

Fehlerrechnung und –schätzung, lineare und nichtlineare Regression

Inhalt

- 0 Einleitung
- 1 Wahrer Wert, Erwartungswert, Messwert und Fehler
- 2 Einzelmessungen
- 3 Wiederholte (viele) Messungen
 - 3.1 Gaußsche Normalverteilung
 - 3.2 Vertrauensbereich = Konfidenzintervall
 - 3.3 Fehler bei unbekannter Streuung → „t-Faktoren“
- 4 (Viele) Messungen in Abhängigkeit einer Variablen: Regressionsanalyse
- 5 Gewichteter Mittelwert
- 6 Fehlerrechnung in der Kernphysik
- 7 Darstellung des Ergebnisses

- 8 Beispiele
 - 8.1 Beispiel: Fehlerfortpflanzung
 - 8.2 Beispiel: Statistischer Fehler bei definiertem Vertrauensniveau
 - 8.3 Beispiele: Regressionsanalyse und Anfitten von Messdaten
 - 8.3.1 Lineare Regression „von Hand“
 - 8.3.2 Lineare Regression mit Excel
 - 8.3.3 Lineare Regression mit Excel nach Linearisierung nichtlinearer Zusammenhänge
 - 8.3.4 Lineare Regression mit qtiplot mit individuellen Fehlern
 - 8.3.5 Nichtlineare Regression mit qtiplot
- 9 Im Script verwendete Notation

0 Einleitung

Physikalische Messungen sind *immer* fehlerbehaftet, weshalb es notwendig ist, den jeweiligen Fehler zumindest grob abzuschätzen und diesen auch anzugeben. Dies ist nicht immer einfach, die wichtigsten Begriffe und Verfahren sollen daher hier kurz wiederholt werden.

1 Wahrer Wert, Erwartungswert, Messwert und Fehler

Im einfachsten Fall entspricht der Fehler Δx einer Messung der Differenz von dem wahren Wert μ (= Erwartungswert falls nur statistische Fehler) und dem Messwert x . Der Begriff „Fehler“ wird aber in der Fehlerrechnung eher im Sinne von „max. möglichem Fehler“ oder „sehr wahrscheinlich max. Fehler“ verwendet.

Wir definieren daher etwas vorsichtiger:

$$|\text{Wahrer Wert } \mu - \text{Messwert } x| = \text{Abweichung} \leq \text{Fehler } \Delta x \quad (\text{bzw. } \Delta x_{\max}) \quad (1)$$

Die Angabe eines Messergebnisses in der Form $x \pm \Delta x$ bedeutet dann, dass der gesuchte wahre Wert μ mit einer (noch zu definierenden) hohen Wahrscheinlichkeit in dem Intervall $[x-\Delta x, x+\Delta x]$ liegt.

Dabei ist egal ob der Fehler aus Sicht des wahren Wertes oder des Messergebnisses betrachtet wird:

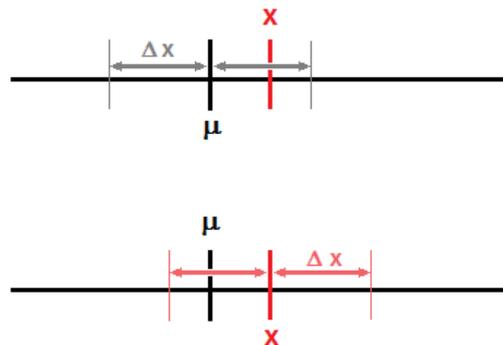


Abb. 1: Abweichung und Fehler

Die Fehlerursachen sind vielfältig, zu unterscheiden (und addieren) sind

- statistische Fehler
 - verschwinden bei Mittelung von unendlich vielen Messungen
 - sind mit statistischen Methoden berechenbar
 - Ursache z.B. „Rauschen“
- systematische Fehler
 - verschwinden *nicht* bei Mittelung von unendlich vielen Messungen
 - haben „immer die gleiche Richtung“, aber *nicht* berechenbar
 - Ursache z.B. falsche Kalibrierung Messmittel

Achtung: Die Unterscheidung ist nicht immer einfach, die Literatur teilweise sogar (anscheinend) widersprüchlich.

Bsp: Stoppuhr + 2 Personen; Viele Geräte/Bauteile Hersteller – ein Gerät/Bauteil bei Anwender; ...

2 Einzelmessungen

Es ist *keine* statistische Betrachtung möglich, der Fehler Δx kann nur *abgeschätzt* werden.

Falls die Messgröße x ungleich der gesuchten Größe $F(x)$ ist, pflanzt sich Fehler entsprechend dem funktionalen Zusammenhang fort (Fehlerfortpflanzung):

$$F + \Delta F^+ = F(x + \Delta x) ; F - \Delta F^- = F(x - \Delta x) \quad F = F(x) + \Delta F^+ / -\Delta F^- \quad (2)$$

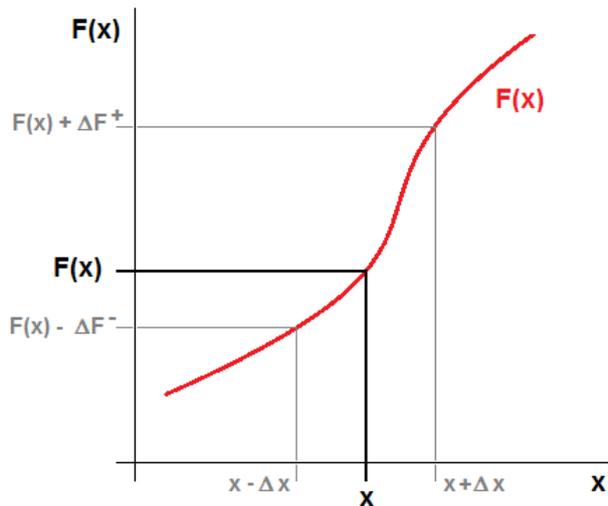


Abb. 2: Fehlerfortpflanzung bei großen Fehlern, eine Unterscheidung ΔF^+ und ΔF^- ist i.A. notwendig.

Für kleine Fehler Δx lässt sich die Funktion $F(x)$ in der Umgebung des Messwertes meist linearisieren, und es gilt $\Delta F^+ = \Delta F^- = \Delta F$ mit

$$\Delta F = \left| \frac{dF}{dx} \right| \Delta x \quad (3)$$

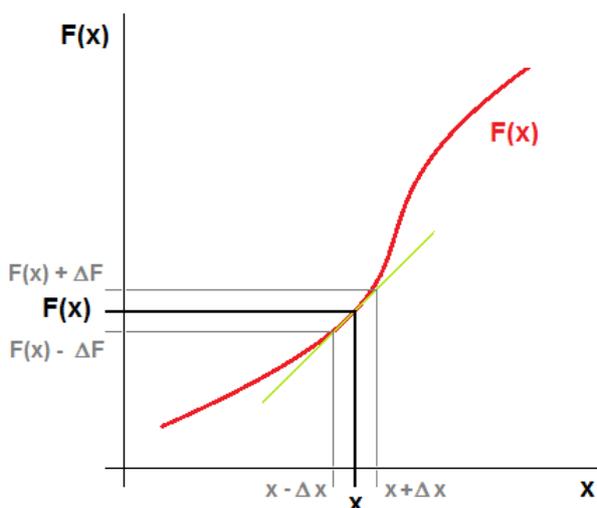


Abb. 3: Fehlerfortpflanzung bei kleinen Fehlern, die Linearisierung erlaubt die Berechnung des Fehlers ΔF mit der Ableitung der Funktion F .

Falls zudem mehrere *verschiedene* Messungen x, y, z, \dots nötig sind um die Größe $F=F(x, y, z, \dots)$ zu bestimmen, addieren sich die Fehlerfortpflanzungen (, an die Stelle der Tangente tritt hier eine Tangentialebene):

a) Maximal möglicher Fehler = Summe fortgeplanzter Einzelfehler:

$$\Delta F_{max} = \left| \frac{\partial F}{\partial x} \right| \Delta x + \left| \frac{\partial F}{\partial y} \right| \Delta y + \left| \frac{\partial F}{\partial z} \right| \Delta z + \dots \quad (4)$$

b) Wahrscheinlichster Fehler bei *unabhängigen* Messgrößen x, y, z, \dots :

$$\Delta F = \sqrt{\left(\frac{\partial F}{\partial x} \Delta x \right)^2 + \left(\frac{\partial F}{\partial y} \Delta y \right)^2 + \left(\frac{\partial F}{\partial z} \Delta z \right)^2 + \dots} \quad (5)$$

(sog. *Gaußsche Fehlerfortpflanzung*: Es ist unwahrscheinlich, dass die Fehler der *unabhängigen* Messgrößen F alle „in die gleiche Richtung“ verfälschen.)

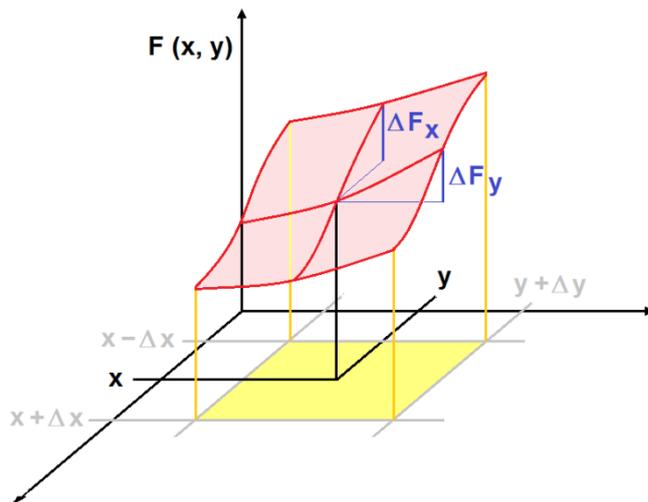


Abb. 4: Fehlerfortpflanzung im Fall von zwei unabhängigen Messgrößen. Bei kleinen Fehlern entspricht die dargestellte Fläche einer Tangentialebene und ΔF ergibt sich als Summe von ΔF_x und ΔF_y .

