Sind solche nicht vorhanden, dann muss mit Mittelwerten von Stichproben gerechnet werden. Wir berechnen den Mittelwert  $m_y$  aus der Stichprobe  $y$  und die Standardabweichung  $s_y$ , die auf  $v$  Freiheitsgraden beruht. Die t-Grösse ist dann

$$t \equiv \frac{y - m_y}{s_y}$$

Sie gehorcht der Studentischen t-Verteilungsfunktion

$$\varphi(t, v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi v} \Gamma\left(\frac{v}{2}\right) \left(1 + \frac{t^2}{v}\right)^{(v+1)/2}}$$

Diese Verteilungsfunktion besitzt die Kenngrößen

$$\mu = 0$$

$$\sigma^2 = \frac{\nu}{\nu + 2}$$

$$\gamma_1 = 0$$

$$\gamma_2 = 3 + \frac{6}{\nu - 4}$$

Für  $n \Rightarrow \infty$  geht die Verteilung in die Normalverteilung über. Die integrale Verteilungsfunktion

$$T(t, \nu) = \int_{-\infty}^t \varphi dt + \int_t^{\infty} \varphi dt = 2 \int_t^{\infty} \varphi(t, \nu) dt$$

ist die Wahrscheinlichkeit, dass bei Abwesenheit systematischer Fehler t-Werte auftreten, die außerhalb des Konfidenzbereiches  $\pm t$  liegen. Diese Funktion findet man in Tabellen. Für sehr umfangreiche Stichproben geht die t-Verteilung in die Normalverteilung über.

Für die (0,1) [Mittelwert 0, Varianz 1] verteilte Testgröße bekommt man

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{(\bar{x}_1 - \bar{x}_2)}}$$

Wenn der Test erfüllt ist, dann sind die beiden Stichproben aus der gleichen Grundgesamtheit.

Zwei Mittelwerte mit  $n_1=9$  und  $n_2=7$  ergeben einen t-Wert von 1,74

Man findet  $T(t, \nu) = T(1,74, 14) = 0,1$

(Die entsprechende Excel Funktion =**TVERT(1,74;14;2)** liefert 0,103784377)

Die Interpretation sagt nun, dass

- Nur in 10% aller Stichproben tritt eine Abweichung  $>10\%$  auf
- Der Test wird von den beiden Messreihen erfüllt mit einer Signifikanz-Grenze von 10%, d.h. systematische Abweichungen sind wenig wahrscheinlich
- Die Mittelwerte müssen als konsistent angesehen werden.

Üblicherweise verwendet man Vertrauensgrenzen (Signifikanz-Grenzen) von 5% und 1%. Nur wenn der t-Test eine Wahrscheinlichkeit kleiner als diese Werte ergibt, nimmt man eine Abweichung an. Größere Signifikanz-Grenzen (z.B. 40%) bedeuten, dass man sehr oft die Hypothese der Gleichheit verwerfen muss, obwohl sie stimmt. Mit sehr kleinen Signifikanz-Grenzen läuft man Gefahr, Gleichheit anzunehmen, obwohl sie nicht da ist.

## F-Test

Der F-Test ist ein zum t-Test analoger Test, der die Konsistenz von Varianzen prüft.

$$F \equiv \frac{s_1^2}{s_2^2} \quad \text{wobei } s_1 > s_2$$

Die Verteilungsfunktion enthält die beiden Freiheitsgrade  $\nu_1 = n_1 - 1$  und  $\nu_2 = n_2 - 1$ . In Tabellen werden für 5% und 1% Signifikanz-Grenzen Werte angegeben.

## Chi-Quadrat-Test

Beim  $\chi^2$ -Test liefert ein allgemeines Kriterium für die Übereinstimmung der Grundgesamtheit mit der Stichprobe. Der  $\chi^2$ -Test taugt für jede Verteilungsfunktion, ist also modellfrei. Das Vorgehen ist folgendermaßen:

1. Man ordnet die Messwerte in Klassen, so dass in jede Klasse mindestens 5 Werte fallen. Die Anzahl Beobachtungen sei  $f_{ex}$
2. Es sollen mindestens  $M \geq 6$  Klassen vorhanden sein.

3. Aus der Hypothese, der Verteilungsfunktion, die man testen möchte, berechnet man die erwartete Häufigkeit  $f_{th}$  für jede Klasse.
4. Man berechnet aus der Klassenbreite, dem Klassenmittelwert und der Dichte der Verteilungsfunktion die erwartete Häufigkeit  $f_{th}$ .

$$\sum_{j=1}^M f_{th,j} = N$$

5. Normierung:

$$\chi^2 \equiv \sum_{j=1}^M \frac{(f_{th,j} - f_{ex,j})^2}{f_{th,j}}$$

6. Die Größe ist ein integrales Maß für die statistischen Abweichungen der beobachteten Häufigkeiten.

Die Verteilungsfunktion lautet:

$$W(\chi^2) = \frac{\chi^{(v-2)} \cdot e^{(-\chi^2/2)}}{2^{v/2} \cdot \Gamma(v/2)}$$

Die kennzeichnenden Größen sind:

$$\mu = v$$

$$\sigma^2 = 2v$$

$$\gamma_1 = \sqrt{8/v}$$

$$\gamma_2 = 3 + 12/v$$

Die  $\chi^2$ -Verteilungsfunktion ist verwandt mit der Poisson-Verteilung: sie geht für große  $v$  in die Normalverteilung über. Wir rechnen nun die Wahrscheinlichkeit für den  $\chi^2$ -Wert aus der Probe aus. Der Test, ob die Stichprobe zur hypothetischen Grundgesamtheit passt, gilt als bestanden, wenn die Wahrscheinlichkeit  $>5\%$  ist.

Für die Anzahl Freiheitsgrade gilt:

$$v = M - K = \text{Anzahl Klassen} - \text{Anzahl Beschränkungen}$$

Solche Beschränkungen sind die Normierung von  $f_{th}$  und die Mittelwertbildung.

Für den  $\chi^2$ -Test ergibt sich demnach

- Normalverteilung:  $v=M-3$
- Poisson-Verteilung:  $v=M-2$
- Modellverteilungen mit a priori-Annahmen über  $\mu$  und  $\sigma$ :  $v=M-1$ .

Wenn ich bei einer Stichprobe mit 8 Messwerten ausrechne, dass der Mittelwert  $\bar{x} = 14,128768432$  und die Standardabweichung  $s_m = 0,142676871234$ , so wäre eine Resultatangabe von

$$\bar{x} = 14,128768432 \pm 0,142676871234$$

$$\frac{\sigma_s}{s} \cong 25\%$$

vollkommen unsinnig. Ist der Fehler nämlich normalverteilt, so erhält man für  $s$ , also sind 4 Stellen für den Mittelwert und 2 für die Abweichung angebracht:

$$\bar{x} = 14,12 \pm 0,14$$

## Fehlerfortpflanzung

Im allgemeinen besteht ein Resultat eines physikalischen Experimentes nicht aus einer einzelnen Messgröße. Wir können also die oben abgeleiteten Verfahren nicht unbesehen auf beliebige experimentelle Resultate ausdehnen. Wie wirken sich also die Einzelfehler auf die Gesamtheit aus?

Als Beispiel soll die Bestimmung einer Gitterkonstanten  $a$  nach der Debye-Scherrer-Methode besprochen werden. Röntgenstrahlen mit der Wellenlänge  $\lambda$  fallen auf eine größere Anzahl kleiner Kristallite mit den Miller'schen Indizes  $h, k, l$ . Für den Glanzwinkel gilt:

$$\sin \Theta = \frac{\lambda}{2a} \sqrt{h^2 + k^2 + l^2}$$

oder

$$a = \frac{\lambda}{2 \sin \Theta} \sqrt{h^2 + k^2 + l^2}$$

Es treten dabei die folgenden Fehler auf:

- (a) Die Wellenlänge  $\lambda$  ist nur auf  $\pm s_\lambda$  bestimmt.
- (b) Der Wert für die Wellenlänge stammt aus einem alten Tabellenwerk, ist also nicht mehr korrekt.
- (c) Die Korrektur für die Adsorption der Probe wird vernachlässigt.
- (d) Der Winkel  $\Theta_{h,k,l}$  wird  $n$ -mal gemessen: der Mittelwert sei  $\overline{\Theta_{h,k,l}}$  und der mittlere Fehler sei  $\bar{\sigma}$ .

(a), (b) und (c) sind systematische Fehler. Bei (b) und (c) lassen sich Vorzeichen und Größe des Fehlers bestimmen. Bei (a) ist der Fehler für unsere Messung ein systematischer, beruht aber ursprünglich auf der Bestimmung eines zufälligen Fehlers, so dass  $\pm s_\lambda$  zur Abschätzung des systematischen Fehlers verwendet werden kann.

### Systematische Fehler

Die direkt gemessenen Größen  $a_1, a_2, \dots, a_N$  seien mit den systematischen Fehlern  $h_1, h_2, \dots, h_N$  verfälscht. Gesucht ist die Größe  $A$ , die eine Funktion der  $N$  gemessenen Werte ist. Wir haben aber die verfälschte Größe  $A_0$  bestimmt. Für sie gilt, nach einer Taylor-Entwicklung

$$A(a_1 + h_1, a_2 + h_2, \dots, a_N + h_N) = A(a_1, a_2, \dots, a_N) + \frac{1}{1!} \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right)_{a_1 \dots a_N} \cdot h_k + \frac{1}{2!} \sum_{k,l=1}^N \left( \frac{\partial^2 A}{\partial a_k \partial a_l} \right)_{a_1 \dots a_N} \cdot h_k \cdot h_l + \dots$$

Da in aller Regel die Fehler klein gegen die gemessenen Werte sind, darf mit der linearen Näherung gerechnet werden.

