

# Fehlerrechnung



Dieser „Fehler“ ereignete sich am 22.10.1895 als ein Zug, der von Granville nach Paris fuhr, nicht wie vorgesehen am Gare Montparnasse anhielt.

## 1 Messgenauigkeit

Wird eine Messgröße  $x$  mehrmals gemessen, so zeigen die Messwerte regellose Schwankungen. Zur grafischen Veranschaulichung dieser Messwertschwankungen dient das *Histogramm*. Als Beispiel sind in Bild 1-1 die gemessenen Schwingungsdauern eines Pendels dargestellt. Die zugehörigen Messwerte sind in nebenstehende Tabelle zusammengestellt.

Schwingungsdauer $T_j$ /s	Häufigkeit $N_j$
1,18	1
1,19	3
1,20	5
1,21	7
1,22	4
1,23	2
1,24	2
1,25	1

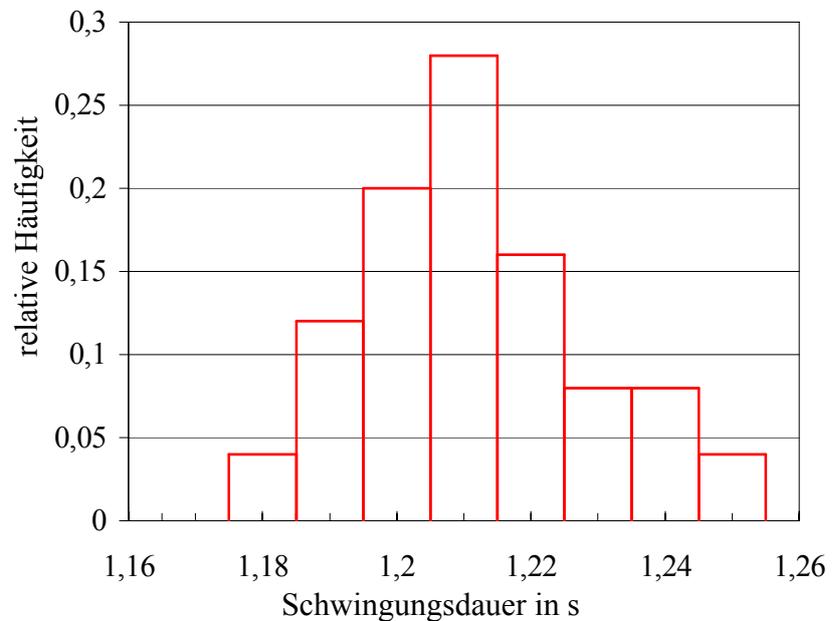


Bild 1-1: Histogramm der gemessenen Schwingungsdauern eines Pendels, Anzahl der Messungen:  $N = 25$ .

Zweckmäßigerweise trägt man anstelle der absoluten Häufigkeit  $N_j$ , mit welcher der Messwert  $x_j$  auftritt die relative Häufigkeit auf, die gegeben ist durch

$$h_j = \frac{N_j}{N}, \quad (1-1)$$

dabei ist  $N$  die Gesamtzahl der Messungen.

*C. F. Gauß* konnte zeigen, dass bei einer großen Zahl von Messungen die Häufigkeitsverteilung  $h(x_j)$  übergeht in die nach ihm benannte *Normalverteilung*

$$h(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (1-2)$$

$h(x) dx$  ist die Wahrscheinlichkeit, dass bei einer Wiederholungsmessung der Messwert  $x$  zwischen  $x$  und  $x + dx$  liegt. Die Funktion  $h(x)$  ist symmetrisch zum *Erwartungswert*  $\mu$ . Sie ist durch den Vorfaktor  $\frac{1}{\sqrt{2\pi\sigma^2}}$  so normiert, dass die Wahrscheinlichkeit 1 ist, bei einer Wie-

derholungsmessung einen Wert  $x$  im Bereich  $-\infty < x < +\infty$  zu finden:  $\int_{-\infty}^{+\infty} h(x) dx = 1$ . Die Breite

der Verteilungsfunktion  $h(x)$  wird bestimmt durch die *Standardabweichung* oder *Streuung*  $\sigma$ ,  $\sigma^2$  wird als *Varianz* bezeichnet. Sie hängt ab von der Genauigkeit des Messverfahrens.

Bild 1-2 zeigt ein Beispiel einer Normalverteilung.

