



---

Physikalisches Praktikum I

Umgang mit Messunsicherheiten (Fehlerrechnung)	
Name:	Matrikelnummer:
Fachrichtung:	Versuchsdatum:
Mitarbeiter/in:	Gruppennummer:
Assistent/in:	Endtestat:

Dieser Fragebogen muss von jedem Teilnehmer **eigenständig** (keine Gruppenlösung!) handschriftlich beantwortet und vor Beginn des Versuchs abgegeben werden. Die Vorbereitung wird zusätzlich durch einen Test bzw. eine mündliche Prüfung über die physikalischen Grundlagen des Versuchs kontrolliert.  
(Version: 24. April 2017)

**Versuchsziel und Versuchsmethode:**

1.) Warum ist ein Messergebnis ohne Angabe einer Messunsicherheit nicht aussagekräftig? Erläutern Sie dies am Beispiel Ihrer Körpergröße!

2.) Wie ist der Mittelwert einer Messreihe und wie die Unsicherheit des Mittelwertes definiert? Worin besteht der Unterschied zur Standardabweichung eines Messwertes?

3.) Was versteht man unter Fehlerfortpflanzung?

4.) Was ist lineare Regression? Zeichnen Sie ein Beispiel!

5.) Was bezeichnet man als statistischen Fehler, was als systematischen Fehler? Erläutern Sie dies am Beispiel der Messung Ihrer Körpergröße!

6.) NUR Physiker: Warum sind bei der Poisson-Verteilung Erwartungswert und Varianz identisch? Warum wird durch eine längere Messung die Unsicherheit des Erwartungswerts verkleinert?

# Umgang mit Messunsicherheiten

## Stichworte

Fehler, Unsicherheiten, statistische und systematische Unsicherheiten, Mittelwert, Standardabweichung, Standardabweichung des Mittelwertes, Fehlerfortpflanzung, Normalverteilung, Poisson-Verteilung, Lineare Regression

## Literatur

- [1] John R. Taylor, *Eine Einführung in die Untersuchung von Unsicherheiten in physikalischen Messungen*, Wiley-VCH (1988), ISBN 978-3527268788
- [2] Hans Joachim Eichler, Heinz-Detlef Kronfeldt, Jürgen Sahm, *Das Neue Physikalische Grundpraktikum*, Springer Berlin Heidelberg (2006), ISBN 978-3-540-21453-3 (Print), Online über Campuslizenz [dx.doi.org/10.1007/978-3-662-49023-5](https://dx.doi.org/10.1007/978-3-662-49023-5)
- [3] Lothar Papula, *Mathematik für Ingenieure und Naturwissenschaftler Band 3*, Springer Fachmedien Wiesbaden (2016), ISBN 978-3-658-11923-2 (Print), Online über Campuslizenz [dx.doi.org/10.1007/978-3-658-11924-9](https://dx.doi.org/10.1007/978-3-658-11924-9)
- [4] William Lichten, *Skriptum Fehlerrechnung* Springer (1988), Berlin, ISBN 978-3540183112
- [5] Wilhelm Walcher, *Praktikum der Physik*, Vieweg+Teubner Verlag (2006), Stuttgart, ISBN 978-3835100466

## 1 Messunsicherheiten

Auch wenn in einer Messung sehr auf Genauigkeit geachtet wird, lassen sich **Unsicherheiten** - wir wollen diese Bezeichnung statt der teils missverständlichen Bezeichnung **Fehler** verwenden - nicht vermeiden. Selbst Naturkonstanten sind nur bis auf eine gewisse Unsicherheit bekannt, innerhalb der ihr **wahrer Wert** mit einer hohen Wahrscheinlichkeit liegen wird. Die Fehlerrechnung liefert Methoden und Werkzeuge, mit denen statistische Unsicherheiten einer Messung quantifiziert werden können, beispielsweise über die Bestimmung eines Mittelwertes, einer Standardabweichung oder Berechnung der Fehlerfortpflanzung. Sie umfasst aber auch den Themenbereich der systematischen Unsicherheiten, auf den wir kurz eingehen werden.

In jedem Bereich der Natur- und Ingenieurwissenschaften ist Fehlerrechnung relevant –

ein Messwert ohne Angabe seiner Unsicherheit ist wertlos!