Also ist das Fehlerfortpflanzungsgesetz für systematische Fehler

$$A(a_1 + h_1, a_2 + h_2, \dots, a_N + h_N) - A(a_1, a_2, \dots, a_N) = \Delta A = \frac{1}{1!} \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right)_{a_1 \dots a_N} \cdot h_k$$

Beim Übergang zu den relativen Fehlern erhält man:

$$\frac{\Delta A}{A} = \sum_{k=1}^N \left( \frac{1}{A} \frac{\partial A}{\partial a_k} \right) h_k$$

### Rechenregeln

- Summe oder Differenz: Wenn  $A = \alpha a_1 \pm \beta a_2$  ist, dann ist  $\Delta A = \alpha h_1 \pm \beta h_2$
- Produkte: Für  $A = a_1 \cdot a_2$  ergibt sich  $\Delta A = a_2 \cdot h_1 + a_1 \cdot h_2$ . Einfacher ist die Gleichung mit

$$\frac{\Delta A}{A} = \frac{h_1}{a_1} + \frac{h_2}{a_2}$$

den relativen Fehlern:  $\frac{\Delta A}{A} = \frac{h_1}{a_1} + \frac{h_2}{a_2}$ .

- Bei beliebigen Potenzen gilt für  $A = \alpha a_1^p \cdot a_2^q$ :  $\frac{\Delta A}{A} = p \frac{h_1}{a_1} + q \frac{h_2}{a_2}$  (Dies erhält man mit der logarithmischen Ableitung)

$$A = \frac{\alpha a_1^p}{a_2^q} \quad \text{ergibt sich:} \quad \frac{\Delta A}{A} = p \frac{h_1}{a_1} - q \frac{h_2}{a_2}$$

- Quotienten: Für  $A = \frac{\alpha a_1^p}{a_2^q}$  ergibt sich:  $\frac{\Delta A}{A} = p \frac{h_1}{a_1} - q \frac{h_2}{a_2}$
- Beim Beispiel in der Einleitung erhält man für die Auswirkung der  $\lambda$ -Fehler:  $\frac{\Delta a}{a} = \frac{\Delta \lambda}{\lambda}$

$$\Delta a = \frac{-\lambda \sqrt{h^2 + k^2 + l^2}}{2 \sin^2 \Theta} \cos \Theta \cdot \Delta \Theta \quad \text{oder}$$

$$\frac{\Delta a}{a} = -\cot \Theta \cdot \Delta \Theta$$

. Daraus schließt man, dass bei  $90^\circ$  (Rückstreuung) der Fehler minimal ist. Hier ist ein Beispiel, das zeigt, dass mit einer geschickten Wahl der experimentellen Anordnung der Messfehler minimal wird.

## Statistische Fehler

### Unabhängigkeit von der Reihenfolge der Auswertung

Wir nehmen wieder an, dass die gesuchte Größe  $A$  aus den direkt gemessenen Größen  $a_1, a_2, \dots, a_N$  bestimmt sei. Diese sind mit zufälligen Fehlern behaftet. Wenn  $a_k$   $n$ -mal gemessen wurde, so bezeichnen wir die  $j$ -te Messung der  $k$ -ten Größe mit  $a_{kj}$ . Der Mittelwert von  $a_k$  ist

$$\bar{a}_k = \frac{1}{n} \sum_{j=1}^n a_{kj}$$

und die Varianz

$$s_{a_k}^2 = \frac{1}{n-1} \sum_{j=1}^n (a_{kj} - \bar{a}_k)^2 = \frac{1}{n-1} \sum_{j=1}^n (\varepsilon_{kj})^2$$

es gibt nun zwei Wege, wie wir von den gemessenen Größen auf das Endresultat kommen können. Sind die beiden Wege äquivalent?

Wir können zuerst jede einzelne Messung mitteln, und dann in unsere Funktion für das Endresultat einsetzen.

$$A_m = A(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n)$$

Wir können auch jeweils die  $j$ -te Wiederholung der Messung einsetzen und dann mitteln.

$$\bar{A} = \frac{1}{n} \sum_{j=1}^n A_j = \frac{1}{n} \sum_{j=1}^n A(a_{1j}, a_{2j}, \dots, a_{nj})$$

Nun lässt sich der Wert von  $A_j$  durch eine Taylor-Entwicklung angeben.

$$A_j = A_m(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) + \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right) \varepsilon_{kj} + \dots$$

Wenn die Abweichungen von den Mittelwerten klein sind gegen die Mittelwerte, haben wir

$$\bar{A} = \frac{1}{n} \sum_{m=1}^n A_m + \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right) \cdot \varepsilon_{kj}$$

Wenn wir nun im 2. Glied die Reihenfolge der Summation vertauschen, dann finden wir

$$\frac{1}{n} \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right) \cdot \sum_{j=1}^n \varepsilon_{kj} = \frac{1}{n} \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right) \cdot 0 = 0$$



Also gilt:

$$\bar{A} = \frac{1}{n} \sum_{j=1}^n A_j = A_m(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_N)$$

Die obige Formel gilt ziemlich allgemein, z.B. auch wenn nicht alle Teilresultate gleich häufig gemessen wurden.

### Gauß'sches Fehlerfortpflanzungsgesetz

Wir nehmen an, dass die Varianzen  $s_k^2$  der gemittelten Messwerte  $a_k$  bekannt seien. Die Varianzen sind, nach der allgemeinen Definition,

$$\begin{aligned} s_A^2 &= \frac{1}{n-1} \sum_{j=1}^n (A_j - \bar{A})^2 = \frac{1}{n-1} \sum_{j=1}^n \left( \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right) \varepsilon_{kj} \right)^2 \\ s_A^2 &= \frac{1}{n-1} \sum_{j=1}^n \left( \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right)^2 \varepsilon_{kj}^2 + \sum_{\substack{l,m=1 \\ l \neq m}}^N \left( \frac{\partial A}{\partial a_l} \right) \left( \frac{\partial A}{\partial a_m} \right) \varepsilon_{lj} \varepsilon_{mj} \right) \\ s_A^2 &= \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right)^2 \frac{\sum_{j=1}^n \varepsilon_{kj}^2}{n-1} + \sum_{\substack{l,m=1 \\ l \neq m}}^N \left( \frac{\partial A}{\partial a_l} \right) \left( \frac{\partial A}{\partial a_m} \right) \frac{\sum_{j=1}^n \varepsilon_{lj} \varepsilon_{mj}}{n-1} \end{aligned}$$

Wenn die Messungen der  $a_k$  voneinander statistisch unabhängig sind, dann summieren sich alle Kreuzprodukte zu 0. Also ist das **Gauß'sche Fehlerfortpflanzungsgesetz**

$$s_A^2 = \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right)^2 s_{a_k}^2$$

Teilen wir beide Seiten mit  $n$ , erhält man das Fehlerfortpflanzungsgesetz für die mittleren Fehler der Mittelwerte, die in der Praxis am meisten angewandte Form.

$$s_A^2 = \sum_{k=1}^N \left( \frac{\partial A}{\partial a_k} \right)^2 s_{\bar{a}_k}^2$$

### Statistische Unabhängigkeit

$$\sum_{j=1}^n \varepsilon_{lj} \varepsilon_{mj}$$

Wenn die Messwerte  $a_k$  statistisch unabhängig sind, dann treten in  $\frac{\sum_{j=1}^n \varepsilon_{lj} \varepsilon_{mj}}{n-1}$  positive und negative Zeichen mit gleicher Häufigkeit auf. Wenn diese Summe null ist, nennt man die beiden Messgrößen  $a_k$  und  $a_m$  statistisch unabhängig. Dies wird mit dem Begriff der Kovarianz beschrieben

$$Kov(x, y) = E\{(x - E(x))(y - E(y))\}$$

wobei der Operator  $E$  den Erwartungswert berechnet. Wieder gilt, dass wenn die Größen  $x$  und  $y$  voneinander statistisch unabhängig sind, dass die Produkte in der Summe für den Erwartungswert mit gleicher Häufigkeit und beide Vorzeichen annehmen.

Wenn wir eine vollständige Korrelation haben, also  $\varepsilon_{kj} = \eta \varepsilon_{lj}$ , dann gilt

$$Kov(\varepsilon_k, \eta \varepsilon_k) = E\{\varepsilon_k \cdot \eta \cdot \varepsilon_k\} = \eta E\{\varepsilon_k^2\} = \eta \cdot Var(\varepsilon_k) = \eta \sigma_k^2$$

Die Kovarianz geht also im Falle einer vollständigen Korrelation in die Varianz über.