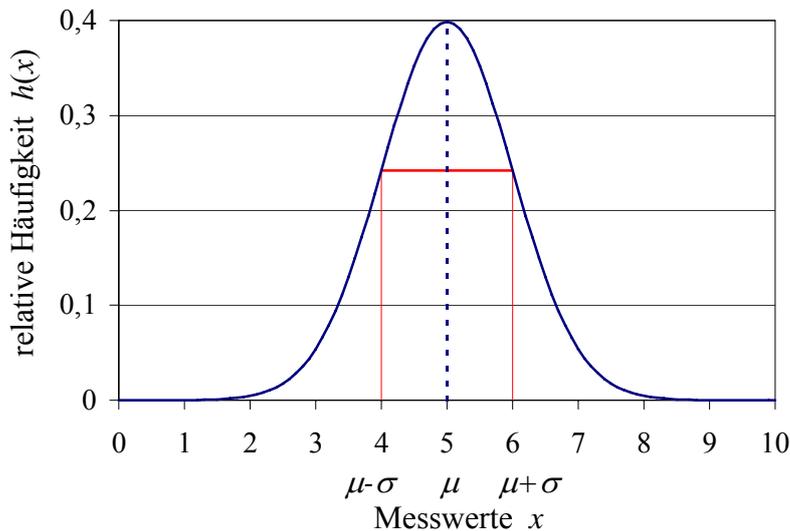


Bild 1-2: Gaußsche Normalverteilung für  $\mu = 5$  und  $\sigma = 1$ .

Die Breite der Verteilungsfunktion kann auch charakterisiert werden durch die auf halber Höhe gemessene Breite  $b_{1/2}$  (engl. **F**ull **W**idth at **H**alf **M**aximum, FWHM), mit der folgender Zusammenhang existiert:

$$b_{1/2} = \text{FWHM} = 2\sqrt{2\ln 2} \sigma = 2,35 \sigma. \quad (1-3)$$

Die Wahrscheinlichkeit, dass ein Messwert in den Bereich  $\mu - \sigma < x < \mu + \sigma$  fällt beträgt

$$P = \int_{\mu - \sigma}^{\mu + \sigma} h(x) dx = 68,3\%. \text{ Wählt man den Bereich doppelt so breit, also } \mu - 2\sigma < x < \mu + 2\sigma,$$

dann beträgt die Wahrscheinlichkeit 95,4 %.

Aus einer endlichen Zahl  $N$  von Messungen lassen sich nach der Gaußschen Theorie der Beobachtungsfehler Schätzwerte für den Erwartungswert  $\mu$  und die Streuung  $\sigma$  berechnen. Der beste Schätzwert für den Erwartungswert ist der *arithmetische Mittelwert*

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (1-4)$$

Er zeichnet sich dadurch aus, dass die Summe der quadratischen Abweichungen zwischen den Messwerten  $x_i$  und diesem Wert minimal wird.

Diese so genannte *minimale Fehlersumme* beträgt

$$FS_{\min} = \sum_{i=1}^N (x_i - \bar{x})^2 = \sum_{i=1}^N x_i^2 - N\bar{x}^2. \quad (1-5)$$

Der beste Schätzwert für die Streuung  $\sigma$  ist die *empirische Standardabweichung* des Messverfahrens bzw. der Messwerte  $x$ :

$$s_x = \sqrt{\frac{FS_{\min}}{N-1}} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}. \quad (1-6)$$

Die relative Unsicherheit der Standardabweichung ist

$$\frac{\Delta s_x}{s_x} = \frac{1}{\sqrt{2(N-1)}}. \quad (1-7)$$

Berechnet man mit Hilfe der Gleichungen (1-4) bis (1-6) den Mittelwert und die Standardabweichung des Datensatzes von Seite 2, so ergibt sich für den Mittelwert der gemessenen Schwingungsdauern  $\bar{T} = 1,2116$  s und für die Standardabweichung des Messverfahrens  $s_T = 0,0172$  s. Bild 1-3 zeigt eine Darstellung der Ergebnisse.

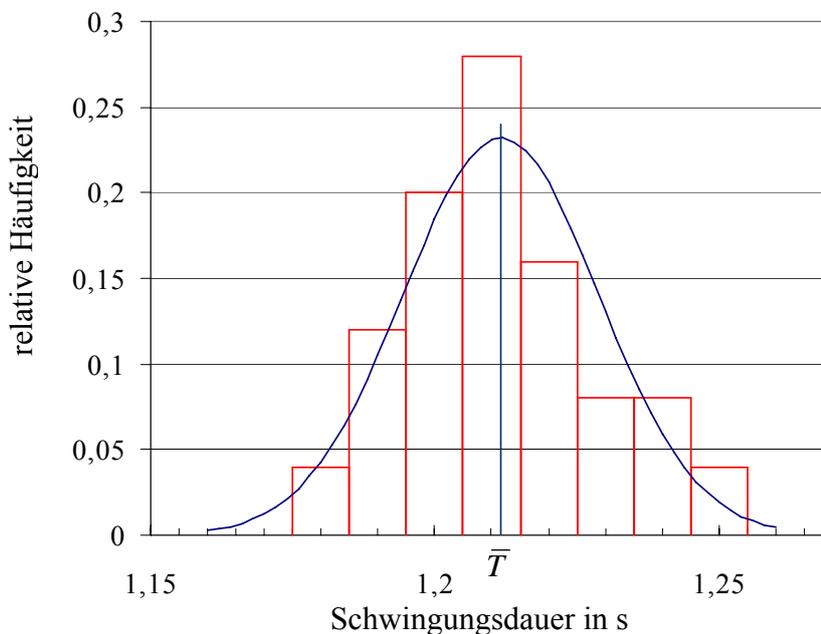


Bild 1-3: Histogramm der Messwerte aus der Tabelle von Seite 2 sowie berechnete Normalverteilung (multipliziert mit der Klassenbreite, hier 0,01 s) mit den Werten  
 $\bar{T} = 1,2116$  s,  
 $s_T = 0,0172$  s,  
 $\Delta \bar{T} = 0,0034$  s.