Nehmen Sie z. B. an, dass Sie in einem Versuch die Elementarladung  $e$  bestimmen wollen. In der Literatur (NIST, CODATA 2014) wird die Elementarladung angegeben als

$$e = (1,602\,176\,620\,8 \pm 0,000\,000\,009\,8) \cdot 10^{-19} \text{ C},$$

bzw., kompakter geschrieben,

$$e = 1,602\,176\,620\,8(98) \cdot 10^{-19} \text{ C}.$$

In dem von Ihnen durchgeführten Versuch haben Sie die Elementarladung zu  $e_{\text{prakt}} = 1,6003 \cdot 10^{-19} \text{ C}$  ermittelt. Ist dieser Wert nun „gut“ oder „schlecht“? Ohne die Angabe einer Unsicherheit kann man hierzu keinerlei Aussage treffen, auch wenn der Messwert schon sehr nahe an dem Literaturwert liegt und phantastisch genau zu sein scheint, da fünf Stellen angegeben werden. Nehmen Sie nun beispielhaft an, dass Sie auch eine Unsicherheit für den Messwert bestimmt haben und sich das Ergebnis dann schreibt als  $e = (1,6003 \pm 0,0005) \cdot 10^{-19} \text{ C}$ . Jetzt kann abgeschätzt werden, was das Ergebnis tatsächlich Wert ist - und hier stellt man schnell fest, dass selbst die dreifache Unsicherheit nicht ausreicht, um den ermittelten Wert mit dem Literaturwert vereinbaren zu können. Damit ist das Ergebnis, obwohl vermeintlich sehr präzise gemessen, ein eher schlechtes Ergebnis und man wird sich über weitere Unsicherheitsquellen Gedanken machen müssen!

Im Folgenden werden wir zuerst beschreiben, wie sich systematische und statistische Unsicherheiten voneinander unterscheiden. Wir werden definieren, wie man Mittelwerte und Standardabweichungen berechnet und wir werden zeigen, wie eine Fehlerfortpflanzung bei physikalischen Größen, die wiederum aus mehreren unsicherheitsbehafteten Messgrößen abgeleitet werden, berechnet wird.

## 1.1 Systematische und statistische Unsicherheiten

### 1.1.1 Systematische Unsicherheiten

Systematische Unsicherheiten haben ihre Ursachen im Messaufbau. Sie sind häufig zeitlich konstant, dadurch reproduzierbar und treten bei Wiederholung in gleicher Richtung und Größe auf. Beispiele dafür sind falsch geeichte Skalen, verschobene Null-Stellungen an Messinstrumenten oder Längenänderungen von Skalen durch die Temperatur der Umgebung. Nehmen Sie z. B. an, dass Sie mit einer Federwaage systematisch eine zu große Masse messen, da die Waage eine ermüdete Feder hat. Systematische Unsicherheiten können durch Kontrolle und Verbesserung der Messapparaturen beseitigt oder zumindest verkleinert werden, sie sind aber häufig schwer und oft nur durch Zufall zu entdecken. Auf das Waagenbeispiel bezogen bedeutet dies, dass man z. B. die Waage mit anderen, kali-

brierten Waagen vergleichen oder mit kalibrierten Massen testen müsste, um die falsche Anzeige zu identifizieren.

Systematische Unsicherheiten werden uns im Praktikum z.B. in der Genauigkeits- bzw. Güteklasse von Messgeräten begegnen. Diese ist entweder auf dem Gerät vermerkt oder den technischen Unterlagen zu entnehmen. Güteklasse 1 bei einem analogen Spannungsmesser bedeutet eine Genauigkeit von  $\pm 1\%$  bezogen auf den Vollausschlag des gewählten Messbereichs, z.B. 10 V. Ein Messwert von 0,5 V hat in diesem Beispiel die relativ große Unsicherheit von  $\pm 0,1$  V. Aus diesem Grund ist es wichtig, den Messbereich nur so groß zu wählen, wie der Messwert erfordert. Bei digital anzeigenden Spannungsmessgeräten bezieht sich die Genauigkeitsangabe auf den Messwert und es kommt noch eine Unsicherheit der letzten Ziffernstelle hinzu, z.B.  $\pm(0,3\%$  vom Messwert +2 Digits). Es kommt also auch hier auf die Wahl des passenden Messbereichs an, damit die beiden Digits möglichst wenig Unsicherheit beitragen.