Das Gauß'sche Fehlerfortpflanzungsgesetz darf nicht angewandt werden, wenn Korrelationen existieren (angezeigt durch die Kovarianz)

- Physikalische Korrelationen: zum Beispiel wenn gleichzeitig Strom und Spannung gemessen werden.
- Algebraische Korrelationen: Die Physikalischen Grundkonstanten sind korreliert. Man muss in diesem Falle die Fehlermatrix berechnen (Siehe Gränicher)

### Angabe von zufälligen Fehlern

Da nach dem Gauß'schen Fehlerfortpflanzungsgesetz positive und negative Fehler gleich häufig vorkommen, muss der Fehler wie folgt angegeben werden:

$$\bar{A} \pm s_{\bar{A}}$$

Es gelten die folgenden Rechenregeln:

- Summen und Differenzen: Mit  $A = a_1 \pm a_2$  ist  $s_A^2 = s_{a_1}^2 + s_{a_2}^2$

- Lineare Funktionen  $A = \sum_{i=1}^n c_i a_i$  ergibt  $s_A^2 = \sum_{i=1}^n c_i^2 s_{a_i}^2$

- Produkte werden mit den relativen Fehlern berechnet. Aus  $A = a_1^{p_1} \cdot a_2^{p_2} \cdot \dots$  erhält man für

$$\left(\frac{s_A}{A}\right)^2 = p_1^2 \left(\frac{s_{a_1}}{a_1}\right)^2 + p_2^2 \left(\frac{s_{a_2}}{a_2}\right)^2 + \dots$$

die Fehlerquadrate

- Für das Bragg'sche Gesetz ergibt sich  $\left(\frac{s_A}{A}\right)^2 = \left(\frac{s_\lambda}{\lambda}\right)^2 + \cot^2 \Theta \cdot s_\Theta^2$

- Bei zusammengesetzten ausdrücken kann man zuerst Teilvarianzen berechnen.

### Maximaler Fehler

Will man nicht so viel rechnen, dann verwendet man die zu der Fehlerfortpflanzung bei systematischen Fehlern verwendete Funktion, bildet aber von jedem Summanden den Betrag.

## Ausgleichsrechnung "Fitten"

Wenn bei einer Messung ein funktionaler Zusammenhang bestätigt werden sollte, möchte man die gemessenen Daten an theoretische oder empirische funktionale Zusammenhänge anpassen. Insbesondere wenn mehr Messdatensätze als freie Parameter existieren, muss eine Anpassung durchgeführt werden.

### Graphische Verfahren

Ein graphischer Ausgleich, d.h. man legt die angepasste Kurve so über die Daten zu legen, dass die Gerade die gemessenen Daten am besten anpaßt. Dazu müssen die Daten so dargestellt werden, dass ein linearer Zusammenhang zwischen den Variablen existiert. Dies

kann zum Beispiel bei einer Exponentialfunktion  $n(t) = A e^{(-Bt)}$ , wie sie bei Zerfallsprozessen auftreten kann, durch logarithmieren und Darstellen der Funktion  $\ln[n(t)] = \ln[A] - Bt$  geschehen.

Praktische Regeln für einen graphischen Ausgleich:

- Die mittleren Fehler der einzelnen Messpunkte einzeichnen (in beide Koordinatenachsen)
- Durchsichtigen Maßstab oder Faden verwenden.
- Endpunkte nicht überschätzen, da sie nicht genauer, oft sogar schlechter bekannt sind als die andern Punkte.

- Curie-Weiss Gesetz:  $\varepsilon = \varepsilon_u + \frac{C}{T - \theta}$  ( $\varepsilon_u$  ist die Untergrund-Suszeptibilität,  $\theta$  die Curie-Weiss-Temperatur) wird in der Darstellung  $\frac{1}{\varepsilon - \varepsilon_u} = \frac{1}{C}(T - \theta)$  linear angepasst.
- Thermoemissionsgesetz:  $j_s = CT^2 e^{(-\phi/kT)}$  ( $\phi$  ist die Austrittsarbeit,  $C$  eine Mengenkonstante) wird in der Darstellung  $\ln\left(\frac{j_s}{T^2}\right) = \ln(C) - \frac{\phi}{k}\left(\frac{1}{T}\right)$  angepasst.

### Numerische Verfahren

Wir betrachten nun zuerst den einfachen Fall einer linearen Beziehung zwischen zwei Größen. Bei einer Feder wären dies die Auslenkung  $z$ , die angewandte Kraft  $F$  und die Lage des Nullpunktes  $z_0$ . Ohne Verlust der Allgemeinheit kann  $z_0=0$  gesetzt werden. Folgende Annahmen können nun gemacht werden:

1. Die Messung der Kraft  $F$  ist fehlerbehaftet
2. Die Messung der Position  $z$  ist fehlerbehaftet
3. Beide Messgrößen sind fehlerbehaftet.

#### Kraft $F$ ist fehlerbehaftet

In diesem Falle können wir den Fehler für den  $i$ -ten Messpunkt wie folgt angeben:

$$F_i - kz_i = \varepsilon_i$$

Das Prinzip der kleinsten Fehlerquadrate verlangt nun, dass

$$\sum_{i=1}^n \varepsilon_i^2 = [\varepsilon \varepsilon] = \text{minimal}$$

ist. Also

$$\frac{\partial}{\partial k} \left( \sum_{i=1}^n \varepsilon_i^2 \right) = \frac{\partial}{\partial k} \left( \sum_{i=1}^n (F_i - k_F z_i)^2 \right) = 2 \sum_{i=1}^n ((F_i - k_F z_i)(-z_i)) = 0$$

Diese Gleichung kann umgeschrieben werden:

$$0 = k_F \sum_{i=1}^n z_i^2 - \sum_{i=1}^n F_i z_i = k_F [z z] - [F z]$$

oder

$$k_F = \frac{[F z]}{[z z]}$$

#### Auslenkung $z$ ist fehlerbehaftet

Wir können die Gleichung für den Zusammenhang zwischen der Auslenkung und der Kraft auch so umschreiben, dass  $z$  unsicher ist. Dann gilt

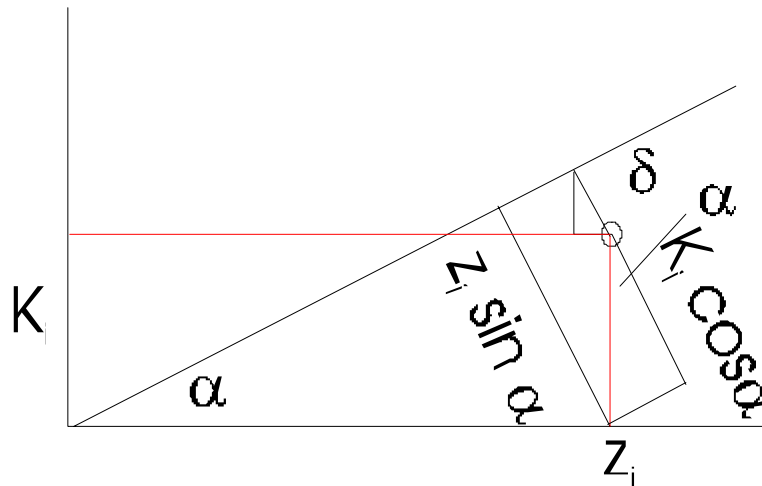
$$z_i - \frac{F_i}{k} = \varepsilon'_i$$

Eine kurze Rechnung zeigt, dass

$$k_z = \frac{[F F]}{[F z]}$$

#### Auslenkung $z$ und Kraft fehlerbehaftet

Wenn beide Messgrößen fehlerbehaftet sind, gilt:



Anhand der Zeichnung findet man heraus, dass

$$F_i \cos \alpha + \delta = z_i \sin \alpha$$

oder

$$\delta = z_i \sin \alpha - F_i \cos \alpha$$

Die Summe der Abstandsquadrate muss null sein, also muss

$$\sum_{i=1}^n \delta_i^2 = \sum_{i=1}^n (z_i \sin \alpha - F_i \cos \alpha)^2$$

minimal sein. Also ist

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha} \sum_{i=1}^n \delta_i^2 = \frac{\partial}{\partial \alpha} \sum_{i=1}^n (z_i \sin \alpha - F_i \cos \alpha)^2 \\ &= \sum_{i=1}^n 2(z_i \sin \alpha - F_i \cos \alpha)(z_i \cos \alpha + F_i \sin \alpha) \\ &= 2 \sum_{i=1}^n (z_i \sin \alpha \cos \alpha - F_i \sin \alpha \cos \alpha + z_i F_i \sin^2 \alpha - z_i F_i \cos^2 \alpha) \\ &= 2 \sum_{i=1}^n ((z_i^2 - F_i^2) \sin \alpha \cos \alpha + z_i F_i (\sin^2 \alpha - \cos^2 \alpha)) \\ &= \sum_{i=1}^n ((z_i^2 - F_i^2) \sin 2\alpha - 2z_i F_i \cos 2\alpha) \\ &= \sin 2\alpha \sum_{i=1}^n (z_i^2 - F_i^2) - 2 \cos 2\alpha \sum_{i=1}^n z_i F_i \end{aligned}$$