Wenn nach einer Messreihe der Mittelwert  $\bar{x}$  der zu messenden Größe bestimmt wird, dann wird bei einer Wiederholung der Messreihe der neu bestimmte Mittelwert sicher nicht mit dem vorigen übereinstimmen. Macht man weitere Messreihen, so werden die jeweils bestimmten Mittelwerte eine statistische Schwankung aufweisen. Ein Maß für die Abweichung des Mittelwerts  $\bar{x}$  vom Erwartungswert  $\mu$  ist die *Standardabweichung des Mittelwerts*

$$s_{\bar{x}} = \Delta\bar{x} = \frac{s_x}{\sqrt{N}}. \quad (1-8)$$

Die Genauigkeit eines Messverfahrens bestimmt die Breite der Häufigkeitsverteilung. Die Standardabweichung  $s_x$  charakterisiert somit die Messgenauigkeit des verwendeten Verfahrens und kann durch Wiederholungsmessungen nicht erhöht werden; dazu muss das Messverfahren geändert werden.

Dagegen erhöhen Wiederholungsmessungen die Genauigkeit, mit welcher der berechnete Mittelwert  $\bar{x}$  mit dem Erwartungswert  $\mu$  übereinstimmt. Die Standardabweichung des Mittelwerts  $\Delta\bar{x}$  nimmt mit steigender Zahl der Messungen ab.

Für unser Zahlenbeispiel von Seite 2 ergibt sich für die Standardabweichung des Mittelwerts  $\Delta\bar{T} = 0,0034$  s. Der *relative Fehler* des Mittelwerts ist  $\frac{\Delta\bar{T}}{\bar{T}} = 0,28\%$ .

Das Ergebnis der Messungen kann folgendermaßen angegeben werden:

$$T = (1,2116 \pm 0,0034) \text{ s, oder } T = 1,2116 \text{ s } (1 \pm 0,28\%).$$

Es sei noch einmal darauf hin gewiesen, dass die Angabe des Ergebnisses in der Form  $x = \bar{x} \pm \Delta\bar{x}$  nicht bedeutet, dass der wahre Wert der zu messenden Größe innerhalb der angegebenen Fehlergrenzen liegt, für unser Beispiel also zwischen 1,2082 s und 1,2150 s! Vielmehr liegt der wahre Wert innerhalb dieser Grenzen nur mit einer Wahrscheinlichkeit von 68,3 % (s. Seite 3). Wird das Messergebnis in der Form  $x = \bar{x} \pm 2\Delta\bar{x}$  angegeben, so liegt der wahre Wert mit einer Wahrscheinlichkeit von 95,4 % innerhalb dieser Grenzen.

## 2 Fehlerfortpflanzung

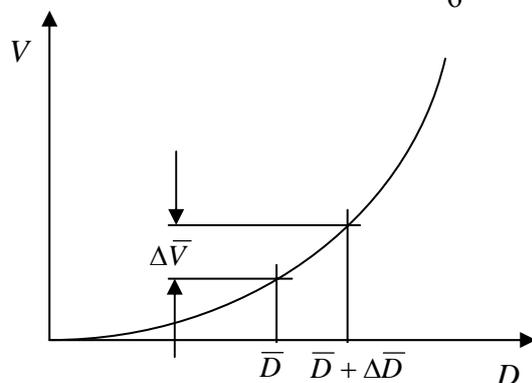
Wird eine physikalische Größe nicht direkt gemessen, sondern aus anderen gemessenen Größen berechnet, dann pflanzen sich die Fehler bis zum Endergebnis fort.

Beispiel:

Das Volumen einer Kugel soll bestimmt werden aus dem gemessenen Durchmesser

$$D = (100 \pm 0,1) \text{ mm.}$$

Das mittlere Kugelvolumen ist  $\bar{V} = \frac{\pi}{6} \bar{D}^3 = 523,6 \text{ cm}^3$ .



Der Fehler des Volumens infolge des fehlerhaften Radius wird mit Hilfe der Differentialrechnung bestimmt:

$$\Delta\bar{V} = \left[ \frac{dV}{dD} \right]_{\bar{D}} \cdot \Delta\bar{D} = \frac{\pi}{2} \bar{D}^2 \Delta\bar{D}.$$

Bild 2-1: Kugelvolumen in Abhängigkeit vom Durchmesser.