Speziell gilt damit für

a) Summen $F = ax + by + cz + \dots$ die Fehler addieren sich:

$$\Delta F_{max} = |a \cdot \Delta x| + |b \cdot \Delta y| + |c \cdot \Delta z| + \dots \quad (6)$$

$$\Delta F = \sqrt{(a \cdot \Delta x)^2 + (b \cdot \Delta y)^2 + (c \cdot \Delta z)^2 + \dots} \quad (7)$$

b) Produkte $F = x^a \cdot y^b \cdot z^c \dots$ die *relativen* Fehler addieren sich:

$$\frac{\Delta F_{max}}{F} = \left| a \cdot \frac{\Delta x}{x} \right| + \left| b \cdot \frac{\Delta y}{y} \right| + \left| c \cdot \frac{\Delta z}{z} \right| + \dots \quad (8)$$

$$\frac{\Delta F}{F} = \sqrt{\left(a \cdot \frac{\Delta x}{x} \right)^2 + \left(b \cdot \frac{\Delta y}{y} \right)^2 + \left(c \cdot \frac{\Delta z}{z} \right)^2 + \dots} \quad (9)$$

3 Wiederholte (viele) Messungen

Hier ist eine Statistik möglich, insbesondere ist

- eine Mittelwertbildung sinnvoll
- der Fehler der Einzelmessung berechenbar (→ Standardabweichung).

Zu unterscheiden ist der

- Fehler der Einzelmessung Δx , und
- Fehler des Mittelwertes $\Delta \bar{x}$ (wird mit Anzahl n der Messungen kleiner!)

Zu beachten:

- Statistische Streuung („Fehler“) bedingt Breite einer Häufigkeitsverteilung
- Definition eines Vertrauensbereichs erlaubt umgekehrt Angabe eines Fehlers auf Basis der Breite der gemessenen Häufigkeitsverteilung

3.1 Gaußsche Normalverteilung

Streuen Messwerte rein zufällig (stochastisch) um den wahren Wert, so nähert sich die gemessene relative Häufigkeitsverteilung mit der Anzahl von Einzelmessungen einer Gaußverteilung $g(x)$ an. Für *unendlich* viele Messungen entspricht der Mittelwert dem wahren Wert. Wie groß ist der verbleibende Fehler bei endlich vielen Messungen?

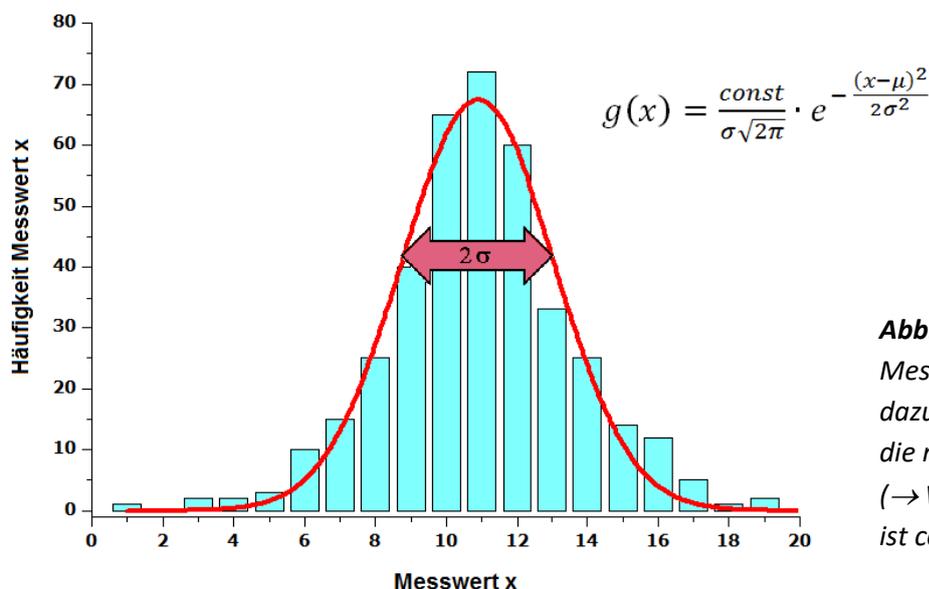


Abb. 5: Häufigkeitsverteilung der Messung eines Wertes x und die dazu angepasste Gaußverteilung. Für die normierte Gaußverteilung (→ Wahrscheinlichkeitsverteilung) ist $const = 1$.

Für unendlich viele Messungen entspricht σ der Standardabweichung s , bei endlich vielen Messungen ist s zumindest eine gute Schätzung für σ .

Je besser die Qualität der Einzelmessungen, desto kleiner wird die Streuung der einzelnen Messwerte also die Standardabweichung s . Für den statistischen Standardfehler δx gilt

$$\delta x = s = \sqrt{\frac{1}{n-1} \cdot \sum_i^n (x_i - \bar{x})^2} \quad (10)$$

Je höher die Anzahl der Messungen, desto genauer wird der daraus bestimmte Mittelwert. Für den Standardfehler des Mittelwertes $\delta \bar{x}$ ergibt sich aus der Fehlertheorie

$$\delta \bar{x} = \frac{s}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \cdot \sum_i^n (x_i - \bar{x})^2} \quad (11)$$

Bei unendlich vielen Messungen wäre der Fehler des Mittelwertes also gleich null. Sinn macht die Fehlerrechnung also nur bei endlich vielen Messungen. Leider ist dann die Standardabweichung s auch nur eine mäßig gute Schätzung für σ und der Fehler $\Delta \bar{x}$ ist i.A. größer als der Standardfehler $\delta \bar{x}$. Für eine genauere statistische Fehlerabschätzung (bzw. –rechnung) bzgl. des Mittelwertes müssen letztlich sog. Vertrauensbereiche definiert und weitere Korrekturen berücksichtigt werden. →

3.2 Vertrauensbereich = Konfidenzintervall

Annahme: Streuung σ (und wahrer Wert μ) seien bekannt.

Da die normierte Gaußverteilung $G(x)$ eine Wahrscheinlichkeitsverteilung ist, ergibt das Integral über die gesamte Verteilung = 1, und über einen Wertebereich die Wahrscheinlichkeit, dass ein Messwert in genau diesen Bereich fällt.

Bsp: Das Integral von $\mu - 2\sigma$ bis $\mu + 2\sigma$ ergibt 0,9545 also gut 95 %. Die Wahrscheinlichkeit, dass der Wert *einer* Messung in dem Intervall $[\mu - 2\sigma, \mu + 2\sigma]$ liegt, ist daher gut 95%.

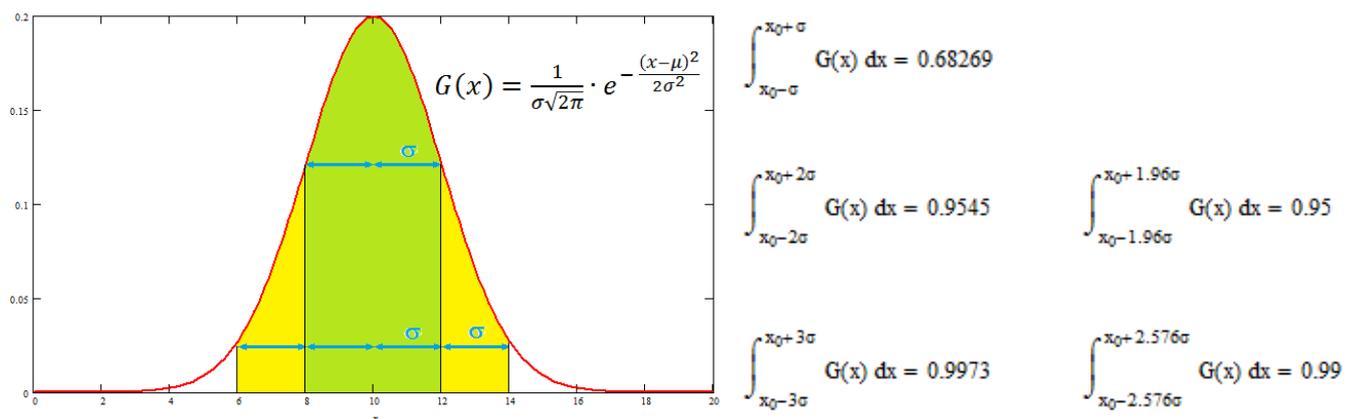


Abb. 6: Zur Definition des Konfidenzintervalls: Die Integration über die normierte Gaußverteilung ergibt für das 1-Sigma-Intervall (grün) eine Wahrscheinlichkeit von ca. 68 %, für das 2-Sigma-Intervall (grün + gelb) eine Wahrscheinlichkeit von gut 95 %. (Die Ergebnisse gelten unabhängig von den in der Grafik gewählten beispielhaften Werten von $\mu = 10$ und $\sigma = 2$.)