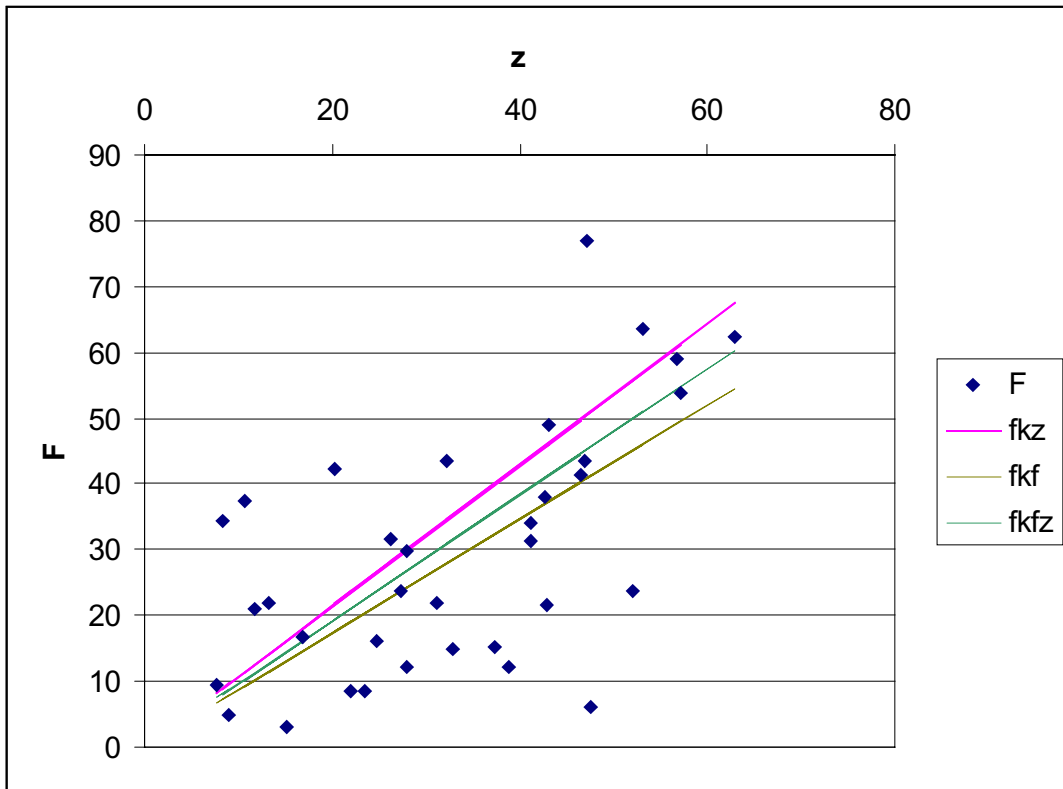
Also wird

$$\tan 2\alpha = \frac{2 \sum_{i=1}^n z_i F_i}{\sum_{i=1}^n z_i^2 - \sum_{i=1}^n F_i^2} = \frac{2[zF]}{[zz] - [FF]}$$

Nun ist aber

$$k = \tan \alpha = \tan \left( \arctan \left( \frac{2[zF]}{[zz] - [FF]} \right) / 2 \right)$$

Als Beispiel ist in der folgenden Figur eine mit Zufallszahlen simulierte Rechnung zu sehen



Hier bedeuten fkz die Messkurve, wenn nur z fehlerbehaftet ist, fkf, diejenige, bei der F einen Fehler hat sowie fkfz diejenige, bei der beide Größen einen Fehler haben. Bei der letzten Kurve wurde die quadratische Anpassung durchgeführt. Wichtig ist, dass dafür die **Federkonstante k auf 1 umskaliert werden muss**.

### Ausgleichsrechnung für allgemeine Beziehungen

Wir betrachten nun die allgemeine Beziehung

$$z_i = \sum_{j=1}^m A_j f_j(x_{1,i} \dots x_{N,i})$$

bei der die unbekannt Koeffizienten  $A_j$  linear in den Gleichungen vorkommen. Wir fordern wieder, dass

$$[\varepsilon \varepsilon] = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n \left\{ \sum_{j=1}^m A_j f_j(x_{1,i} \dots x_{N,i}) - z_i \right\}^2 = \text{Minimum}$$

ist. Die Gleichung kann auch in Matrixschreibweise geschrieben werden:

$$\begin{pmatrix} f_1(x_{1,1} \dots x_{N,1}) & \dots & f_m(x_{1,1} \dots x_{N,1}) \\ \vdots & & \vdots \\ f_1(x_{1,n} \dots x_{N,n}) & \dots & f_m(x_{1,n} \dots x_{N,n}) \end{pmatrix} \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

Mit der Abkürzung

$$B_{i,j} = f_j(x_{1,i} \dots x_{N,i})$$

wird die Gleichung

$$\begin{pmatrix} B_{1,1} & \dots & B_{1,m} \\ \vdots & & \vdots \\ B_{n,1} & \dots & B_{n,m} \end{pmatrix} \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

oder

$$B\vec{A} = \vec{z}$$

Die Lösung dieses überbestimmten Gleichungssystems im Sinne der Minimierung der Fehlerquadrate ist die Lösung des  $m \times m$ - Gleichungssystems

$$(B^T B)\vec{A} = B^T \vec{z}$$

also

$$\vec{A} = (B^T B)^{-1} B^T \vec{z}$$

Bei allgemeinen funktionalen Abhängigkeiten sucht man, mit welcher Methode auch immer, Näherungswerte für die unbekannt Parameter. Ist zum Beispiel

$$g_i(A, B, C, \dots, x_i, y_i, \dots) - L_i = \varepsilon_i$$

und  $A_0, B_0, C_0$  Näherungswerte, so kann eine Taylor-Entwicklung durchgeführt werden.

$$g_i(A, B, C, \dots) = g_{i,0}(A_0, B_0, C_0, \dots) + \left( \frac{\partial g_i}{\partial A} \right)_{A_0, B_0, C_0, \dots} (A - A_0) + \dots$$

Mit  $a_i = \left( \frac{\partial g_i}{\partial A} \right)_{A_0, B_0, C_0, \dots}$ ,  $b_i = \left( \frac{\partial g_i}{\partial B} \right)_{A_0, B_0, C_0, \dots}$ , u.s.w.,  $\ell_i = L_i - g_{i,0}$  und  $\xi = A - A_0$ ,  $\eta = B - B_0$  u.s.w. ist die Gleichung wieder in linearisierter Form vorhanden

$$a_i \xi + b_i \eta + \dots - \ell_i = \varepsilon_i$$

Die oben skizzierte Methode kann zur Lösung verwendet werden. Weiter Informationen können im Buch von Zurmühl gefunden werden.

### Verwenden von Excel zum Fitten

Die Anpassung von empirisch gefundenen oder theoretisch begründeten funktionalen Zusammenhängen an gemessene Daten ist nicht in jedem Falle einfach. Bei lineare Zusammenhängen funktionieren die meisten Programme sehr gut. Bei nichtlinearen Abhängigkeiten, die insbesondere über mehrere Größenordnungen betrachtet werden müssen, sind meistens Probleme zu befürchten.

Die Datenanpassung beruht auf den im obigen Kapitel angegebenen Verfahren, unter Umständen auch auf wesentlich involvierteren Verfahren wie sie in den Referenzen 1 und 2 beschrieben werden.

In dieser Vorlesung sollen Sie mit den Möglichkeiten einer Tabellenkalkulation für diese Anpassungen vertraut gemacht werden. Natürlich können auch andere Programmpakete, oft sogar besser, als Excel, das hier als Beispiel genommen wird, verwendet werden. Excel ist jedoch unschlagbar, wenn es um das Antasten an ein Anpassungsverfahren geht.

Als Beispiel verwende ich pV-Daten des Gases CO<sub>2</sub>. Diese Daten wurden mit einem, sich im Aufbau befindlichen Experiment des Grundpraktikums. Das Experiment läuft so ab, dass mit einer Mikrometerschraube das Probenvolumen, ausgehend von einem unbekannt Volumen  $V_0$  verändert wird. Eine weitere Unbekannte ist die Molzahl, da diese nur schlecht zu bestimmen ist. Wir starten mit einem leeren Bildschirm.

In diesen Bildschirm tragen wir nun oben den Titel ein

Microsoft Excel - APrechnung.xls

Datei Bearbeiten Ansicht Einfügen Format Extras Daten Fenster ?

Arial 10 F K U

B4 =  $V = V_0 - (50-x) \cdot 1.539 \cdot 10^{-7} \text{ m}^3$

	A	B	Formeln	D	E	F	G
1	Rohdaten	CO2					
2							
3	Gleichungen						
4	Volumen	$V = V_0 - (50-x) \cdot 1.539 \cdot 10^{-7} \text{ m}^3$					
5	Druck	100 mV = 150 bar = 15 Mpa					
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							

Markieren Sie den Zielbereich, und drücken Sie die

Es ist eine gute Praxis, die relevanten Gleichungen im Seitenkopf anzugeben. Hier enthalten die Zellen B4 und B5 die Umrechnungsgleichungen von den gemessenen Daten auf die physikalischen Größen.

Microsoft Excel - APrechnung.xls

Datei Bearbeiten Ansicht Einfügen Format Extras Daten Fenster ?

Arial 10 F K U

B9 =  $p = nRT / (V - nb) - a \cdot n^2 / V^2$

	A	B	C	D	E	F	G
1	Rohdaten	CO2					
2							
3	Gleichungen						
4	Volumen	$V = V_0 - (50-x) \cdot 1.539 \cdot 10^{-7} \text{ m}^3$					
5	Druck	100 mV = 150 bar = 15 Mpa					
6							
7							
8	Reales Gas	$(p + a \cdot n^2 / V^2) \cdot (V - nb) = nRT$					
9		$p = nRT / (V - nb) - a \cdot n^2 / V^2$					
10							
11							
12							
13							
14							
15							
16							
17							

Markieren Sie den Zielbereich, und drücken Sie die

In der Zelle B8 geben wir die Gleichung des realen Gases ein, in der Zelle B9 die Auflösung nach p.