Der relative Fehler wird  $\frac{\Delta \bar{V}}{\bar{V}} = 3 \frac{\Delta \bar{D}}{\bar{D}}$ , er ist also drei Mal so groß wie der relative Fehler des Durchmessers. Dabei stammt der Faktor 3 ganz offensichtlich von der Potenz im Gesetz  $V = \frac{\pi}{6} D^3$ .

Hängt eine physikalische Größe  $f(x, y, z, \dots)$  von mehreren Messgrößen ab, dann ist der wahrscheinlichste Wert der indirekt bestimmten Größe

$$\bar{f} = f(\bar{x}, \bar{y}, \bar{z}, \dots). \quad (2-1)$$

Nach dem *Fehlerfortpflanzungsgesetz* von Gauß ist die Standardabweichung der Größe  $f$  bzw. des indirekten Messverfahrens für  $f$

$$s_f = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 s_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 s_y^2 + \left(\frac{\partial f}{\partial z}\right)^2 s_z^2 + \dots}, \quad (2-2)$$

dabei sind  $\left(\frac{\partial f}{\partial x}\right)$ ,  $\left(\frac{\partial f}{\partial y}\right)$ ,  $\left(\frac{\partial f}{\partial z}\right)$ , ... die partiellen Ableitungen der Funktion  $f(x, y, z, \dots)$  nach den Teilgrößen  $x, y, z, \dots$  an der Stelle  $\bar{x}, \bar{y}, \bar{z}, \dots$

Ersetzt man nach Gl. (1-8) die Standardabweichungen  $s_x, s_y, s_z, \dots$  durch die Standardabweichungen der Mittelwerte  $\Delta \bar{x}, \Delta \bar{y}, \Delta \bar{z}, \dots$ , dann ergibt sich

$$s_f = \sqrt{\left[\left(\frac{\partial f}{\partial x}\right)^2 \Delta \bar{x}^2 + \left(\frac{\partial f}{\partial y}\right)^2 \Delta \bar{y}^2 + \left(\frac{\partial f}{\partial z}\right)^2 \Delta \bar{z}^2 + \dots\right] N}. \quad (2-3)$$

Der mittlere Fehler bzw. die Standardabweichung des Mittelwertes  $\bar{f}$  beträgt

$$s_{\bar{f}} = \Delta \bar{f} = \frac{s_f}{\sqrt{N}} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 \Delta \bar{x}^2 + \left(\frac{\partial f}{\partial y}\right)^2 \Delta \bar{y}^2 + \left(\frac{\partial f}{\partial z}\right)^2 \Delta \bar{z}^2 + \dots}. \quad (2-4)$$

Häufig ist in der Praxis die Zahl der Messungen zu gering, so dass die Gleichungen (2-2) und (2-4), die eigentlich nur für eine große Anzahl  $N$  von Messungen gültig sind, zu kleine Fehler liefern. Daher ist es in vielen Fällen angebracht, die *Größtfehler* zu berechnen.

Der absolute Größtfehler des Mittelwertes  $\bar{f}$  beträgt

$$\Delta \bar{f}_{\max} = \left| \frac{\partial f}{\partial x} \right| \Delta \bar{x} + \left| \frac{\partial f}{\partial y} \right| \Delta \bar{y} + \left| \frac{\partial f}{\partial z} \right| \Delta \bar{z} + \dots \quad (2-4)$$

Der relative Größtfehler  $\frac{\Delta \bar{f}_{\max}}{\bar{f}}$  lässt sich besonders einfach berechnen bei reinen Potenzfunktionen vom Typ  $f = x^k \cdot y^m \cdot z^n$ :

$$\frac{\Delta \bar{f}_{\max}}{\bar{f}} = |k| \left| \frac{\Delta \bar{x}}{\bar{x}} \right| + |m| \left| \frac{\Delta \bar{y}}{\bar{y}} \right| + |n| \left| \frac{\Delta \bar{z}}{\bar{z}} \right|. \quad (2-5)$$

### 3 Ausgleichsgerade - lineare Regression

Wenn zwischen zwei Messgrößen  $y$  und  $x$  ein linearer Zusammenhang vermutet, bzw. durch eine Theorie vorhergesagt wird, lässt sich durch die Messpunkte eine optimale Gerade, die so genannte Ausgleichsgerade legen (Bild 3-1). Dabei wird angenommen, dass die Fehler der  $x$ -Werte vernachlässigbar sind. Wird die Ausgleichsgerade formuliert als

$$y = a_0 + a_1 x,$$

so ist für den Messwert  $x_i$  der Fehler zwischen dem Funktionswert  $y$  und dem Messwert  $y_i$  gegeben durch

$$a_0 + a_1 x_i - y_i.$$

Wieder wird gefordert, dass die Summe aller quadratischen Abweichungen ein Minimum sein soll:

$$\sum_{i=1}^N (a_0 + a_1 x_i - y_i)^2 = \text{Min!}$$

Partielles Ableiten nach  $a_0$  und  $a_1$  und Nullsetzen der Ableitungen führt auf ein System von Normalgleichungen

$$a_0 N + a_1 \sum x_i = \sum y_i$$

$$a_0 \sum x_i + a_1 \sum x_i^2 = \sum x_i y_i,$$

das nach den üblichen Methoden der linearen Algebra gelöst wird.