Andere systematische Unsicherheiten könnten sein, dass z.B. eine Temperatur wegen schlechter Wärmeleitung stets zu gering gemessen wird oder bei einer elektrischen Messung durch mangelhafte Isolation ein versteckter Strom fließt und der Messwert deshalb stets zu groß ist.

### 1.1.2 Statistische Unsicherheiten

Zufällige Unsicherheiten lassen sich im Gegensatz zu systematischen nicht vermeiden. Innerhalb einer Messreihe unterscheiden sie sich nach Größe und Betrag. Zufällige Unsicherheiten können beispielsweise entstehen, wenn Sie die Schwingungsdauer eines Pendels mit einer Stoppuhr bestimmen und dabei die Start-/ Stopp-Taste mal etwas zu früh oder etwas zu spät betätigen. Die Auswirkung solcher statistischer Unsicherheiten auf das Messergebnis kann durch Wiederholung der Messung und Mittelwertbildung minimiert werden und, was hier besonders wichtig ist, mathematisch oder durch Abschätzen bestimmt werden.

Bei der Durchführung einer Messung kommt es also darauf an, die Quellen der Messunsicherheit richtig einzuschätzen und den Messvorgang danach auszurichten. So macht es wenig Sinn, die relativ große Spannung einer Batterie mit ein- und demselben Instrument mehrmals zu messen, da statistische Einflüsse sehr klein sind und schnell fluktuieren (thermisches Rauschen). Beim Messen der Intensität eines radioaktiven Materials ist aber Wiederholung bzw. Messung über einen ausreichend langen Zeitraum sehr sinnvoll, da die atomaren Zerfallsprozesse ihrer Natur nach vollkommen statistisch vonstatten gehen.

## 2 Mittelwert, Standardabweichungen und Fehlerfortpflanzung

Bei reproduzierbaren Vorgängen mit statistischen Unsicherheiten bietet sich an, eine Messung nicht nur einmal, sondern mehrmals hintereinander durchzuführen. Dadurch wird die Berechnung des sogenannten Mittelwertes als besten Wert der Messreihe möglich. Bei ausreichend vielen Messvorgängen kann zudem die wahrscheinliche Unsicherheit der einzelnen Datenwerte und die wahrscheinliche Unsicherheit des Mittelwertes ermittelt werden.

### 2.1 Mittelwert $\bar{x}$

Aus einer Reihe von  $n$  verschiedenen Messungen einer Messreihe wird der Bestwert als der arithmetische Mittelwert  $\bar{x}$  berechnet.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

### 2.2 Standardabweichung eines Messwertes einer Messreihe

Die mittlere Unsicherheit der Einzelmessung, auch Standardabweichung oder mittlere Schwankung genannt, ist ein Maß für die Abweichung des Einzelmesswertes  $x_i$  vom Mittelwert  $\bar{x}$ :

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

Die Standardabweichung der Einzelwerte ist die Quadratwurzel der quadratischen Abweichung der Einzelwerte vom Mittelwert, mit dem Faktor  $\frac{1}{n-1}$  normiert. Die Normierung mit  $\frac{1}{n-1}$  statt  $\frac{1}{n}$  ist hierbei notwendig, da es keinen Sinn machen würde im Fall von nur einem vorliegenden Messwert, also  $n = 1$ , eine Unsicherheit zu ermitteln. Diese bleibt deshalb undefiniert, es sind immer mindestens zwei Messungen notwendig, um eine Unsicherheit anzugeben. Das Quadrat der Standardabweichung wird Varianz genannt. Diese Größe ist in Statistik-Auswerteprogrammen zu finden, wird im Praktikum jedoch nicht verwendet.

Mit der Angabe der Standardabweichung wird zugleich ein sogenannter “Vertrauensbereich“ (auch: “Konfidenzintervall“) definiert. Bei normalverteilten Zufallswerten (s. Kapitel 4.1) bedeutet dies, dass ca. 68% aller Messwerte  $x$  im Bereich  $\bar{x} \pm \sigma_{\bar{x}}$  liegen werden – die restlichen 32% aber außerhalb (s. Abbildung 6)!