Microsoft Excel - APrechnung.xls

File Edit View Insert Format Extras Data Window ?

Arial 10 F K U

0,001

	A	B	C	D	E	F	G
1	<b>Rohdaten</b>	CO2					
2							
3	<b>Gleichungen</b>						
4	Volumen	$V = V_0 - (50-x) \cdot 1.539e-7 \text{ m}^3$					
5	Druck	100 mV = 150 bar = 15 Mpa					
6							
7							
8	Reales Gas	$(p+an^2/V^2)(V-nb)=nRT$					
9		$p=nRT/(V-nb)-an^2/V^2$					
10							
11	Unbekannte						
12	V0	0,001		m <sup>3</sup>			
13	n=	0,001		mol			
14							
15							
16							
17							

75,6C (3) 75,6C (4) 5,4C (2) Tabelle1

Bearbeiten

Wir geben nun in den Zellen A12 und A13 die Bezeichnungen der unbekannt Größen ein. Die Zellen B12 und B13 enthalten Anfangswerte, die Zellen C12 und C13 die physikalischen Einheiten.

Microsoft Excel - APrechnung.xls

File Edit View Insert Format Extras Data Window ?

Arial 10 F K U

H11 = 0,00003148

	E	F	G	H	I	J	K	L
1								
2								
3		Konstanten						
4			Tk	304,21	K			
5			RHO <sub>k</sub>	0,466	kg/l			
6								
7			M	0,04401	kg/mol			
8			P <sub>k</sub>	7382500	Pa			
9			R	8,314	J/k mol			
10		vdW-Paramet	a	2,69E-01	m <sup>6</sup> *Pa/mol <sup>2</sup>			
11			b	3,15E-05	m <sup>3</sup> /mol			
12								
13								
14								
15								
16								
17								

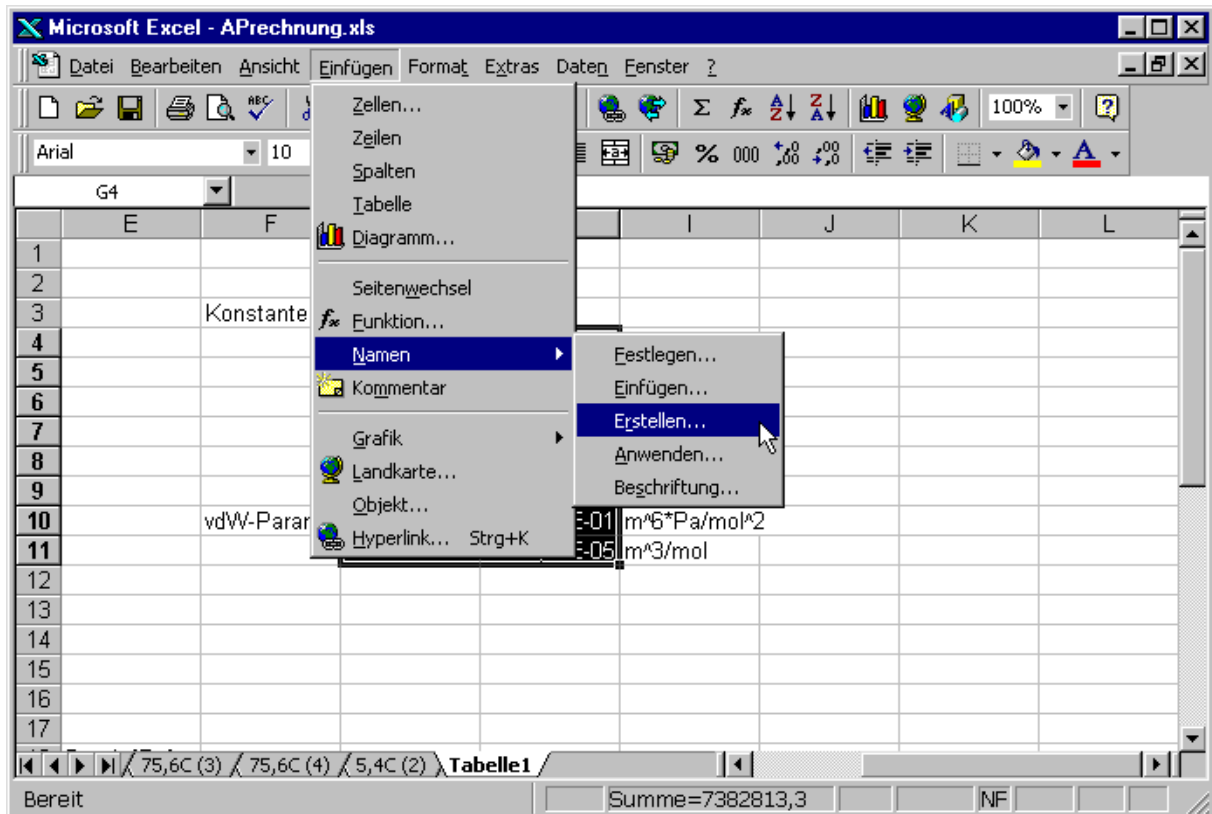
75,6C (3) 75,6C (4) 5,4C (2) Tabelle1

Bereit

Die Zellen F3 bis I11 enthalten die Literaturwerte. Wir markieren nun den Bereich G4 bis H11.

Mit dem Befehl Einfügen - Namen - Erstellen... werden die Bezeichnungen der Namen festgelegt.





Wir wählen nun die Funktion aus linker Spalte aus.



Damit werden die Namen erzeugt, die dann als absolute Referenz in Formeln verwendet werden.

In die Zelle B16 wird die Temperatur des Gases eingegeben. Beachten Sie, dass mit absoluten Temperaturen gerechnet werden muss. Die Umrechnung kann, wie hier gezeigt, direkt in Excel durchgeführt werden.

Microsoft Excel - APrechnung.xls

File Edit View Insert Format Extras Data Window ?

Arial 10 F K U

20,25

	A	B	C	D	E	F	G
18	Mikrometerschraube [mm]	Druck am Messgerät [mV]					
19	20,25	92,9					
20	20,3	85					
21	20,35	78,5					
22	20,4	72,7					
23	20,5	63,3					
24	20,58	57,3					
25	20,75	48					
26	21	38,8					
27	21,5	27,5					
28	22	21,3					
29	22,5	17,3					
30	23	14,5					
31	24	11,1					
32	25	8,9					
33	30	4,7					
34	35	3,2					

75,6C (3) 75,6C (4) 5,4C (2) Tabelle1

Bearbeiten NF

Als nächstes geben wir die Messwerte ein, für die Mikrometerschraube in mm und für den Druck in mV. Dies sind die Einheiten, die wir ablesen. Es ist eine schlechte Praxis, schon vor dem Notieren der Werte Umrechnungen durchzuführen. Wenn bei diesen Umrechnungen nämlich ein Fehler passiert, dann ist er nicht mehr auszumergen, Wenn die Umrechnungen erst im nachhinein geschehen, haben solche Fehler, wenn sie entdeckt werden, keine Konsequenz.

Microsoft Excel - APrechnung.xls

ARCTAN  $\times$   $\checkmark$   $=$   $=(50-A19)*0,0000001539$

	A	B	C	D	E	F
18	Mikrometerschraube [mm]	Druck am Messgerät [mV]	Verringerung			
19	20,25	92,9	00001539			
20	20,3	85				
21	20,35	78,5				
22	20,4	72,7				
23	20,5	63,3				
24	20,58	57,3				
25	20,75	48				
26	21	38,8				
27	21,5	27,5				
28	22	21,3				
29	22,5	17,3				
30	23	14,5				
31	24	11,1				
32	25	8,9				
33	30	4,7				
34	35	3,2				

Bearbeiten

Als nächstes berechnen wir in der Spalte C die Volumenverringering. Die Formel ist abgeleitet aus der in Zelle B4. Dabei wird zuerst das Volumen  $V_0$  außer Acht gelassen, also nur  $\Delta V = (50-x)*1.539e-7$  ausgerechnet.

Microsoft Excel - APrechnung.xls

C19  $=$   $=(50-A19)*0,0000001539$

	A	B	C	D	E	F
18	Mikrometerschraube [mm]	Druck am Messgerät [mV]	Verringerung			
19	20,25	92,9	4,579E-06			
20	20,3	85	4,571E-06			
21	20,35	78,5	4,563E-06			
22	20,4	72,7	4,555E-06			
23	20,5	63,3	4,540E-06			
24	20,58	57,3	4,528E-06			
25	20,75	48	4,502E-06			
26	21	38,8	4,463E-06			
27	21,5	27,5	4,386E-06			
28	22	21,3	4,309E-06			
29	22,5	17,3	4,232E-06			
30	23	14,5	4,155E-06			
31	24	11,1	4,001E-06			
32	25	8,9	3,848E-06			
33	30	4,7	3,078E-06			
34	35	3,2	2,309E-06			

Bereit

Summe=6,662E-05

Wir markieren nun den Bereich, in den die Formel kopiert werden sollte. Mit Strg-U füllen wir den Bereich.