Mit der Koeffizienten-Determinante

$$\Delta = N \sum x_i^2 - \left( \sum x_i \right)^2 \quad (3-1)$$

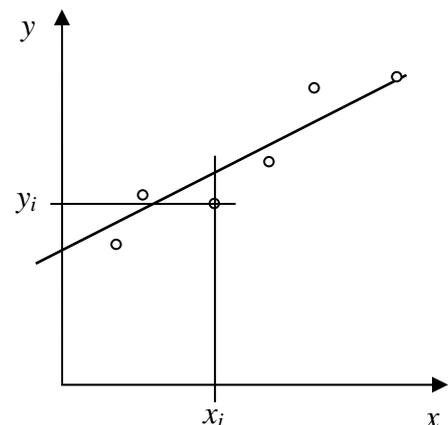


Bild 3-1: Lineare Regression.

ergibt sich für den Ordinatenabschnitt

$$a_0 = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{\Delta} \quad (3-2)$$

und für die Steigung der Geraden

$$a_1 = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\Delta} \quad (3-3)$$

Die Standardabweichung der  $y$ -Werte beträgt

$$s_y = \sqrt{\frac{1}{N-2} \sum_{i=1}^N (a_0 + a_1 x_i - y_i)^2} \quad (3-4)$$

Die Standardabweichungen (Unsicherheiten) der nach den Gleichungen (3-2) und (3-3) berechneten Werte für den Ordinatenabschnitt und die Steigung betragen

$$s_{a_0} = s_y \sqrt{\frac{\sum x_i^2}{\Delta}} \quad (3-5)$$

und

$$s_{a_1} = s_y \sqrt{\frac{N}{\Delta}} \quad (3-6)$$

Es ist häufig die Frage, ob die beiden Messgrößen  $x$  und  $y$  tatsächlich korreliert sind, ob also eine Funktion vom Typ  $y = a_0 + a_1 x$  einen sinnvollen Zusammenhang darstellt. Eine Antwort darauf liefert die *Kovarianz*

$$s_{xy} = \frac{1}{N-1} \sum (x_i - \bar{x})(y_i - \bar{y}) \quad (3-7)$$

Falls die Messwerte von  $x$  und  $y$  nicht korreliert sind, nähert sich die Kovarianz für eine große Zahl von Messungen dem Wert null. Eine quantitative Beschreibung der Korrelation ist möglich mit dem *Korrelationskoeffizienten*

$$r = \frac{s_{xy}}{s_x s_y}$$

Mithilfe der Gleichungen (1-6) und (3-7) ergibt sich

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (3-8)$$

oder

$$r = \frac{\sum x_i y_i - N \bar{x} \bar{y}}{\sqrt{(\sum x_i^2 - N \bar{x}^2)(\sum y_i^2 - N \bar{y}^2)}} \quad (3-9)$$

Liegen alle Messpunkte auf einer Geraden, dann ist der Korrelationskoeffizient  $r = \pm 1$  (je nach Steigung). Liegt keine Korrelation vor, dann ist  $r = 0$ . Als Faustregel kann man sagen, dass bei  $|r| > 0,8$  eine Korrelation vorliegt.

**Beispiel:**

Die Heizleistung eines Hauses in Abhängigkeit von der äquivalenten Außentemperatur (berücksichtigt Sonneneinstrahlung und Wind) wird untersucht. Eine lineare Regression soll mit Hilfe von Excel durchgeführt werden.

	A	B	C	D	E	F
1	Tag-Nummer	äquiv. Temp	Heizleistung			
2		x/Celsius	y/kW			
3	1	0,8	85		-5,79	82,68
4	2	0,4	81		0,61	1,48
5	3	3,2	67		0,86	5,08
6	4	-3	93		90,35	15,00
7	5	1,2	81		2327,36	386,41
8	6	-0,7	88			
9	7	-2	102			
10	8	0,6	73			
11	9	4,2	65			
12	10	3,5	64			
13	11	0,2	78			
14	12	2	65			
15	13	0,7	81			
16	14	1,4	74			
17	15	2,6	65			
18	16	4,4	52			
19	17	3,4	59			

Nach Eingabe der Messwerte muss ein Ausgabebereich in Form einer Matrix mit 5 Zeilen und 2 Spalten markiert werden (hier: E3 bis F7). Mit Hilfe des Funktionsassistenten wird die Statistik-Funktion RGP aufgerufen. Eingabe: =RGP(y-Werte;x-Werte;1;1). Matrixbefehle werden ausgeführt mit Strg + ↑ + Enter.