**Hinweis:** EXCEL enthält verschiedene Varianten der Berechnung der Standardabweichung mit Normierungsfaktor  $\frac{1}{n}$  (STABW.N und STABWNA) und mit Normierungsfaktor

$\frac{1}{n-1}$  (STABW.S und STABWA). Näheres siehe EXCEL-Hilfe.

### 2.3 Standardabweichung des Mittelwertes $\sigma_{\bar{x}}$

Die Standardabweichung des Mittelwertes  $\sigma_{\bar{x}}$  ist um den Faktor  $1/\sqrt{n}$  kleiner als die Standardabweichung  $\sigma_x$ :

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} \quad (3)$$

oder

$$\sigma_{\bar{x}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (4)$$

Nach Abschluss einer Messreihe wird für die bestimmte physikalische Messgröße  $x$  ein Ergebnis in der Form

$$x = \bar{x} \pm \sigma_{\bar{x}} \quad (5)$$

mit dem ermittelten Mittelwert der Messreihe  $\bar{x}$  und der Unsicherheit  $\sigma_{\bar{x}}$  dargestellt. Alternativ ist eine Darstellung mit der Unsicherheit in relativer Form möglich:

$$x = \bar{x} \pm \frac{\sigma_{\bar{x}}}{|\bar{x}|} \cdot 100 \% \quad (6)$$

Die relative Darstellung der Unsicherheit bietet für viele Leser einen intuitiveren Zugang um festzustellen, wie klein oder groß die Unsicherheit des ermittelten Wertes ist.

Die Standardabweichung des Mittelwertes ist die relevante Größe für die Angabe der zufälligen (oder statistischen) Messabweichung. Nur sie sagt aus, innerhalb welchem Unsicherheitsbereich der ermittelte Bestwert, der Mittelwert, streut. Bei Wiederholung der gesamten Messreihe würde man mit 68% Wahrscheinlichkeit erneut einen Mittelwert innerhalb des  $1\text{-}\sigma_{\bar{x}}$ -Vertrauensbereichs erhalten.

Wenn die Standardabweichung des Mittelwertes nur aus wenigen Messwerten gebildet wird, in der Regel nimmt man  $n < 30$ , muss sie mit einem Korrekturfaktor  $t$ , der aus der sogenannten Student-t-Verteilung resultiert, skaliert werden. Der Grund dafür ist, dass die der Berechnung der Standardabweichung zugrundeliegende Annahme, dass die Messwerte normalverteilt sind, bei wenigen Messwerten nicht erfüllt sein muss, beziehungsweise die Normalverteilung für wenige Messwerte unsymmetrisch sein kann. Im Praktikum verzichten wir auf diese zusätzliche Skalierung der Standardabweichung mit dem Student-Faktor, um die Auswertung einfacher zu halten.

## 2.4 Addition von Unsicherheiten

Setzt sich die Unsicherheit einer Messgröße aus voneinander unabhängigen, unkorrelierten Beiträgen zusammen, so wird die Gesamt-Unsicherheit nicht durch arithmetische Addition, sondern durch die Wurzel der quadratischen Addition gebildet, im Fall der **systematischen** und der **statistischen** Unsicherheit also:

$$\sigma_{\bar{x}} = \sqrt{\sigma_{x,\text{stat}}^2 + \sigma_{x,\text{syst}}^2} \quad (7)$$

Diese Art der Berechnung berücksichtigt, dass sich unkorrelierte Unsicherheiten gegenseitig partiell aufheben können und gibt dem größeren Beitrag das größere Gewicht.

## 2.5 Mittelwert und Standardabweichung – ein Beispiel

Ein Fadenpendel (Abbildung 1) wird ausgelenkt und die Periodendauer mit einer Stoppuhr stets am (rechten) Umkehrpunkt gemessen.

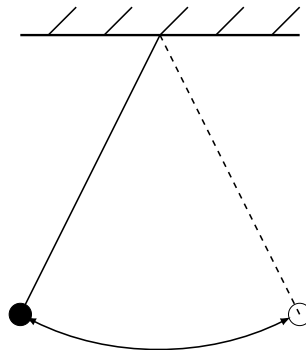


Abbildung 1: Beispiel Fadenpendel: Nach einer Periode wird jeweils am Umkehrpunkt mit einer Stoppuhr die Periodendauer gemessen.