In der Spalte D wird nun das Volumen ausgerechnet. In der Formel  $=\$B\$12-C19$  bedeutet die Zellenreferenz  $\$B\$12$  das Volumen  $V_0$ .  $\$B$  bedeutet, dass beim Kopieren sich diese

Microsoft Excel - APrechnung.xls

ARCTAN  $\times$   $\checkmark$  = =B\$12-C19

	A	B	C	D	E	F
18	Mikrometerschraube [mm]	Druck am Messgerät [mV]	Verringerung	Volumen [m³]		
19	20,25	92,9	4,579E-06	=B\$12-C19		
20	20,3	85	4,571E-06			
21	20,35	78,5	4,563E-06			
22	20,4	72,7	4,555E-06			
23	20,5	63,3	4,540E-06			
24	20,58	57,3	4,528E-06			
25	20,75	48	4,502E-06			
26	21	38,8	4,463E-06			
27	21,5	27,5	4,386E-06			
28	22	21,3	4,309E-06			
29	22,5	17,3	4,232E-06			
30	23	14,5	4,155E-06			
31	24	11,1	4,001E-06			
32	25	8,9	3,848E-06			
33	30	4,7	3,078E-06			
34	35	3,2	2,309E-06			

Spaltenreferenz nicht verändern sollte. \$12 bedeutet, dass die Zeilenreferenz beim Kopieren nicht verändert werden soll. \$B\$12 bedeutet also eine absolute Referenz. C19 sagt, dass beim kopieren die Zellenreferenz relativ aufzufassen ist. Man kann auch Kombinationen verwenden, bei der nur die Spalten oder nur die Zeilen absolute Referenzen sind.

Microsoft Excel - APrechnung.xls

ARCTAN  $\times$   $\checkmark$  = =B19/100\*150\*100000

	A	B	C	D	E	F
18	Mikrometerschraube [mm]	Druck am Messgerät [mV]	Verringerung	Volumen [m³]	Druck [Pa]	
19	20,25	92,9	4,579E-06	9,954E-04	=B19/100*150	
20	20,3	85	4,571E-06	9,954E-04		
21	20,35	78,5	4,563E-06	9,954E-04		
22	20,4	72,7	4,555E-06	9,954E-04		
23	20,5	63,3	4,540E-06	9,955E-04		
24	20,58	57,3	4,528E-06	9,955E-04		
25	20,75	48	4,502E-06	9,955E-04		
26	21	38,8	4,463E-06	9,955E-04		
27	21,5	27,5	4,386E-06	9,956E-04		
28	22	21,3	4,309E-06	9,957E-04		
29	22,5	17,3	4,232E-06	9,958E-04		
30	23	14,5	4,155E-06	9,958E-04		
31	24	11,1	4,001E-06	9,960E-04		
32	25	8,9	3,848E-06	9,962E-04		
33	30	4,7	3,078E-06	9,969E-04		
34	35	3,2	2,309E-06	9,977E-04		

Die Spalte E enthält den in Pascal umgerechneten Druck. Dabei wird die in der Zelle B5 angegebene Umrechnung verwendet.

Microsoft Excel - APrechnung.xls

Formelzeile: F19 =  $=\$B\$13*R*Temperatur/(D19-\$B\$13*b)-a*\$B\$13^2/D19^2$

	A	B	C	D	Formeln	F
18	Mikrometerschraube [mm]	Druck am Messgerät [mV]	Verringerung	Volumen [m <sup>3</sup> ]	Druck [Pa]	Berechneter D
19	20,25	92,9	4,579E-06	9,954E-04	1,394E+07	2,326E+03
20	20,3	85	4,571E-06	9,954E-04	1,275E+07	
21	20,35	78,5	4,563E-06	9,954E-04	1,178E+07	
22	20,4	72,7	4,555E-06	9,954E-04	1,091E+07	
23	20,5	63,3	4,540E-06	9,955E-04	9,495E+06	
24	20,58	57,3	4,528E-06	9,955E-04	8,595E+06	
25	20,75	48	4,502E-06	9,955E-04	7,200E+06	
26	21	38,8	4,463E-06	9,955E-04	5,820E+06	
27	21,5	27,5	4,386E-06	9,956E-04	4,125E+06	
28	22	21,3	4,309E-06	9,957E-04	3,195E+06	
29	22,5	17,3	4,232E-06	9,958E-04	2,595E+06	
30	23	14,5	4,155E-06	9,958E-04	2,175E+06	
31	24	11,1	4,001E-06	9,960E-04	1,665E+06	
32	25	8,9	3,848E-06	9,962E-04	1,335E+06	
33	30	4,7	3,078E-06	9,969E-04	7,050E+05	
34	35	3,2	2,309E-06	9,977E-04	4,800E+05	

In der Spalte F wird nun der Druck mit der van der Waals Gas-Theorie ausgerechnet.  $\$B\$13$  ist die zu berechnende Molzahl, D19 das Volumen, das implizit das Volumen  $V_0$  enthält.

Microsoft Excel - APrechnung.xls

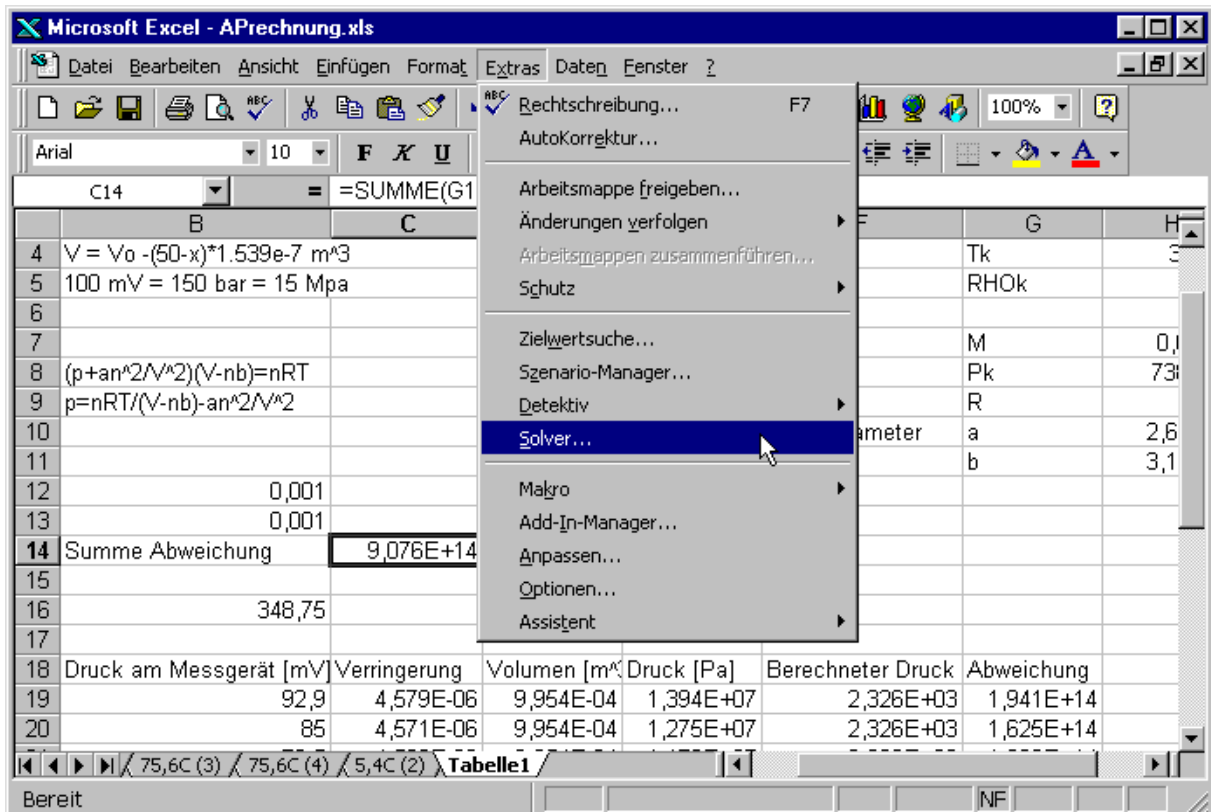
Formelzeile: ARCTAN =  $=ABS(F19-E19)^2$

	B	C	D	E	F	G	H
18	Druck am Messgerät [mV]	Verringerung	Volumen [m <sup>3</sup> ]	Druck [Pa]	Berechneter Druck	Abweichung	
19	92,9	4,579E-06	9,954E-04	1,394E+07	2,326E+03	$=ABS(F19-E19)^2$	
20	85	4,571E-06	9,954E-04	1,275E+07	2,326E+03		
21	78,5	4,563E-06	9,954E-04	1,178E+07	2,326E+03		
22	72,7	4,555E-06	9,954E-04	1,091E+07	2,326E+03		
23	63,3	4,540E-06	9,955E-04	9,495E+06	2,326E+03		
24	57,3	4,528E-06	9,955E-04	8,595E+06	2,326E+03		
25	48	4,502E-06	9,955E-04	7,200E+06	2,326E+03		
26	38,8	4,463E-06	9,955E-04	5,820E+06	2,326E+03		
27	27,5	4,386E-06	9,956E-04	4,125E+06	2,326E+03		
28	21,3	4,309E-06	9,957E-04	3,195E+06	2,326E+03		
29	17,3	4,232E-06	9,958E-04	2,595E+06	2,326E+03		
30	14,5	4,155E-06	9,958E-04	2,175E+06	2,325E+03		
31	11,1	4,001E-06	9,960E-04	1,665E+06	2,325E+03		
32	8,9	3,848E-06	9,962E-04	1,335E+06	2,325E+03		
33	4,7	3,078E-06	9,969E-04	7,050E+05	2,323E+03		
34	3,2	2,309E-06	9,977E-04	4,800E+05	2,321E+03		