Das Ergebnis wird in folgender Form ausgegeben (auf die Bedeutung der letzten beiden Zeilen wird hier nicht eingegangen):

$a_1$	$a_0$
Stand.abw. $s_{a1}$	Stand.abw. $s_{a0}$
Best.maß $R = r^2$	Stand.abw. $s_y$

Excel gibt anstelle des Korrelationskoeffizienten  $r$  dessen Quadrat, das so genannte *Bestimmtheitsmaß* aus. Eine Darstellung der Anpassung zeigt Bild 3-2. Der Korrelationskoeffizient beträgt  $r = -0,93$ .

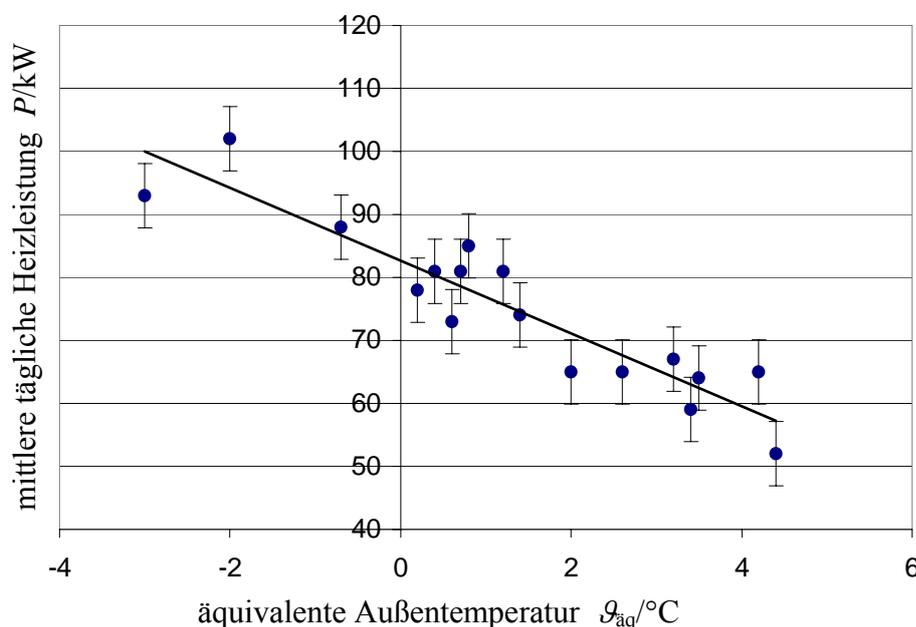


Bild 3-2: Beispiel einer linearen Regression mit  $r = -0,93$ .

### Vorgehen bei nicht linearen Zusammenhängen

Nicht lineare Zusammenhänge lassen sich häufig mithilfe einer Koordinatentransformation in lineare Zusammenhänge überführen. Ein Beispiel zeigt Bild 3-3, wo die gemessene spektrale Linienbreite  $\Delta f_M$  eines Halbleiterlasers dargestellt ist in Abhängigkeit vom Spiegelabstand  $d$  eines Fabry-Perot-Interferometers. Die eigentlich interessierende wahre Linienbreite  $\Delta f_L$  eines Lasers wird bei einer solchen Messung verfälscht (verbreitert) zur gemessenen Linienbreite

$$\Delta f_M = \Delta f_L + \frac{c}{2Fd}, \quad (3-10)$$

dabei ist  $c$  die Lichtgeschwindigkeit und  $F$  die so genannte Finesse des Instruments. Gesucht ist die wahre Linienbreite  $\Delta f_L$  des Lasers. Die in Bild 3-3 a) dargestellten Messpunkte lassen zwar einen Zusammenhang nach Gleichung (3-10) vermuten aber nicht ohne weiteres beweisen.

Trägt man  $\Delta f_M(x)$  auf, wobei  $x = \frac{1}{d}$ , dann geht Gleichung (3-10) in eine Geradengleichung über:

$$\Delta f_M = \Delta f_L + \frac{c}{2F} x, \quad (3-11)$$

und ermöglicht eine lineare Regression nach Bild 3-3 b). Als Ergebnis der Anpassung folgt:

$\Delta f_L = 21$  MHz (Ordinatenabschnitt),

Steigung  $a_1 = \frac{c}{2F} = 95,8$  MHz · mm

und damit für die Finesse des Spektrometers  $F = 1566$ .

Der Korrelationskoeffizient beträgt  $r = 0,998$ .