Der Fehler des Mittelwertes nimmt mit dem Faktor $\frac{1}{\sqrt{n}}$ ab. Daher ist die Wahrscheinlichkeit, dass der aus n Messungen bestimmte Mittelwert in dem Intervall $[\mu - \frac{2\sigma}{\sqrt{n}}, \mu + \frac{2\sigma}{\sqrt{n}}]$ liegt, auch entsprechend bei gut 95%.

Umgekehrt heißt das (vgl. Kap.1), dass der wahre Wert μ mit einer Wahrscheinlichkeit von gut 95% im Intervall $[\bar{x} - \frac{2\sigma}{\sqrt{n}}, \bar{x} + \frac{2\sigma}{\sqrt{n}}]$ liegt.

Der Fehler läge demnach mit einer statistischen Sicherheit von gut 95% bei $\Delta\bar{x} = \frac{2\sigma}{\sqrt{n}}$ (und mit einer statistischen Sicherheit von genau 95% bei $\Delta\bar{x} = \frac{1,96\sigma}{\sqrt{n}}$).

3.3 Fehler bei unbekannter Streuung → „t-Faktoren“

Streuung σ (und wahrer Wert μ) sind i.A. jedoch *nicht* bekannt, sondern müssen aus einer *endlichen* Zahl von Messungen abgeschätzt werden.

In diesem Fall muss eigentlich an Stelle der Gaußverteilung eine student-t-Verteilung (nach W.S. Gosset) angesetzt werden. Für $n \rightarrow \infty$ geht diese in die Gaußverteilung über, aber für eine (kleine) endliche Zahl von Messungen ergeben sich (deutliche) Abweichungen. Diese werden bei der Berechnung von Vertrauensintervallen, und damit der Fehler, mit den sog. t-Faktoren berücksichtigt. Die t-Faktoren sind von der Anzahl der Messungen n und der gewählten statistischen Sicherheit P abhängig (vgl. Tab.1 und *tp_Faktoren.xls*).

Für den von P und n abhängigen Fehler des Mittelwertes $\Delta\bar{x}$ ergibt sich letztlich

$$\Delta\bar{x} = t_p \cdot \delta\bar{x} = t_p \cdot \frac{s}{\sqrt{n}} = t_p \cdot \sqrt{\frac{1}{n(n-1)} \cdot \sum_i^n (x_i - \bar{x})^2} , \quad (12)$$

wobei s der Standardabweichung der gemessenen Einzelwerte entspricht.

Tab. 1: Einige t_p Faktoren

Bsp:

n = 10 Messungen

95 % stat. Sicherheit

→ $t_p = 2,30$

→ $\Delta\bar{x}_{95\%,n=10} = 2,30 \cdot \frac{s}{\sqrt{10}}$

n \ P	68,3%	95%	99%
2	1,84	12,7	63,7
3	1,32	4,30	9,92
4	1,20	3,20	5,84
5	1,15	2,80	4,6
10	1,06	2,30	3,25
20	1,03	2,10	2,86
100	1,00	1,98	2,63
∞	1,00	1,96	2,58

Die t_p -Faktoren entsprechen also in etwa einem (von der Anzahl der Messungen abhängigen) „effektiven Vielfachen“ der Streuung in der Gaußverteilung, welches für die Berechnung des Vertrauensbereichs angesetzt werden muss.

4 (Viele) Messungen in Abhängigkeit einer Variablen: Regressionsanalyse

Oft wird mit einer Messung nicht nur ein Wert x oder ein Funktionswert $F(x, \dots)$ bestimmt (vgl. Kap.3), sondern die Funktion oder deren Parameter selbst. Dazu werden

1. möglichst viele Wertepaare $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ gemessen,
2. eine passende Funktion (mit zunächst noch freien Parametern) gewählt und
3. meist mit einer geeigneten Software unter Verwendung der Messdaten die Parameter bestimmt, mit denen sich die Messdaten am besten beschreiben lassen.

Die meiste Software (z.B. auch Excel, Origin, qtiplot, mathcad, ...) verwendet hierfür standardmäßig die *Methode der kleinsten Fehlerquadrate* (least squares fit) und liefert neben den gesuchten Parametern auch deren *Standardfehler*.

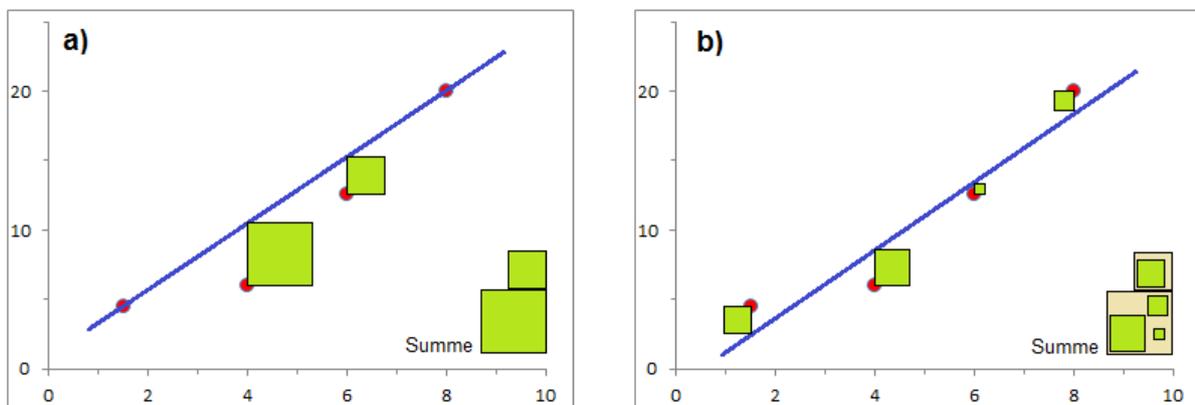


Abb. 7: Zur Methode der kleinsten Quadrate: Die Parameter der Geradengleichung werden so lange angepasst, bis die Summe der Abweichungsquadrate (grün) minimal wird. Aus der verbleibenden Summe (\rightarrow Residuen) bestimmen sich die vom Algorithmus angegebene Fehler. In Abb. b) ist die bestmögliche Anpassung dargestellt, für die in a) angenommene Gerade ist die Summe der Abweichungsquadrate offensichtlich größer.

Für den Fall eines linearen Zusammenhangs $F(x) = m \cdot x + b$ werden z.B. die Parameter m und b sowie die jeweiligen Standardfehler δm und δb bestimmt. Zudem wird der Standardfehler δF der daraus zu berechnenden Größe $F(x)$ angegeben. Der *Standardfehler* gibt (wie schon in Kap.3) an, dass der wahre Wert z.B. des Parameters m mit einer Wahrscheinlichkeit von ca. 68% im dem Intervall $[m - \Delta m, m + \Delta m]$ liegt, sofern hinreichend viele Messpunkte (> 100) bestimmt wurden.

Für ein höheres Vertrauensniveau oder/und eine geringe Anzahl von Messungen muss der Standardfehler wieder mit den entsprechenden t_p Faktoren (vgl. *tp_Faktoren_Tabelle.pdf*) multipliziert werden, wobei hier der Wert bei $f = r - 2$ zu nehmen ist. (Bsp.: $r = 5$ Wertepaare, 95 % stat. Sicherheit, $\rightarrow t_p = 3,18$)

Die oben beschriebene Bestimmung des Fehlers gilt für lineare wie nichtlineare Regressionen, wobei im Falle einer nichtlinearen Regression auch oft von *Fitten* gesprochen wird. Sind die Fehler der Einzelmessungen alle gleich groß, lässt sich eine lineare Regression sogar analytisch durchführen und die Parameter und Fehler „von Hand“ ausrechnen (vgl. Bsp. in

Kap. 8.3.1). Einfacher ist sicher die Verwendung einer geeigneten Software wie z.B. Excel. Sind die Fehler der Einzelmessungen *unterschiedlich* groß (z.B. durch Wechsel des Messbereichs oder verschiedenen Zählraten in der Kernphysik), ist es sinnvoll die Einzelmesswerte entsprechend ihren jeweiligen Fehlern unterschiedlich zu wichten (→Fehlerbalken). Hierfür benötigt man eine etwas „professionellere“ Software, wie z.B. Origin oder qtiplot, die hier entsprechend erweiterte Algorithmen verwenden. Auch in diesem Fall werden neben den bestimmten (Fit-) Parametern deren Standardfehler ausgegeben, welche ggf. wie in 3.3 beschrieben interpretiert werden müssen.

Achtung: Auch hier wird nur der *statistische* Fehler aufgrund der Streuung der einzelnen Messwerte berechnet. Ein anzunehmender *systematischer* Fehler ist ggf. zu addieren!

Anmerkung:

Auch wenn nur ein Parameter (z.B. eine Proportionalitätskonstante) zu bestimmen ist, ist eine Regression (z.B. bei vielen Messungen bei unterschiedlichen Werten x) einer einfachen Mittelwertbildung (nach ebenso vielen Messungen bei dem gleichen Wert x) u.a. aus den folgenden Gründen grundsätzlich vorzuziehen!

- Abweichungen von dem angenommenen Zusammenhang werden ggf. deutlich (z.B. Gültigkeitsbereich einer Linearisierung).
- Systematische Fehler (z.B. Kalibrierungsfehler) können ggf. erkannt werden .