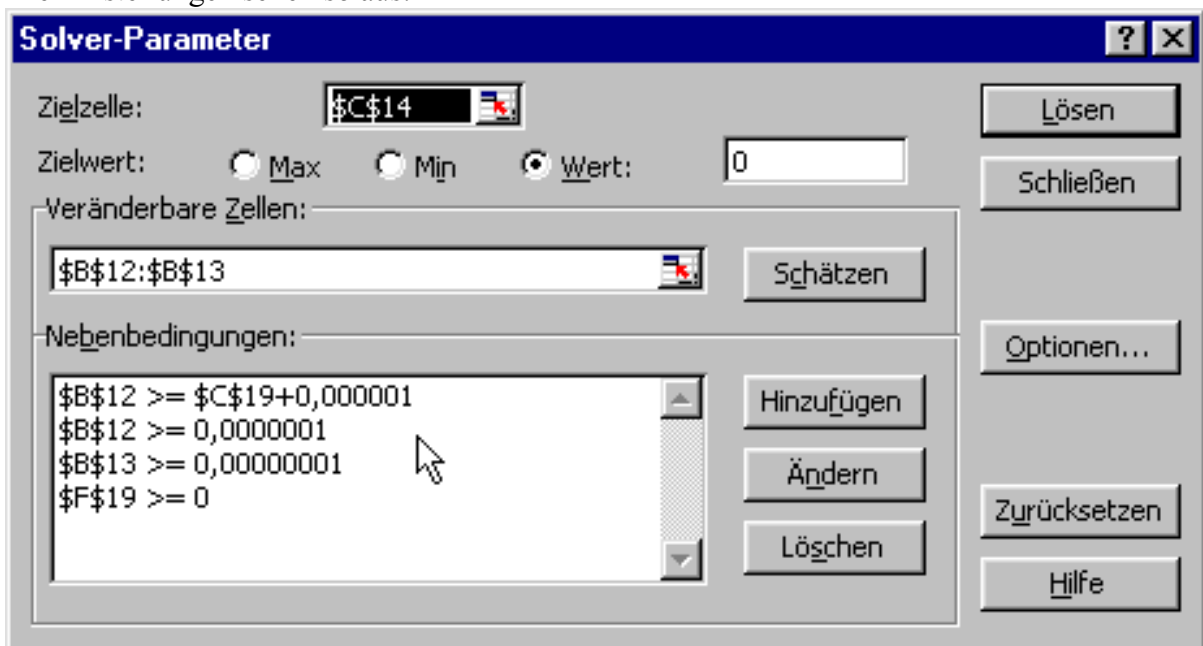
Die Spalte G enthält nun die Berechnung des Fehlers. Hier verwenden wir das Fehlerquadrat.

In der Zelle B14 wird nun die Summe aller Fehler eingetragen.

Nun können von Hand die freien Parameter verändert werden. Das erlaubt einem, den realistischen Parameterbereich abzuschätzen. Zur automatischen Abschätzung benötigen wir den Solver.



Die Einstellungen sehen so aus:



Die Zielzelle enthält die Referenz auf die Zelle, die minimiert werden soll. Der Solver funktioniert besser, wenn nicht nach einem Minimum, das ja auch lokal sein kann, sondern nach dem Zielwert 0 gesucht wird. Veränderbare Zellen sind die Molzahl und das Volumen  $V_0$ . Die Nebenbedingungen verhindern, dass bei der Suche durch 0 geteilt wird. Unter den Optionen kann der Suchalgorithmus bestimmt werden.

Nach dem Starten des Lösungsprozesses läuft die Suche. Ist sie fündig geworden, wird der folgende Dialog angezeigt.

Wir sagen OK und unsere Datei sieht nun so aus:

Microsoft Excel - APrechnung.xls

File Edit View Insert Format Extras Data Window ?

Formula bar: C14 = =SUMME(G19:G34)

	A	B	C	D	E	F
4	Volumen	$V = V_0 - (50-x) \cdot 1.539e-7 \text{ m}^3$				
5	Druck	100 mV = 150 bar = 15 Mpa				
6						
7						
8	Reales Gas	$(p + a \cdot n^2 / V^2)(V - nb) = nRT$				
9		$p = nRT / (V - nb) - a \cdot n^2 / V^2$				
10						vdW-Paramet
11	Unbekannte					
12	V0	4,82483E-06		m³		
13	n=	0,001797963		mol		
14		Summe Abweichung	1,418E+14			
15						
16	Temperatur	348,75		K		
17						
18	Mikrometerschraube [mm]	Druck am Messgerät [mV]	Verringerung	Volumen [m³]	Druck [Pa]	Berechneter [
19	20,25	92,9	4,579E-06	2,463E-07	1,394E+07	7,631
20	20,3	85	4,571E-06	2,540E-07	1,275E+07	7,630

Sheet: 75,6C (2) / 75,6C (3) / 75,6C (4) / 5,4C (2) / Tabelle1

Beachten Sie die Zielzelle, und die beiden veränderbaren Zellen. Sie zeigen nun das Resultat des Fits.

Es hilft oft, sich eine graphische Darstellung des Fits anzusehen. Dazu markieren wir die Spalte mit dem Volumen sowie die beiden Druckspalten.

Microsoft Excel - APrechnung.xls

File Edit View Insert Format Extras Data Window ?

Formula bar: D18 = Volumen [m³]

	B	C	D	E	F	G	H
18	Druck am Messgerät [mV]	Verringerung	Volumen [m³]	Druck [Pa]	Berechneter Druck	Abweichung	
19	92,9	4,579E-06	2,463E-07	1,394E+07	7,631E+06	3,974E+13	
20	85	4,571E-06	2,540E-07	1,275E+07	7,630E+06	2,622E+13	
21	78,5	4,563E-06	2,617E-07	1,178E+07	7,619E+06	1,728E+13	
22	72,7	4,555E-06	2,694E-07	1,091E+07	7,599E+06	1,093E+13	
23	63,3	4,540E-06	2,848E-07	9,495E+06	7,538E+06	3,832E+12	
24	57,3	4,528E-06	2,971E-07	8,595E+06	7,473E+06	1,260E+12	
25	48	4,502E-06	3,233E-07	7,200E+06	7,302E+06	1,049E+10	
26	38,8	4,463E-06	3,617E-07	5,820E+06	7,008E+06	1,411E+12	
27	27,5	4,386E-06	4,387E-07	4,125E+06	6,384E+06	5,104E+12	
28	21,3	4,309E-06	5,156E-07	3,195E+06	5,804E+06	6,806E+12	
29	17,3	4,232E-06	5,926E-07	2,595E+06	5,295E+06	7,290E+12	
30	14,5	4,155E-06	6,695E-07	2,175E+06	4,856E+06	7,186E+12	
31	11,1	4,001E-06	8,234E-07	1,665E+06	4,149E+06	6,170E+12	
32	8,9	3,848E-06	9,773E-07	1,335E+06	3,613E+06	5,189E+12	
33	4,7	3,078E-06	1,747E-06	7,050E+05	2,179E+06	2,172E+12	
34	3,2	2,309E-06	2,516E-06	4,800E+05	1,556E+06	1,157E+12	

Sheet: 75,6C (2) / 75,6C (3) / 75,6C (4) / 5,4C (2) / Tabelle1

Status bar: Summe=190383128

Wir drücken nun auf den Knopf für den Diagrammassistenten.

Nun wählen wir

**Diagramm-Assistent - Schritt 1 von 4 - Diagrammtyp**

Standardtypen | Benutzerdefinierte Typen

Diagrammtyp: Säule, Balken, Linie, Kreis, Punkt (XY), Fläche, Ring, Netz, Oberfläche, Blase, Kurs

Untertyp: [Vier Diagramm-Vorschau-Optionen]

Punkte mit Linien ohne Datenpunkte

Schaltfläche gedrückt halten für Beispiel

Abbrechen < Zurück Weiter > Ende

F	G	H
eter Druck	Abweichung	
7,631E+06	3,974E+13	
7,630E+06	2,622E+13	
7,619E+06	1,728E+13	
7,599E+06	1,093E+13	
7,538E+06	3,832E+12	
7,473E+06	1,260E+12	
7,302E+06	1,049E+10	
7,008E+06	1,411E+12	
6,384E+06	5,104E+12	
5,804E+06	6,806E+12	
5,295E+06	7,290E+12	
4,856E+06	7,186E+12	
4,149E+06	6,170E+12	
3,613E+06	5,189E+12	
2,179E+06	2,172E+12	
1,556E+06	1,157E+12	

Ein Liniendiagramm ohne Datenpunkte aus und drücken Weiter.