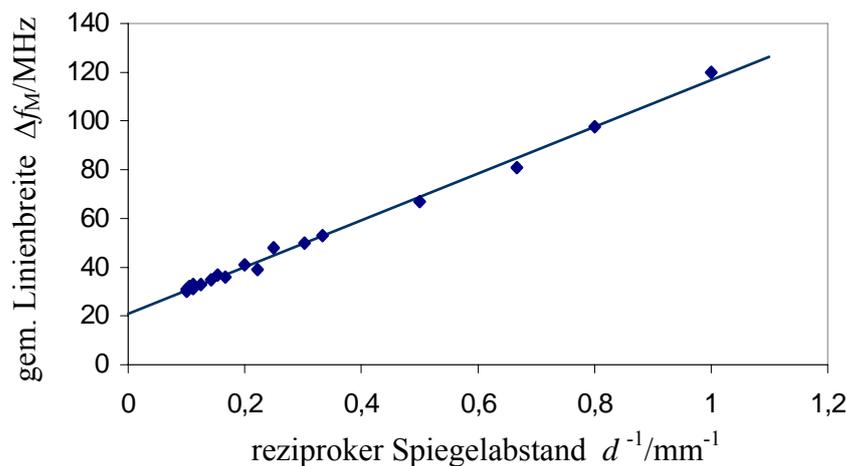
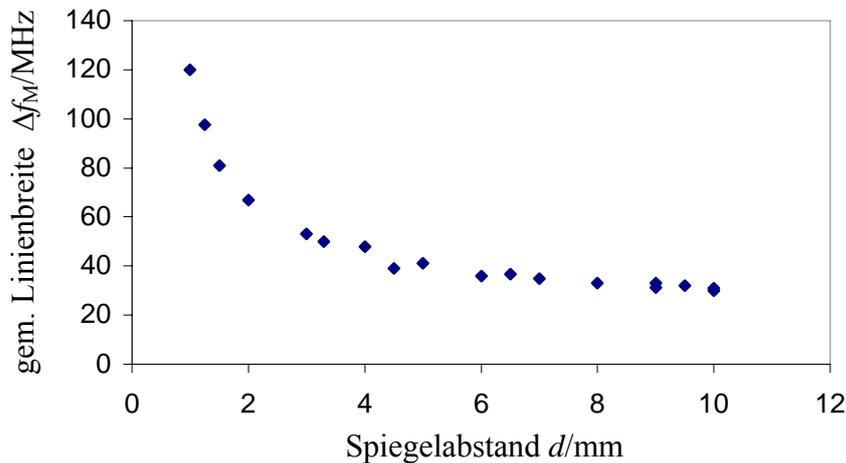


Bild 3-3: Transformation einer nicht linearen in eine lineare Funktion.

Programme wie Excel, Winstat etc. erlauben selbstverständlich auch nicht lineare Anpassungen von theoretischen Funktionen an experimentelle Ergebnisse. Dennoch sind Darstellungen der Zusammenhänge in Form von Geraden am besten geeignet, physikalische Modelle zu überprüfen.

**Beispiel:** Erwärmung eines Körpers mit der Anfangstemperatur  $T_A$  in einem Ofen der Temperatur  $T_O$ .

Wärmestrom vom Ofen in den Körper, wenn  $T$  die Körpertemperatur ist:  
 $\dot{Q} = kA(T_O - T)$ ,  $A$ : Oberfläche,  $k$ : Wärmedurchgangskoeffizient.

Mit  $dQ = C dT$  ( $C$ : Wärmekapazität) folgt für die Temperaturänderung die Dgl.

$$\frac{dT}{dt} = \frac{kA}{C}(T_O - T), \text{ mit der Lösung}$$

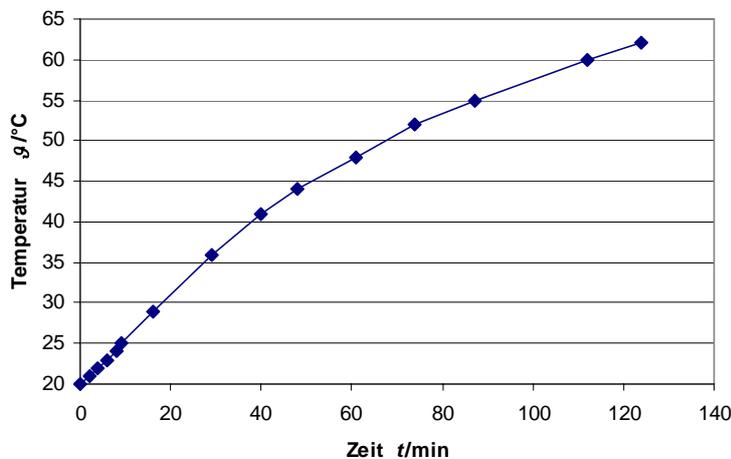
$$T(t) = T_A + (T_O - T_A)(1 - e^{-t/\tau}), \text{ wobei } \tau = \frac{C}{kA}. \quad (3-12)$$

Die linearisierte Funktion lautet

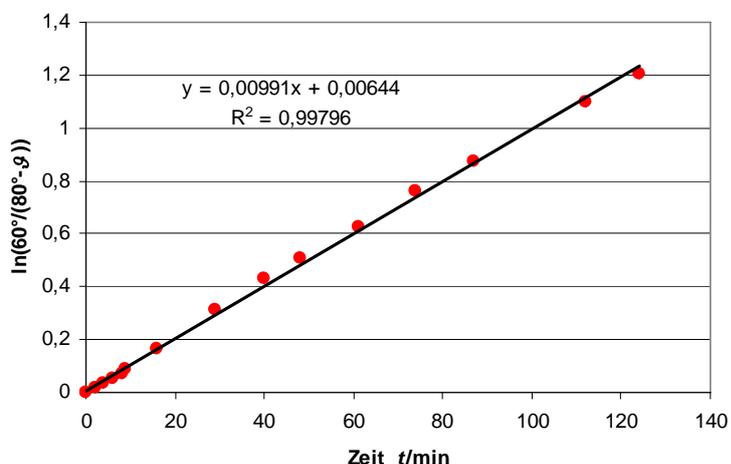
$$\ln \frac{T_O - T_A}{T_O - T} = \frac{1}{\tau} \cdot t.$$