5 Gewichteter Mittelwert

Haben voneinander unabhängige und meist auch unterschiedliche Messungen *verschiedene* Fehler, macht ein gewöhnlicher Mittelwert wenig Sinn. Hier sollte ein gewichteter Mittelwert gebildet werden. Als Wichtungsfaktoren w_i nimmt man bei statistischen Einzelfehlern die Kehrwerte der Fehlerquadrate, womit die vertrauenswürdigeren Ergebnisse mit den kleineren Fehlern entsprechend stärker gewichtet werden:

$$\bar{x}_{gew} = \frac{\sum_i w_i \cdot x_i}{\sum_i w_i} \quad \Delta \bar{x}_{gew} = \sqrt{\frac{1}{\sum_i w_i}} \quad w_i = \frac{1}{(\Delta x_i)^2} \quad (13)$$

Bsp: Zwei verschiedene Messungen der gleichen Größe ergeben die Werte

$X_1 = 32 \pm 0,5$
 $X_2 = 34 \pm 2$
 → $\bar{x}_{gew} = 32,1$
 → $\Delta \bar{x}_{gew} = 0,49$

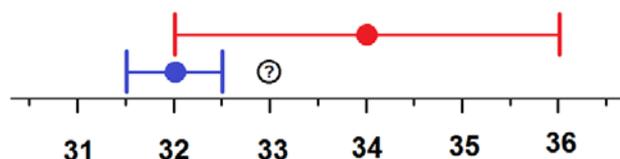


Abb. 8: Messwerte mit Fehlerbalken: Offensichtlich ist der arithmetische Mittelwert von 33 nicht sinnvoll, da er außerhalb des Fehlerbalkens der genaueren Messung liegt.

Bei nicht statistisch berechneten absoluten Einzelfehlern kann eine lineare Wichtung bei der Mittelwertbildung gemäß $w_i = \Delta x_i^{-1}$ evtl. sinnvoller sein.

6 Fehlerrechnung in der Kernphysik

Die Messung sehr vieler (N) Einzelereignisse, also die Messung einer Zählrate $z = \frac{N}{t}$ über eine lange Zeit reduziert den Fehler in z mit einem Faktor $\frac{1}{\sqrt{N}}$, da entsprechend mehr Einzelereignisse gemessen werden. Wie ist aber im Falle der Messung einer Zählrate der absolute Fehler anzusetzen, und wie groß sind die Fehlerbalken bei einer grafischen Darstellung zu wählen?

Da die Zerfallswahrscheinlichkeit eines einzelnen Atomkerns sehr klein ist, bzw. die Zahl von Kernen sehr groß ist gegen die Zahl von gemessenen Zerfällen, kann die Standardabweichung hier über einen „Umweg über die Poisson-Verteilung“ abgeschätzt werden. Eine etwas schwierige Herleitung ergibt letztlich das einfache Ergebnis:

$$\delta N = \sqrt{N} \quad \text{bzw.} \quad \delta z = \frac{\delta N}{t} = \frac{\sqrt{N}}{t}. \quad (14)$$

Der Fehler ΔN oder Δz (\rightarrow Fehlerbalken) ergibt sich dann entsprechend dem angesetzten Vertrauensniveau mit $\Delta N = t_p \cdot \delta N$ bzw. $\Delta z = t_p \cdot \delta z$.

Für eine Sicherheit von z.B. 95% ist $\Delta N \cong 2 \cdot \delta N$ bzw. $\Delta z \cong 2 \cdot \delta z$, da hier die Anzahl der Messungen der Zahl der Einzelereignisse N entspricht und i.d.R. sehr groß ist.

\rightarrow Bei gleicher Aktivität wird also der Fehler Δz mit $\frac{1}{\sqrt{N}}$ kleiner, da N linear mit der Messzeit t zunimmt!

\rightarrow Der absolute Fehler ΔN nimmt zwar mit N zu, der relative Fehler $\frac{\Delta N}{N}$ nimmt aber auch hier mit $\frac{1}{\sqrt{N}}$ ab.

7 Darstellung des Ergebnisses

Zu einem Messergebnis gehört auch immer der Fehler. Geben Sie daher Ihr Messergebnis abschließend immer mit Fehler an. Wird bei einer statistischen Auswertung abweichend vom Standardfehler ein höheres Vertrauensniveau als 68,3% gewählt, sollte dies ggf. angegeben werden.

Eine physikalische Größe besteht aus einer Zahl und einer Einheit. Wählen Sie sinnvolle Einheiten, d.h. gegebenenfalls Vorsätze wie k für „kilo“ oder m für „milli“ oder passende Zehnerpotenzen. Achten Sie darauf, dass zu vergleichende Werte, d.h. Ergebnis und Fehler, in gleichen Einheiten bzw. Zehnerpotenzen dargestellt werden.

Der Fehler wird üblicherweise mit ein bis zwei relevanten Stellen angegeben und grundsätzlich aufgerundet. Die Zahl der anzugebenden relevanten Stellen für das Ergebnis ist entsprechend der Größe des Fehlers zu wählen.

Als Zahl der relevanten Stellen bezeichnet man die Anzahl aussagekräftiger Ziffern ohne führende Nullen. Drei relevante Stellen haben z.B. die Zahlen 12,3 | 1,23 | 0,0123 | $1,23 \cdot 10^{-4}$.

Beispiele für eine falsche (schlechte) und richtige (sinnvolle) Angabe des Ergebnisses:

<i>falsch / schlecht</i>	richtig / sinnvoll
$g = (9,823456 \pm 0,481234) \text{ m/s}^2$ $g = (9,823456 \pm 0,48) \text{ m/s}^2$ $g = 9,8 \pm 0,5 \text{ m/s}^2$	$g = (9,8 \pm 0,5) \text{ m/s}^2$
$x = 30 \pm 22 \cdot 10^{-2}$	$x = 30,0 \pm 0,22$
$d = 0,023 \text{ m} \pm 0,976 \text{ mm}$	$d = 23,0 \text{ cm} \pm 0,1 \text{ cm}$ $d = (23,0 \pm 0,1) \text{ cm}$

Tab. 2: Zur Darstellung eines Messergebnisses

Um ganz allgemein die Qualität einer Messung zu beschreiben, ist die Angabe des relativen Fehlers $\Delta x/x$ sinnvoll, üblicher Weise in %. Für die Beispiele oben ergibt sich z.B.:

$$\frac{\Delta g}{g} = 0,0510 < 0,06 \rightarrow < 6\% \quad \text{bzw.} \quad \frac{\Delta x}{x} = 0,007333 \leq 0,0074 \rightarrow 0,8\% \quad (\text{oder } < 1\%)$$

Je nach Zusammenhang ist auch die Angabe eines Fehlerintervalls sinnvoll, also z.B.:

Der Wert für g liegt mit einer Wahrscheinlichkeit von 95 % im Intervall $[9,3 \text{ m/s}^2, 10,3 \text{ m/s}^2]$.

8 Beispiele

8.1 Beispiel: Fehlerfortpflanzung

Es wird die Geschwindigkeit v eines vorbeifahrenden Zugs bestimmt, indem mit einer Stoppuhr die Zeit $t = 2,85$ s gemessen wird, in welcher der Zug die Strecke $s = 50$ m zurücklegt. Eine Wiederholung der Messung ist nicht möglich, die Fehler der Einzelmessungen werden mit $\Delta t = 0,2$ s und $\Delta s = 0,5$ m abgeschätzt.

Entsprechend Definition der Geschwindigkeit ergibt sich:

$$v = \frac{s}{t} = \frac{50 \text{ m}}{2,85 \text{ s}} = 17,54 \frac{\text{m}}{\text{s}} = 63,16 \frac{\text{km}}{\text{h}}$$

Entsprechend Gleichung (4) gilt für den maximalen Fehler:

$$\Delta v = \left| \frac{\partial v}{\partial s} \right| \cdot \Delta s + \left| \frac{\partial v}{\partial t} \right| \cdot \Delta t$$

Die partiellen Ableitungen lauten explizit:

$$\frac{\partial v}{\partial s} = \frac{\partial}{\partial s} v = \frac{\partial}{\partial s} \left(\frac{s}{t} \right) = \frac{1}{t} \quad \text{und} \quad \frac{\partial v}{\partial t} = \frac{\partial}{\partial t} v = \frac{\partial}{\partial t} \left(\frac{s}{t} \right) = -\frac{s}{t^2}$$

Für den maximalen Fehler ergibt sich damit

$$\begin{aligned} \Delta v &= \left| \frac{1}{t} \right| \cdot \Delta s + \left| -\frac{s}{t^2} \right| \cdot \Delta t = \left| \frac{1}{2,85 \text{ s}} \right| \cdot 0,5 \text{ m} + \left| -\frac{50 \text{ m}}{(2,85 \text{ s})^2} \right| \cdot 0,2 \\ &= 0,175 \frac{\text{m}}{\text{s}} + 1,23 \frac{\text{m}}{\text{s}} = 1,405 \frac{\text{m}}{\text{s}} \leq 1,5 \frac{\text{m}}{\text{s}} \cong 5,1 \frac{\text{km}}{\text{h}} \end{aligned}$$

Das Endergebnis ist demnach $v = (17,5 \pm 1,5) \frac{\text{m}}{\text{s}}$ bzw. $v = (63,1 \pm 5,1) \frac{\text{km}}{\text{h}}$.