**Diagramm-Assistent - Schritt 2 von 4 - Diagramm-Quelldaten**

Datenbereich | Reihe

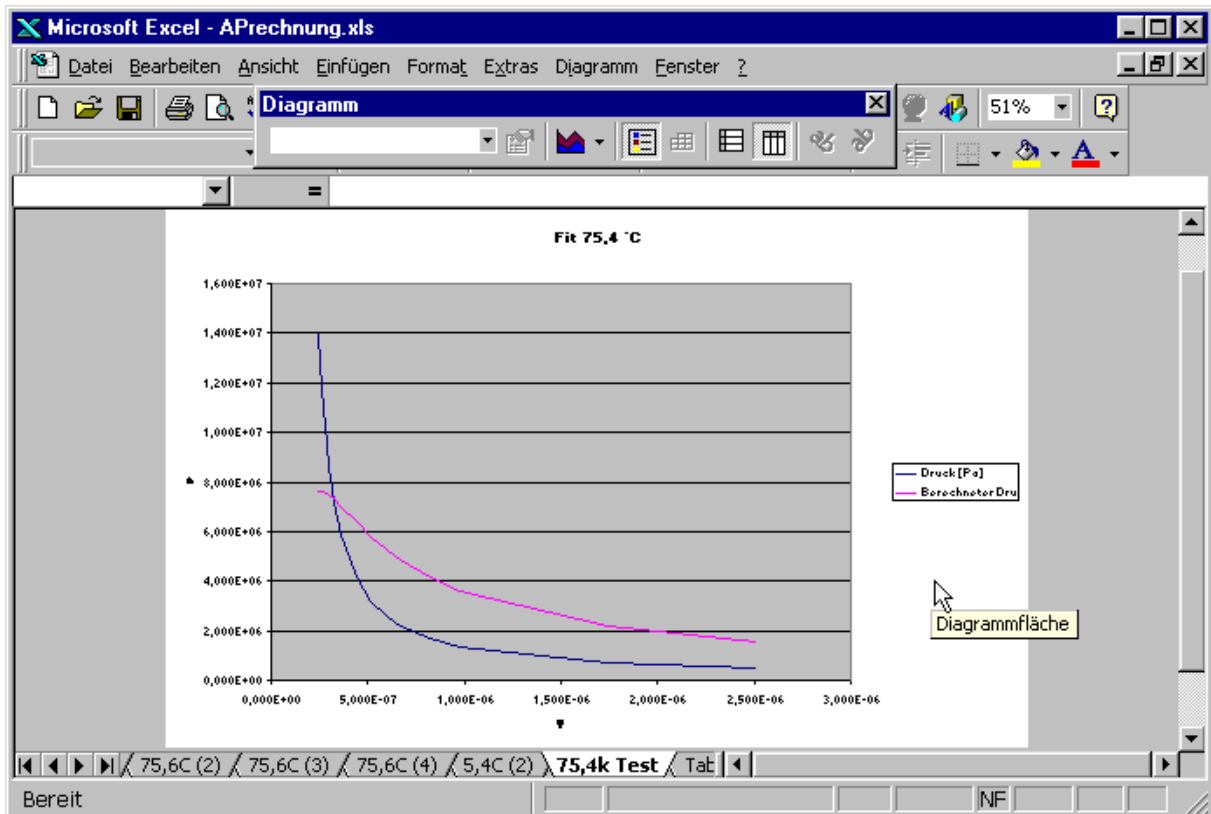
Datenbereich:

Reihe in:  Zeilen  Spalten

Abbrechen < Zurück Weiter > Ende

Wir beschriften im nächsten Dialog das Diagramm  
Und fügen es als neue Seite ein.





Sie sehen, dass der Fit nicht sehr gut ist.

Ein Grund ist das  $1/x$ -Verhalten der Kurve. Wir glätten das Ganze, indem wir logarithmieren.

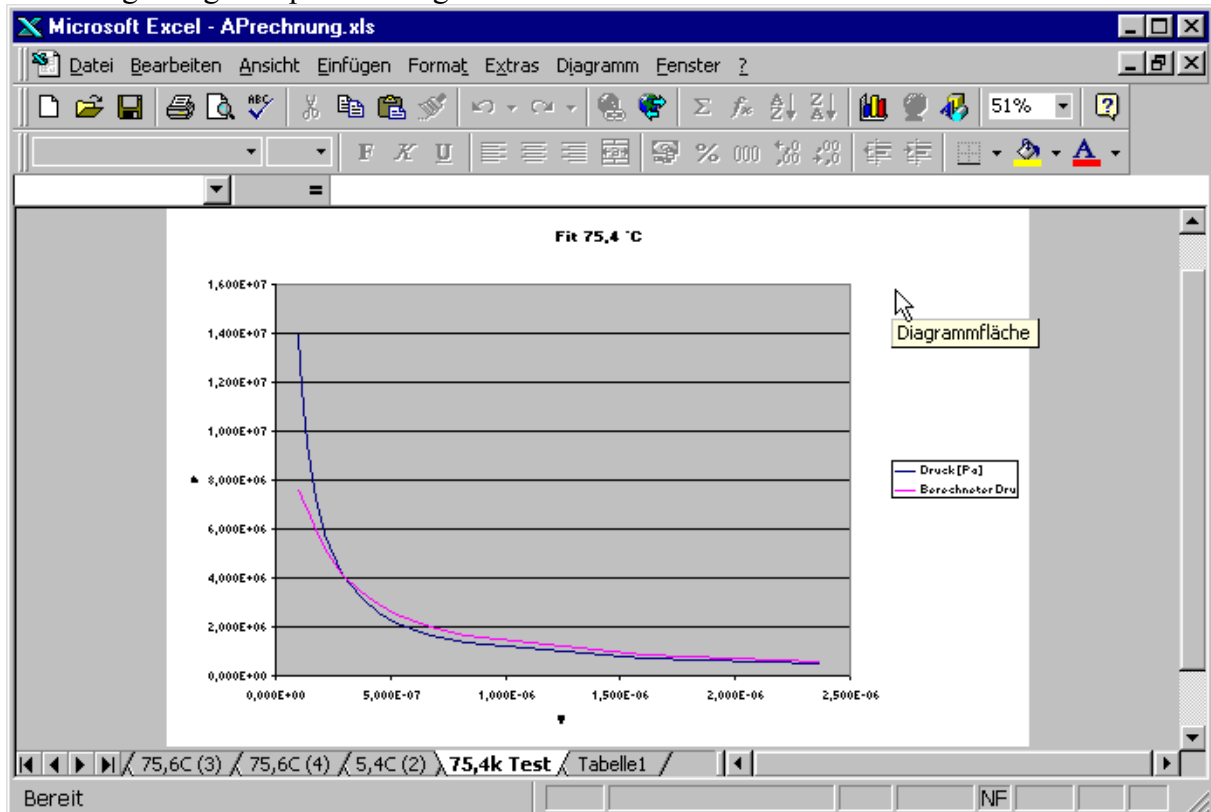
	B	C	D	E	F	G	H
4	$V = V_0 - (50-x) \cdot 1.539e-7 \text{ m}^3$					Tk	3
5	100 mV = 150 bar = 15 Mpa					RHok	
6							
7						M	0,1
8	$(p+an^2/V^2)(V-nb)=nRT$					Pk	73
9	$p=nRT/(V-nb)-an^2/V^2$					R	
10					vdW-Parameter	a	2,6
11						b	3,1
12	4,82483E-06		m <sup>3</sup>				
13	0,001797963		mol				
14	Summe Abweichung	1,020E+14					
15							
16	348,75		K				
17							
18	Druck am Messgerät [mV]	Verringerung	Volumen [m <sup>3</sup> ]	Druck [Pa]	Berechneter Druck	Abweichung	
19	92,9	4,579E-06	2,463E-07	1,394E+07	7,631E+06	=ABS(LN(F19	
20	85	4,571E-06	2,540E-07	1,275E+07	7,630E+06	2,622E+13	

Nun fitten wir wieder und erhalten

Unbekannte

$V_0$  4,67852E-06 m<sup>3</sup>  
 $n$  0,000633095 mol  
 Summe Abweichung 1,558E+00

Der dazugehörige Graph sieht folgendermaßen aus:



Spielen Sie mit der beiliegenden Datei APrechnung.xls. Sie enthält die oben gezeigten Rechnungen.

## Literatur

1. H. Gränicher, Grundlagen physikalischer Messungen, VDF Zürich
2. Philipp Bevington, Data Reduction and Error Analysis for the Physical Sciences (Mc Graw Hill, New York 1969).
3. R. Zurmühl, Praktische Mathematik für Ingenieure und Physiker, Springer-Verlag 1963, p 290 ff.
4. Statistische Datenauswertung und Fehlerrechnung (<http://www.we.fh-osnabrueck.de/fbwe/vorlesung/edv1/statx.html>), E.D. Schmitter, Fachhochschule Osnabrück (Postscript-Datei, <http://www.we.fh-osnabrueck.de/fbwe/vorlesung/edv1/statx.ps>)
5. Rechnergestützte Analyse von Daten (<http://www.we.fh-osnabrueck.de/fbwe/vorlesung/edv2/skript4.html>), E.D. Schmitter, Fachhochschule Osnabrück (Postscript-Datei, <http://www.we.fh-osnabrueck.de/fbwe/vorlesung/edv2/skript4.ps>)
6. Nichtlineare Regression mit neuronalen Netzen (<http://www.we.fh-osnabrueck.de/fbwe/vorlesung/edv2/nn/nn.html>), E.D. Schmitter, Fachhochschule Osnabrück (Postscript-Datei, <http://www.we.fh-osnabrueck.de/fbwe/vorlesung/edv2/nn/nn.ps>)