Die Periodendauer wird fünfzigmal hintereinander gemessen. Aufgrund der Reaktionszeit des Experimentators sind die Messwerte leicht unterschiedlich:

$i$	1	2	3	4	5	6	7	8	9	...	49	50
$T_i/\text{s}$	2.44	2.45	2.45	2.34	2.54	2.41	2.26	2.38	2.52	...	2.38	2.47

Als Mittelwert ergibt sich im vorliegenden Fall  $\bar{T} = 2,45$  s und für die Standardabweichung der einzelnen Messwerte  $\sigma_{T_i} = 0,06$  s. Die Standardabweichung des Mittelwertes beträgt  $\sigma_{\bar{T}} = 0,01$  s. Die statistische Unsicherheit des Mittelwertes ist, wie zu erwarten, deutlich geringer als die statistische Unsicherheit der einzelnen Messwerte.



Das Ergebnis lautet deshalb:

$$T = (2,45 \pm 0,01) \text{ s} \quad (8)$$

Die Darstellung der Messwerte als Histogramm zeigt deren statistische Verteilung. Für die Darstellung wurde eine Unterteilung in 7 sog. Klassen mit einer Breite von 0,05 s gewählt. Jeder Balken mit der Breite 0,05 s repräsentiert die Anzahl der in das jeweilige Intervall fallenden Messwerte. Die Abbildung zeigt, dass - trotz der noch geringen Anzahl von nur 50 Messwerten - die Häufigkeitsverteilung annähernd der erwarteten Gaußfunktion folgt.

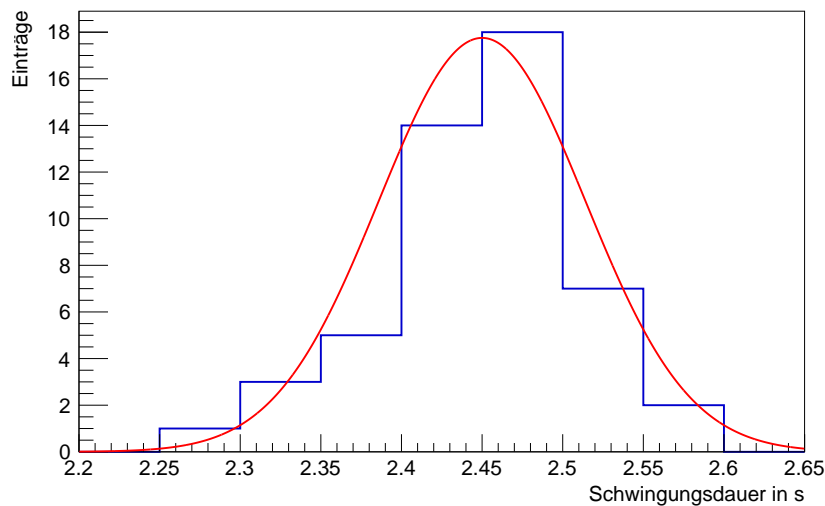


Abbildung 2: Beispiel Fadenpendel: Häufigkeitsverteilung der Messwerte und Gaußfunktion mit einem Maximum bei  $\bar{T}$  und einer Breite von  $\sigma_T$

## 2.6 Signifikante Stellen

Abgesehen von reinen Abzähl-Ergebnissen, die durch natürliche Zahlen ohne Unsicherheit anzugeben sind, liefern Messgeräte Werte mit einer bestimmten Anzahl signifikanter (oder auch gültiger) Stellen: Mithilfe eines Meterstabs mit Millimeterunterteilung kann eine Länge auf z.B. 1873 mm oder 1,873 m oder 0,001 873 km bestimmt werden - immer hat das Ergebnis 4 signifikante Stellen! **Führende Nullen** (egal ob vor oder hinter dem Komma) werden nicht gezählt. Fügen wir aber am **Ende eine Null** hinzu (1873,0 mm oder 1,8730 m), so bedeutet dies eine Erhöhung der Präzision auf Zehntel-Millimeter und deshalb eine Erweiterung auf 5 gültige Stellen. Bei der Darstellung von Messergebnissen ist also darauf zu achten, dass 1,8 m und 1,800 m nicht dasselbe ist! Man geht implizit davon aus, dass die letzte Stelle durch eine Rundung im Bereich der Messunsicherheit

zustande gekommen ist. 1,8 m bedeutet demnach, dass der exakte Wert zwischen 1,75 m und 1,85 m liegt, während die Angabe 1,800 m bedeutet, dass der exakte Wert zwischen 1,7995 m und 1,8005 m liegt - ein erheblicher Unterschied in der Unsicherheit!