Sind statistische Schwankungen für den abgeschätzten Fehler verantwortlich, und insbesondere die Messgrößen s und t voneinander unabhängig, verwendet man besser die Gaußsche Fehlerfortpflanzung entsprechend Gleichung (5):

$$\begin{aligned} \Delta v &= \sqrt{\left(\frac{\partial v}{\partial s} \right)^2 \cdot \Delta s^2 + \left(\frac{\partial v}{\partial t} \right)^2 \cdot \Delta t^2} = \sqrt{\left(\frac{1}{t} \right)^2 \cdot \Delta s^2 + \left(-\frac{s}{t^2} \right)^2 \cdot \Delta t^2} = \\ &= \sqrt{\left(\frac{1}{2,85 \text{ s}} \right)^2 \cdot 0,5^2 \text{ m}^2 + \left(-\frac{50 \text{ m}}{(2,85 \text{ s})^2} \right)^2 \cdot 0,2^2 \text{ s}^2} = 1,24 \frac{\text{m}}{\text{s}} \cong 4,47 \frac{\text{km}}{\text{h}} \end{aligned}$$

Das Endergebnis ist jetzt $v = (17,5 \pm 1,3) \frac{\text{m}}{\text{s}}$ bzw. $v = (63,1 \pm 4,5) \frac{\text{km}}{\text{h}}$.

8.2 Beispiel: Statistischer Fehler bei definiertem Vertrauensniveau

Die Tiefe eines Brunnens wird bestimmt, indem man kleine Steine hineinfallen lässt und die Zeit t bis zum „Platsch“ mit einer Stoppuhr misst. Die Tiefe des Brunnens x ergibt sich dann aus dem Zusammenhang $x = \frac{1}{2}gt^2$, wobei die Erdbeschleunigung mit $g = 9,81 \frac{m}{s^2}$ hier als genau bekannt angenommen wird.

Da von Hand gestoppt wird, ist die Zeitmessung ist sicher viel ungenauer als der systematische Fehler der Stoppuhr, der mit 0,01 s angenommen wird. Um eine größere Genauigkeit zu erreichen, wird die Messung daher 10 mal wiederholt.

Die Messwerte sind:

Messung Nr.	Zeit t / s
1	2,6
2	2,2
3	2,4
4	2,7
5	2,5
6	2,2
7	2,4
8	2,5
9	2,3
10	2,6

Mittelwert:

$$\bar{t} = \frac{1}{n} \sum_i^n t_i = 2,440 \text{ s}$$

Standardfehler einer Einzelmessung nach Gleichung (10):

$$\delta t = s = \sqrt{\frac{1}{n-1} \cdot \sum_i^n (t_i - \bar{t})^2} = 0,1625 \text{ s}$$

Standardfehler des Mittelwertes nach Gleichung (11):

$$\delta \bar{t} = \frac{s}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \cdot \sum_i^n (t_i - \bar{t})^2} = 0,05138 \text{ s}$$

Für den Fehler soll ein Vertrauensniveau von 95 % angesetzt werden. Entsprechend Tab.1 in Kap. 3.3 ergibt sich ein t_p -Wert von 2,3 und damit entsprechend Gleichung (12) für den Fehler des Mittelwertes

$$\Delta \bar{t}_{95\%,n=10} = t_p \cdot \frac{s}{\sqrt{n}} = 2,3 \cdot 0,05138 \text{ s} = 0,1182 \text{ s} .$$

Nach Addition des systematischen Fehlers (obwohl hier eigentlich vernachlässigbar) ergibt sich für den Fehler der Zeitmessung $\Delta \bar{t} = \Delta \bar{t}_{95\%,n=10} + \Delta t_{\text{sys.}} = 0,1282 \text{ s}$. Der wahrscheinlichste Wert für die Tiefe des Brunnens berechnet sich aus dem Mittelwert der Zeitmessung zu

$$x = \frac{1}{2}g\bar{t}^2 = \frac{1}{2} \cdot 9,81 \frac{m}{s^2} \cdot 2,44^2 \cdot s^2 = 29,20 \text{ m} .$$

Der Fehler für die Tiefe x des Brunnens ergibt sich aus dem Fehler der Zeitmessung und Fehlerfortpflanzung:

$$\Delta x = \left| \frac{\partial x}{\partial \bar{t}} \right| \cdot \Delta \bar{t} = 2 \cdot \frac{1}{2}g\bar{t} \cdot \Delta \bar{t} = 9,81 \frac{m}{s^2} \cdot 2,44 \text{ s} \cdot 0,1282 \text{ s} = 3,07 \text{ m} \leq 3,1 \text{ m}$$

Die Tiefe des Brunnens beträgt demnach $x = (29,2 \pm 3,1) \text{ m}$, bzw. x liegt mit einer Wahrscheinlichkeit von 95% im Intervall [26,1 m ; 32,3 m]. Der relative Fehler liegt bei über 10%.

8.3 Beispiele: Regressionsanalyse und Anfiten von Messdaten

8.3.1 Lineare Regression „von Hand“

Gegeben seien n Messwerte y_i (x_i) die vermutlich in einem linearen Zusammenhang stehen. Die Aufgabe, diesen durch eine lineare Gleichung der Form $y = m \cdot x + b$ darzustellen, führt bei Anwendung der Methode der kleinsten Quadrate zu einem lösbaeren Gleichungssystem mit dem Ergebnis:

$$m = \frac{n \cdot \sum(x_i \cdot y_i) - \sum x_i \cdot \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2} \quad b = \frac{\sum x_i^2 \cdot \sum y_i - \sum x_i \cdot \sum(x_i \cdot y_i)}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

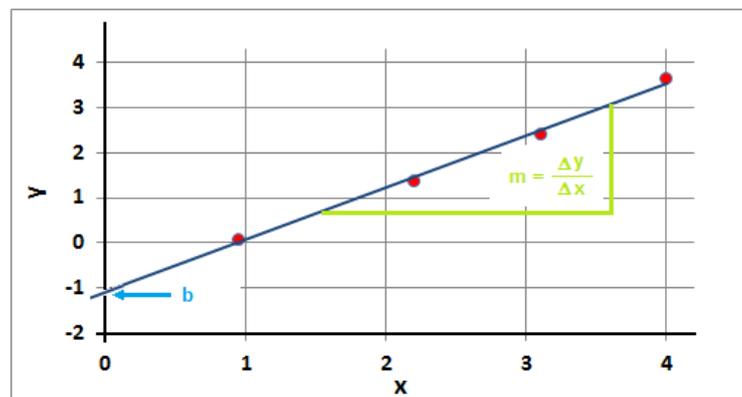
Für die Standardfehler gilt:

$$\delta m = \sqrt{\frac{n}{n \cdot \sum x_i^2 - (\sum x_i)^2}} \cdot \Delta y \quad \delta b = \sqrt{\frac{\sum x_i^2}{n \cdot \sum x_i^2 - (\sum x_i)^2}} \cdot \Delta y$$

$$\delta y = \sqrt{\frac{\sum y_i^2 - \frac{1}{n}(\sum y_i)^2 - \frac{m}{n} \cdot (n \cdot \sum(x_i \cdot y_i) - \sum x_i \cdot \sum y_i)}{n - 2}}$$

Bsp.: Messung von 4 Wertepaaren y_i (x_i)

x	y
0,95	0,08
2,2	1,38
3,1	2,41
4,0	3,65



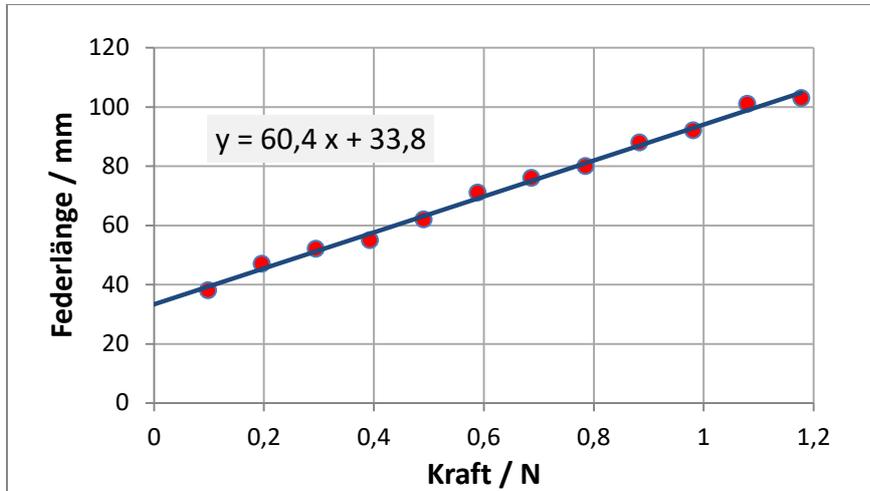
Die etwas mühselige Rechnung (vgl. *Lineare_Regression_mit_Excel.xls*) ergibt folgende Werte:

$$m = 1,16 \quad b = -1,10 \quad \delta y = 0,124 \quad \delta m = 0,055 \quad \delta b = 0,153$$

Die oben angegebenen Fehler entsprechen dem statistischen Standardfehler. Für ein höheres Vertrauensniveau von z.B. 95% muss der Standardfehler wieder mit dem entsprechenden t_p Faktor multipliziert werden, wobei hier der Wert bei $f = r - 2 = 4 - 2 = 2$ zu nehmen ist. Für eine statistische Sicherheit von 95% sind hier die Fehler (vgl. *tp_Faktoren_Tabelle.pdf*) mit $t_p = 4,3$ zu multiplizieren, also $\Delta m_{95\%} = 4,3 \cdot 0,055 = 0,24$ und $\Delta b_{95\%} = 4,3 \cdot 0,153 = 0,66$. Damit gilt mit einer Sicherheit von 95% für Steigung und Y-Achsenabschnitt $m = 1,16 \pm 0,24$ und $b = -1,10 \pm 0,66$.