**Experimente:**

1. Erwärmung von ca. 1,2 l Wasser mit der Anfangstemperatur  $T_A = 20 \text{ }^\circ\text{C}$  in einem Ofen mit der Ofentemperatur  $T_O = 80 \text{ }^\circ\text{C}$ .



a)



b)

Bild 3-4: Erwärmung von Wasser, dargestellt in  
a) linearem,  
b) logarithmischem Maßstab.

2. Garen von ca. 1,2 kg Fleisch mit der Anfangstemperatur  $T_A = 20\text{ °C}$  in einem Ofen mit der Ofentemperatur  $T_O = 80\text{ °C}$ .

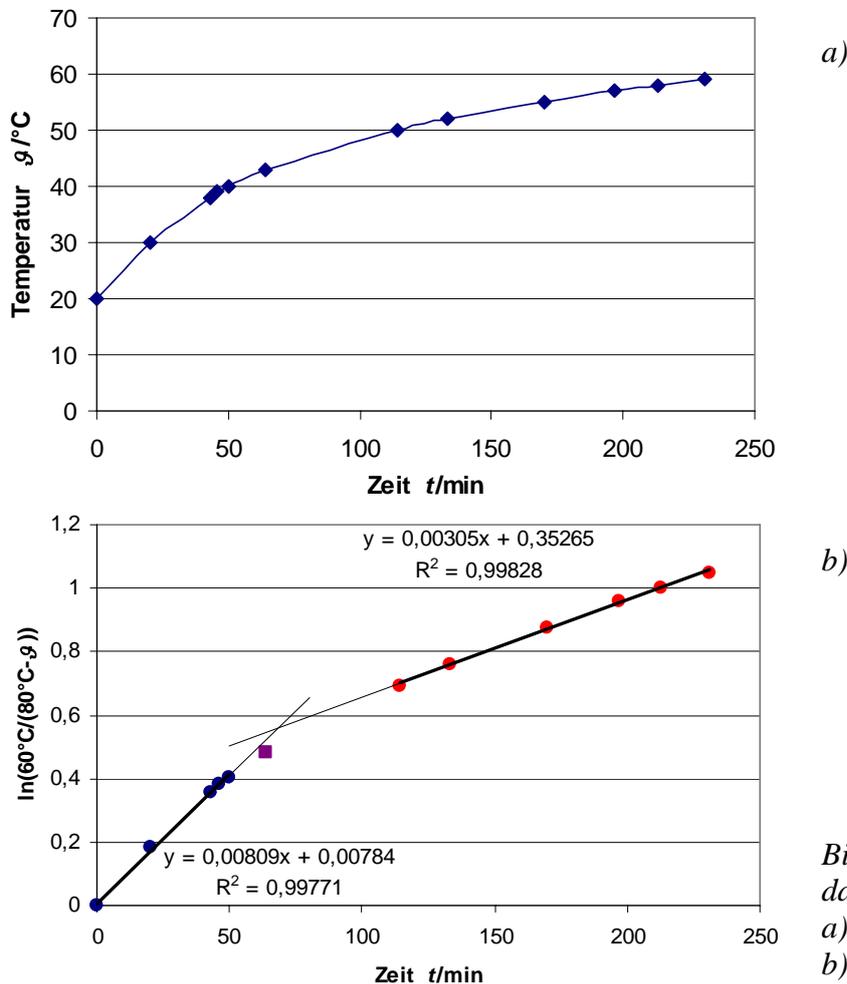


Bild 3-4: Garen von Fleisch, dargestellt in  
a) linearem,  
b) logarithmischem Maßstab.

Aus der logarithmischen Darstellung geht klar hervor, dass zwei Exponentialfunktionen mit unterschiedlichen Zeitkonstanten den Vorgang beschreiben. Das der Gl. (3-12) zugrunde liegende Modell gilt also offenbar beim Fleisch nicht.

Grund: Fleisch ist im Wesentlichen nur ein Proteingemisch. Bei etwa  $45\text{ °C}$  beginnen die Proteine zu „denaturieren“. Sie liegen zunächst in der so genannten Sekundärstruktur vor, wobei Aminosäuren über größere Abstände miteinander wechselwirken (Wasserstoffbrücken, van-der-Waals-Kräfte etc). Bei erhöhter Temperatur lösen sich diese Bindungen auf und die Proteine gehen in die Primärstruktur über, bei der die Aminosäuren wie auf einer Perlenschnur angeordnet sind.