Eine physikalische Ergebnisgröße, die aus verschiedenen Messgrößen berechnet wird - z.B. die Geschwindigkeit als Quotient aus Weg und Zeit - kann deshalb nicht mehr signifikante Stellen haben als die ungenaueste Messgröße! Die vielen Stellen, die ggf. bei der Division auf dem Taschenrechner erscheinen, müssen also auf die niedrigste Anzahl signifikanter Stellen gerundet werden.

Dieselben Überlegungen gelten für die Angabe der berechneten Messunsicherheit, die niemals präziser sein kann, als der Messwert selbst. Eine Standardabweichung oder eine systematische Unsicherheit ist deshalb auf den Stellenwert **aufzurunden**, der durch die letzte signifikante Stelle des Messwerts gegeben ist. Die Zahl der signifikanten Stellen eines Mittelwerts ergeben sich wiederum aus der Standardabweichung. Wenn im obigen Beispiel (Fadenpendel) die Messwerte sehr stark streuen und z.B. eine Standardabweichung  $\sigma_{\bar{T}} = 0,13$  s ermittelt wird, d.h. die Unsicherheit bereits die erste Nachkommastelle des Mittelwerts betrifft, macht es keinen Sinn, den Mittelwert mit zwei Nachkommastellen anzugeben! Das Ergebnis muss dann gerundet werden zu:

$$T = (2,5 \pm 0,2) \text{ s} \tag{9}$$

Eine Ausnahme bilden hier Werte, bei denen eine sehr intensive Unsicherheitsanalyse durchgeführt wurde, wie zum Beispiel bei der Bestimmung der Elementarladung am Anfang des Textes. Im Rahmen eines Praktikum wird dies aber nie der Fall sein.

### 3 Fehlerfortpflanzung

Wenn ein Ergebnis aus verschiedenen gemessenen physikalischen Größen abgeleitet wird, die allesamt mit Unsicherheiten behaftet sind, muss berechnet werden, wie sich die Unsicherheiten auf das Ergebnis auswirken (Fehlerfortpflanzung). Nehmen wir als Beispiel an, dass aus einer gemessenen Geschwindigkeit und Masse die kinetische Energie eines Objektes als  $E_{\text{kin}} = \frac{1}{2}mv^2$  bestimmt werden soll. Die kinetische Energie weist nun ebenfalls eine Unsicherheit  $\Delta E_{\text{kin}}$  auf, die eine Funktion der Unsicherheiten  $\Delta m$  sowie  $\Delta v$  ist. Die Fehlerfortpflanzung kann auf zweierlei Weise berücksichtigt werden: Sind statistische Unsicherheiten der Messwerte nicht bekannt, ist die Abschätzung des sog. Größtfehlers möglich. Bei Kenntnis der statistischen Messabweichungen, die normalverteilt und unabhängig voneinander sind, ist die sogenannte Gaußsche Fehlerfortpflanzung anwendbar.

### 3.1 Größtfehlerabschätzung

Der Größtfehler einer Messreihe ist definiert als die Differenz zwischen dem Mittelwert und dem Messwert, der am meisten vom Mittelwert abweicht:

$$\Delta x = \max |x_i - \bar{x}| \quad (10)$$

Die **Größtfehlerabschätzung** ergibt sich aus der Addition der gewichteten Größtfehler der einzelnen Messgrößen. Der Gewichtungsfaktor ist hierbei die partielle Ableitung nach der jeweiligen Messgröße. Dies ist plausibel, da die Ableitung bzw. Steigung einer Funktion ein Maß dafür ist, wie stark sich der  $y$ -Wert ändert, wenn der  $x$ -Wert variiert wird:

$$\Delta y = \sum_{i=1}^n \left| \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} \right| \Delta x_i \quad (11)$$

Hier muss unbedingt unterschieden werden zwischen den statistischen Unsicherheiten der Mittelwerte  $\sigma_{\bar{x}}$ , die wir in Abschnitt 2.3 behandelt haben, und den Größtfehlern  $\Delta x_i$ . Wurde nur eine einzelne Messung durchgeführt, muss der Größtfehler anhand der für das Messgerät spezifizierten Genauigkeit oder der Skalenaufösung abgeschätzt werden.