8.3.2 Lineare Regression mit Excel

Die Länge einer hängenden Feder wird als Funktion daran aufgehängter Gewichte gemessen:



Mit Hilfe der Funktion „RGP“ von Excel (vgl. *Lineare_Regression_mit_Excel.xls*) wird eine Geradengleichung $y = m \cdot x + b$ angepasst. Man erhält als Ergebnis:

- 1) Y-Achsenabschnitt $b = (33,3 \pm 1,0)$ mm, entspricht der Länge der Feder ohne Belastung.
- 2) Steigung der Ausgleichsgeraden $m = (60,8 \pm 1,4)$ N/m, entspricht der Federkonstanten.

Die oben angegebenen Fehler entsprechen dem statistischen Standardfehler. Für ein höheres Vertrauensniveau von z.B. 99% muss der Standardfehler wieder mit dem entsprechenden t_p Faktor multipliziert werden, wobei hier der Wert bei $r = 12$ bzw. bei $f = r - 2 = 12 - 2 = 10$ zu nehmen ist. Der Tabellenwert (vgl. *tp_Faktoren_Tabelle.pdf*) für t_p ist 3,169. Der Fehler in der Steigung m , also der Federkonstanten, beträgt bei einer geforderten Sicherheit von 99% damit $\Delta m = t_p \cdot \delta m = 3,169 \cdot 1,4 \text{ N/m} = 4,44 \text{ N/m}$.

3) Ist z.B. aus konstruktiven Gründen nicht die Federkonstante, sondern die tatsächliche Länge der Feder entscheidend, verwenden wir den von der Funktion RGP ausgegebene Standardfehler „delta_y“ von in diesem Fall 1,634 mm und multiplizieren diesen mit dem t_p Faktor 3,169 von oben.

$$\Delta y_{99\%} = t_p \cdot \delta y = 3,169 \cdot 1,634 \text{ mm} = 5,178 \text{ mm} \leq 5,2 \text{ mm}$$

Mit 99% Sicherheit ist der mit den aus der Messung bestimmten Parametern berechnete Wert für die Federlänge mit einem Fehler von $\leq 5,2$ mm behaftet.

8.3.3 Lineare Regression mit Excel nach Linearisierung nichtlinearer Zusammenhänge

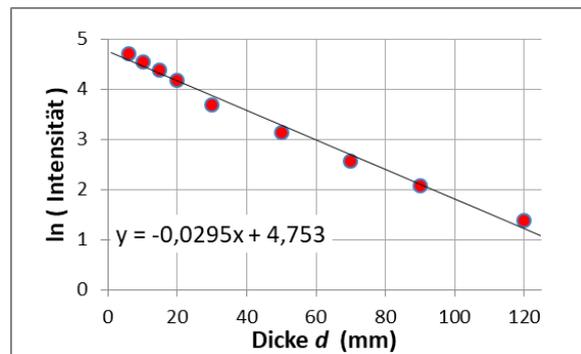
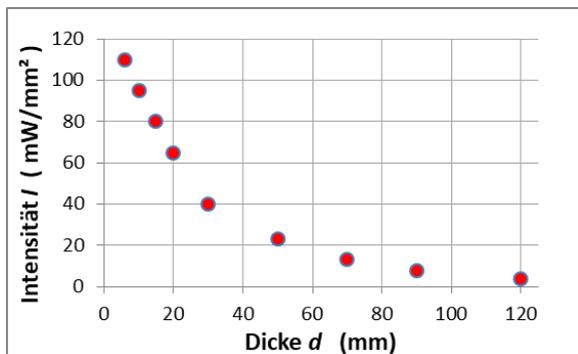
Oft lassen sich mit Hilfe der linearen Regression auch die Parameter von nichtlinearen funktionalen Zusammenhängen bestimmen, indem man zuvor durch eine geeignete mathematische Umformung eine Linearisierung vornimmt. Zwei häufig vorkommende Fälle sollen hier kurz beschrieben werden:

a) Exponentiell abfallende Funktionen

Die Linearisierung erfolgt in dem man den Logarithmus über der Variablen $x = d$ aufträgt.

Bsp.: Lambert-Beersches Gesetz

$$I = I_0 \cdot e^{-\alpha \cdot d} \quad \rightarrow \quad \ln(I) = \ln(I_0) - \alpha \cdot d =: m \cdot x + b$$



Die Lineare Regression (vgl. *Lineare_Regression_mit_Excel.xls*) ergibt für den linearen Dämpfungsparameter

$$\alpha = (0,0295 \pm 0,0012) \cdot 1/\text{mm},$$

und für die Anfangsintensität

$$I_0 = e^{4,753} \cdot \frac{\text{mW}}{\text{mm}^2} = (115,9 \pm 8,3) \cdot \frac{\text{mW}}{\text{mm}^2}$$

Der Fehler für I_0 ergibt sich aus dem Fehler für den Y-Achsenabschnitt $b = \ln(I_0)$ und Fehlerfortpflanzung entsprechend $I_0 = e^{\ln(I_0)} = e^b$:

$$\delta I_0 = \left| \frac{\partial I_0}{\partial b} \right| \cdot \delta b = \left| \frac{\partial}{\partial b} e^b \right| \cdot \delta b = e^b \cdot \delta b = 115,9 \cdot 0,071 = 8,223 < 8,3$$

Die hier angegebenen Fehler entsprechen wieder dem statistischen Standardfehler. Für ein höheres Vertrauensniveau muss ggf. wie in Kap. 8.3.2 beschrieben verfahren werden.

Anmerkung:

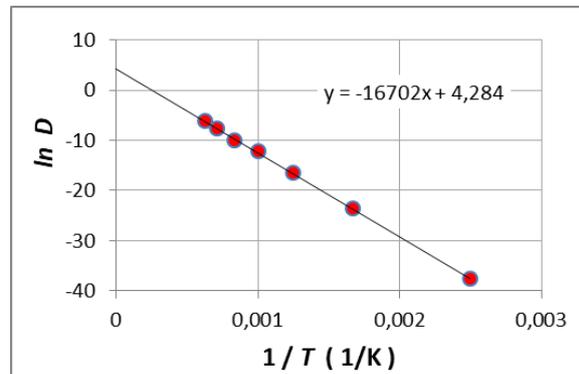
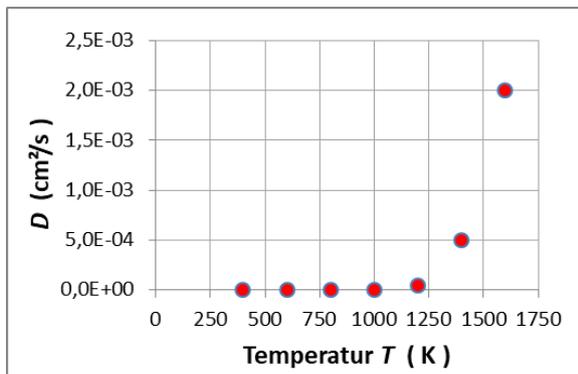
Die Linearisierung wie oben beschrieben funktioniert nicht mehr, wenn z.B. ein zusätzlicher additiver Offset I_{off} bei der Messung vorliegt und die gemessene Intensität eigentlich mit $I = I_0 \cdot e^{-\alpha \cdot d} + I_{\text{off}}$ beschrieben werden muss. In einem solchen Fall muss man den Offset versuchen separat zu messen und vor der Auswertung abzuziehen oder eine nichtlineare Regression mit der passenden Funktion durchführen. Dieser Fall tritt z.B. durch die sog. Nullrate bei Messungen einer Radioaktivität auf.

b) Thermische Prozesse die durch den sog. Boltzmannfaktor bestimmt sind

Die Linearisierung erfolgt, indem man den Logarithmus über der Variablen $x = 1/T$ aufträgt.

Bsp.: Temperaturabhängige Diffusionskonstante

$$D = D_0 \cdot e^{-\frac{Q_a}{RT}} \quad \rightarrow \quad \ln(D) = \ln(D_0) - \frac{Q_a}{R} \cdot \frac{1}{T} =: m \cdot x + b$$



Die Lineare Regression (vgl. *Beispiele_8.3.xls*) liefert nach Umrechnung die gesuchten Werte für den Frequenzfaktor $D_0 = (72,5 \pm 11) \frac{\text{cm}^2}{\text{s}}$ und die molare Aktivierungsenergie $Q_a = (139 \pm 1) \frac{\text{kJ}}{\text{mol}}$.

Der Standardfehler für D_0 berechnet sich wie bei a) mittels Fehlerfortpflanzung aus dem Fehler des Parameters $b = \ln(D_0)$ bzw. $D_0 = e^{\ln(D_0)} = e^b$:

$$\delta D_0 = \left| \frac{\partial D_0}{\partial b} \right| \cdot \delta b = \left| \frac{\partial}{\partial b} e^b \right| \cdot \delta b = e^b \cdot \delta b = 72,5 \cdot 0,142 = 10,29 < 11$$

Die Linearisierung von D wurde in Einheiten cm^2/s durchgeführt. Das Ergebnis für D_0 und der oben berechnete Fehler sind daher noch mit dieser Einheit zu multiplizieren.

Der Standardfehler für Q_a berechnet sich formal auch mittels Fehlerfortpflanzung aus dem Fehler des Parameters $m = \frac{-Q_a}{R}$:

$$Q_a = -R \cdot m \rightarrow$$

$$\delta Q_a = \left| \frac{\partial Q_a}{\partial m} \right| \cdot \delta m = |-R| \cdot \delta m = 8,3145 \frac{\text{J}}{\text{mol} \cdot \text{K}} \cdot 103,4 \cdot \text{K} = 859 \frac{\text{J}}{\text{mol}} < 1 \frac{\text{kJ}}{\text{mol}}$$

Es verbleibt also für δQ_a wie auch für Q_a nur die Multiplikation mit dem Faktor R .

Die hier angegebenen Fehler entsprechen wieder dem statistischen Standardfehler. Für ein höheres Vertrauensniveau muss ggf. wie in Kap. 8.3.2 beschrieben verfahren werden.

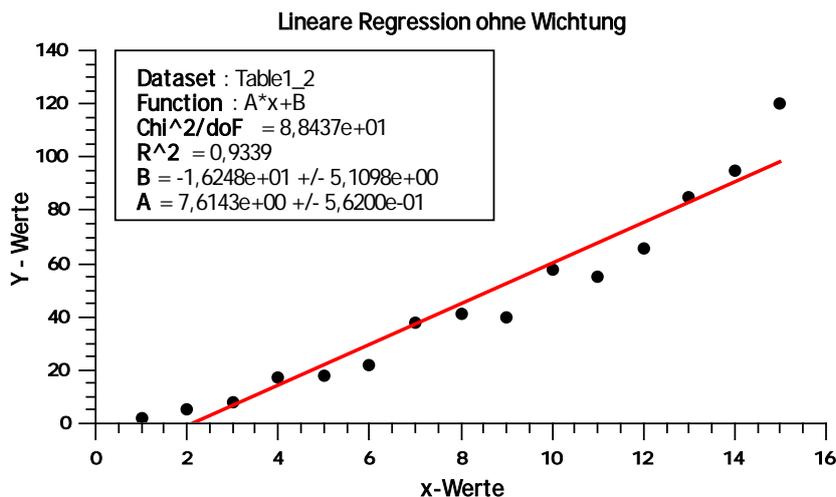
8.3.4 Lineare Regression mit individuellen Fehlern

Sind die Fehler der Einzelmessungen *unterschiedlich* groß (z.B. durch Wechsel des Messbereichs oder verschiedenen Zählraten in der Kernphysik), ist es sinnvoll die Einzelmesswerte entsprechend ihren jeweiligen Fehlern unterschiedlich zu wichten.

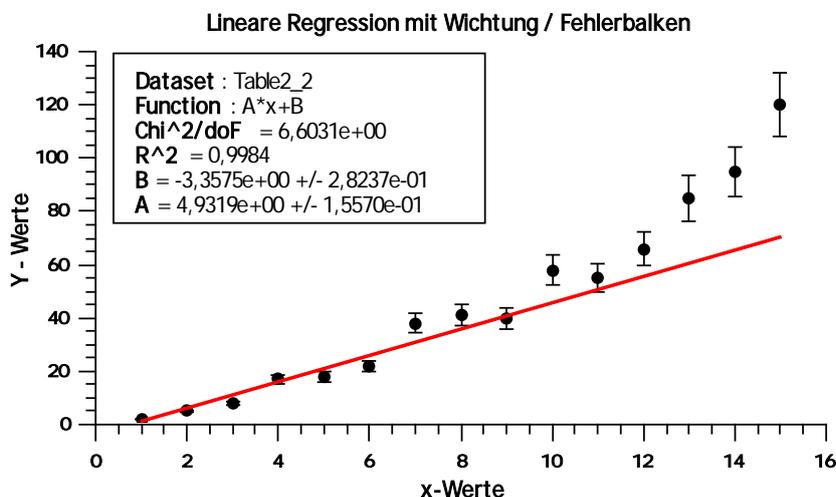
Hierfür benötigt man eine etwas „professionellere“ Software, wie z.B. Origin oder qtiplot, die hier entsprechend erweiterte Algorithmen verwenden. Wie unten gezeigt wird, ergeben sich bei unterschiedlicher Wichtung der Messpunkte entsprechend ihren jeweiligen Fehlern für die Fitparameter unterschiedliche Ergebnisse.

Bei qtiplot erfolgt die Wichtung durch das Einfügen von Fehlerbalken. (Dazu Grafik anwählen, und dann im Menü Graph → Add Error Bars ... wählen. Letztlich wird dann eine zusätzliche Spalte mit den Fehlern in der Datentabelle eingefügt, die dann auch frei editiert werden können.) Grundsätzlich können Fehlerbalken für x- und y-Werte angelegt werden.

Die folgenden Beispiele (vgl. *Fit_linear_gewichtet.qti*) zeigen die Auswirkung einer Wichtung durch Einfügen von Fehlerbalken (10% des Messwertes y) bzgl. der Messwerte y:



Einfache lineare Regression:
Alle Werte werden bei Berechnung der Ausgleichsgeraden gleich berücksichtigt.



Gewichtete lin. Regression nach Einfügen von Fehlerbalken: Die Messwerte y sind hier mit einem relativen Fehler von 10% behaftet.

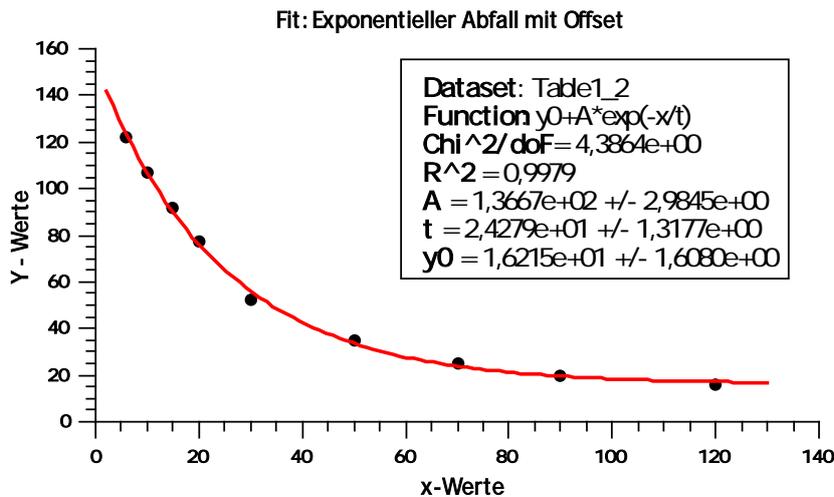
Die Ausgleichsgerade fällt deutlich flacher aus, da die „oberen“ Messwerte aufgrund ihres Fehlers weniger stark berücksichtigt werden.

8.3.5 Nichtlineare Regression mit qtiplot

Der Vorteil einer professionelleren Software wie Origin oder qtiplot ist vor allem, dass fast beliebige nichtlineare Funktionen wählbar bzw. editierbar sind. Insbesondere lassen sich dadurch auch Funktionen, die sich schlecht oder gar nicht linearisieren lassen, an Messdaten anpassen (fitten).

Im Folgenden werden nur kurz die Ergebnisse der Beispiele anbei dargestellt:

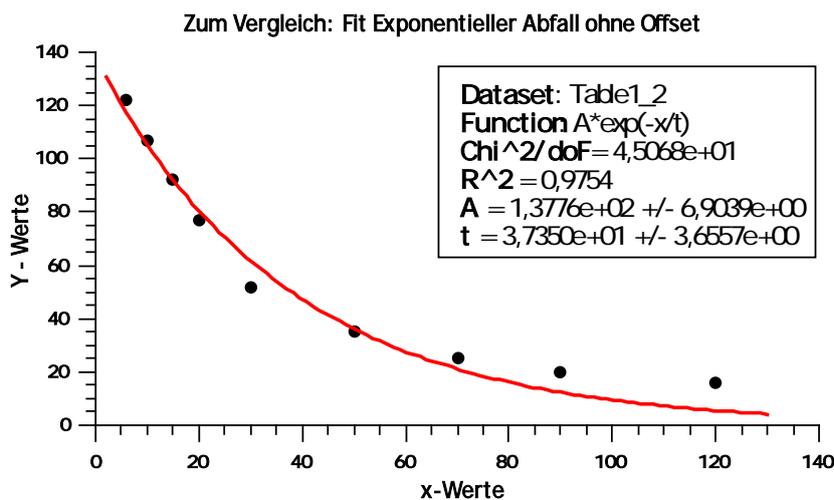
a) Exponentieller Abfall mit und ohne Offset (vgl. *Exponentieller_Abfall.qti*)



$$y(x) = y_0 + A \cdot e^{-\frac{x}{t}}$$

Durch Berücksichtigung des offensichtlich vorhandenen Offsets, werden die Messpunkte gut durch den Fit beschrieben.

Die Qualität des Fits ist mit $R^2 = 0,9979$ entsprechend hoch, die Fehler relativ klein.



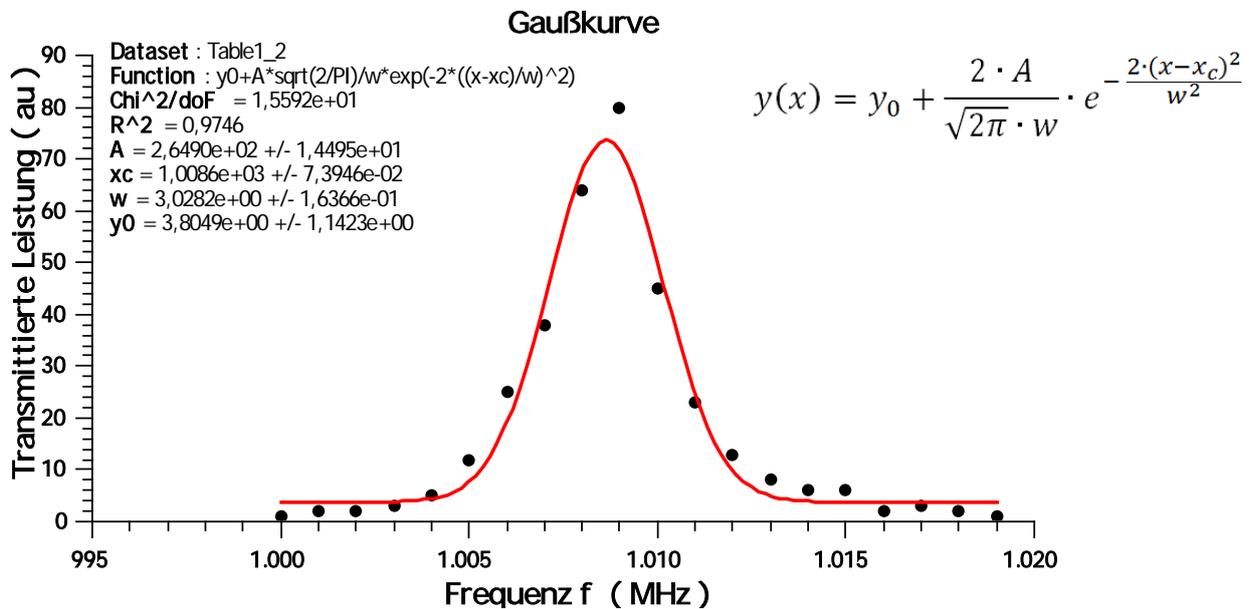
$$y(x) = A \cdot e^{-\frac{x}{t}}$$

Ohne Berücksichtigung des Offsets, werden die Messpunkte deutlich schlechter durch den Fit beschrieben.

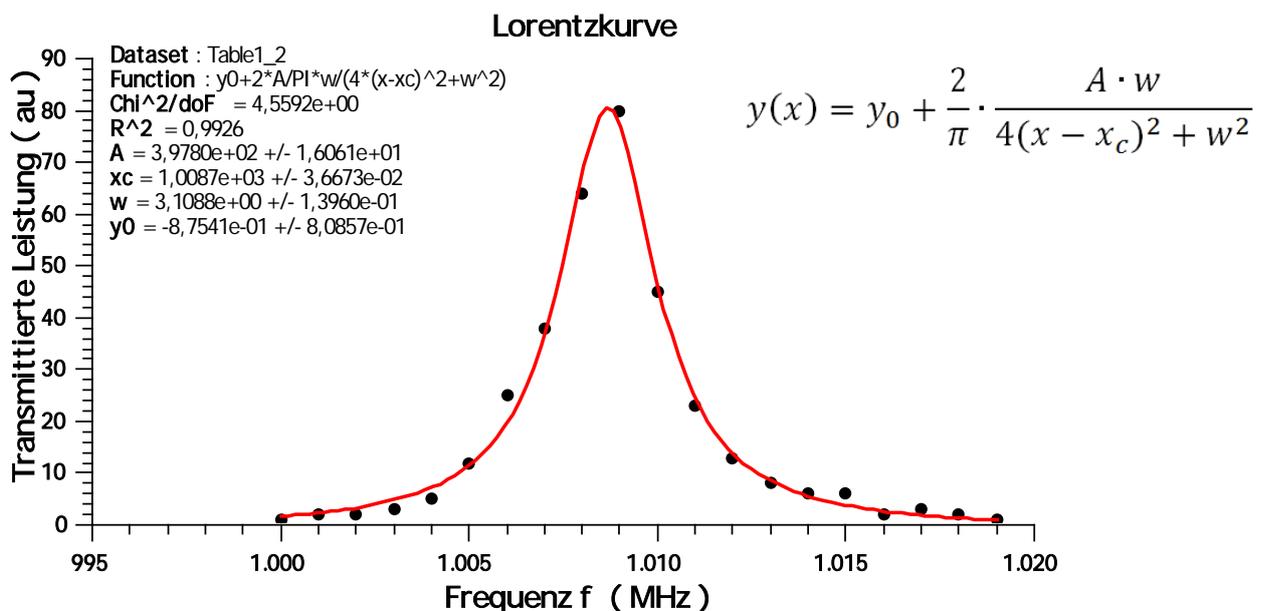
Die Qualität des Fits ist mit $R^2 = 0,9754$ schlechter und die Fehler für die Parameter A und t entsprechend deutlich größer.

b) Gaußkurve oder Lorentzkurve (vgl. *Gauss_Lorentz.qti*)

Messdaten ist i.A. nicht einfach anzusehen, ob sie besser mit einer Gaußkurve oder einer Lorentzkurve beschrieben werden. Hier hilft das Anfitzen und der anschließende Vergleich von beiden Modellen. Man erhält dadurch nicht nur eine besserer Beschreibung der Messung, sondern i.A. auch Informationen über die zugrundeliegenden physikalischen Prozesse.



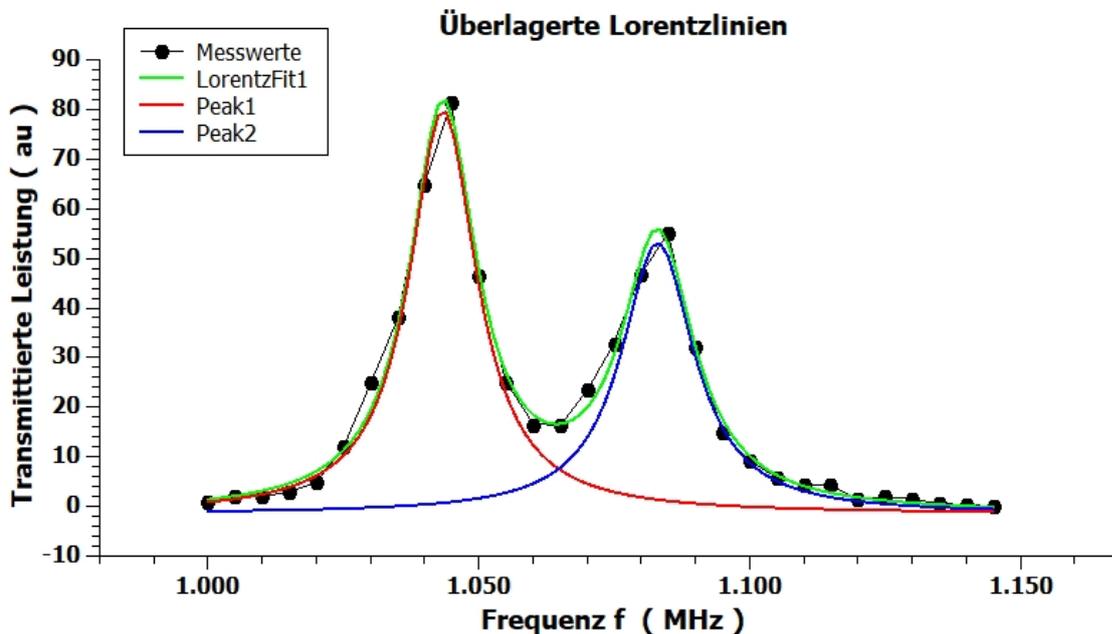
Fit einer Gaußkurve an fiktive Messdaten. Die wichtigen Ergebnisse sind mit der Kurvenlage $x_c = 1008,6$ und der Kurvenbreite $w = 3,03$ gegeben. Die Streuung σ entspricht $w/2$ also ca. 1,5.



Fit einer Lorentzkurve an die gleichen Daten mit deutlich besserem Ergebnis, entsprechend liegt die Qualität des Fits mit $R^2 = 0,993$ deutlich höher. Die Kurvenlage (\rightarrow Resonanzfrequenz in MHz) liegt bei $x_c = 1008,7$. Der Parameter $w = 3,11$ gibt die sog. Halbwertslinienbreite ($\rightarrow \Delta f_{-3dB}$) in MHz an. Damit bestimmt sich z.B. die Güte des Resonators zu $f/\Delta f_{-3dB} = 324$.

c) Fit von überlagerten Kurven, hier: Spektrum mit zwei Lorentzlinien (z.B. Resonanzmoden)

(vgl. *Multiple_Lorentz.qti.*)



In diesem Beispiel überlagern sich zwei Lorentzlinien (oder – kurven), diese sind zwar noch zu unterscheiden, aber nicht getrennt auswertbar. Hier hilft nur ein Fit mit dem Ansatz für zwei überlagerte Linien. Als Ergebnis erhält man dann die Parameter für zwei Linien:

$$\begin{array}{ll}
 A1 = 1,9293e+03 \pm 3,0999e+01 & A2 = 1,3993e+03 \pm 3,3655e+01 \\
 xc1 = 1,0434e+03 \pm 8,5631e-02 & xc2 = 1,0830e+03 \pm 1,3233e-01 \\
 w1 = 1,5086e+01 \pm 2,9119e-01 & w2 = 1,6289e+01 \pm 4,6280e-01
 \end{array}$$

In diesem Fall also die Resonanzlinien zweier Moden bei 1043,4 MHz und 1083,0 MHz mit leicht unterschiedlichen Amplituden aber vergleichbarer Linienbreite von ca. 1,5 MHz.

9 Im Script verwendete Notation

Δx	Fehler
δx	Standardfehler (Statistik)
s	Standardabweichung, oder Wegstrecke
σ	Streuung
x, y, z	Messwerte
μ	wahrer Wert (Erwartungswert)
f	Anzahl der Freiheitsgrade
n	Anzahl von Messungen
r	Anzahl Wertepaare bei lin. Regression
w_i	Wichtungsfaktoren
t_p	t_p - bzw. t - Faktor bzgl. Student-t-Verteilung