#### Beispiele

- a) Aus den Messgrößen  $x$  und  $t$  wird eine Geschwindigkeit  $v = \frac{x}{t}$  bestimmt. Gesucht ist die Unsicherheit von  $v$ . Als Mittelwerte und Größtfehler wurden folgende Werte errechnet:

$$\begin{aligned} \bar{x} &= 6,2 \text{ cm} & \Delta x &= 0,1 \text{ cm} \\ \bar{t} &= 3,1 \text{ s} & \Delta t &= 0,1 \text{ s} \end{aligned}$$

Daraus ergibt sich für die Geschwindigkeit  $v = \frac{6,2 \text{ cm}}{3,1 \text{ s}} = 2,0 \frac{\text{cm}}{\text{s}}$ . Die partiellen Ableitungen zur Bestimmung der Unsicherheit der Geschwindigkeit bestimmen sich zu

$$\begin{aligned} \frac{\partial v}{\partial x} &= \frac{1}{t}, \\ \frac{\partial v}{\partial t} &= -\frac{x}{t^2}, \end{aligned}$$

womit für den Größtfehler der Geschwindigkeit folgt

$$\begin{aligned}\Delta v &= \left| \frac{1}{\bar{t}} \right| \cdot \Delta x + \left| -\frac{\bar{x}}{\bar{t}^2} \right| \cdot \Delta t \\ &= \frac{0,1 \text{ cm}}{3,1 \text{ s}} + \frac{6,2 \cdot 0,1 \text{ cm} \cdot \text{s}}{(3,1)^2 \text{ s}^2} = 0,096 \text{ cm/s} \approx 0,1 \text{ cm/s}\end{aligned}$$

Damit folgt als Ergebnis:

$$v = (2,0 \pm 0,1) \text{ cm/s}$$

- b) Die Erdbeschleunigung  $g$  soll aus der Schwingungsdauer  $T$  und der Länge  $l$  eines mathematischen Pendels bei kleinen Auslenkungen des Pendels bestimmt werden. Welche der beiden Größen  $l$  und  $T$  muss mit größerer Sorgfalt gemessen werden?

Die Erdbeschleunigung ist gegeben durch  $g = 4\pi^2 l / T^2$ . Die partiellen Ableitungen zur Bestimmung des Größtfehlers der Erdbeschleunigung sind damit:

$$\begin{aligned}\frac{\partial g}{\partial l} &= \frac{4\pi^2}{T^2}, \\ \frac{\partial g}{\partial T} &= -2 \frac{4\pi^2 l}{T^3}.\end{aligned}$$

Der Größtfehler für die Erdbeschleunigung ergibt sich damit zu

$$\Delta g = 2 \frac{4\pi^2 l}{T^3} \cdot \Delta T + \frac{4\pi^2}{T^2} \cdot \Delta l = 2 \frac{\bar{g}}{T} \cdot \Delta T + \frac{\bar{g}}{l} \cdot \Delta l,$$

bzw. der relative Größtfehler wird zu

$$\frac{\Delta g}{\bar{g}} = 2 \frac{\Delta T}{T} + \frac{\Delta l}{l}.$$

Daraus folgt, dass die Messung der Schwingungsdauer  $T$  besonders sorgfältig erfolgen muss, da die Unsicherheit der Zeitmessung doppelt so stark in den Größtfehler  $\Delta \bar{g}$  von  $\bar{g}$  eingeht wie die Unsicherheit der Längenmessung.

## 3.2 Gaußsche Fehlerfortpflanzung

Die Gaußsche Fehlerfortpflanzung wird angewandt, wenn von allen Komponenten einer physikalischen Größe angenommen werden kann, dass diese aus normalverteilten und voneinander unabhängigen Messwerten gebildet werden, d.h. dass Mittelwerte und Standardabweichungen der Mittelwerte gebildet werden können. Die Unsicherheit  $\sigma_y$  der zu

bestimmenden physikalischen Größe  $f(x_1, \dots, x_n)$  wird dann geschrieben als

$$\sigma_y = \sqrt{\sum_{i=1}^n \left( \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} \right)^2 \sigma_{x_i}^2} \quad (12)$$

Auch in diesem Fall wird die jeweilige Ableitung der zu berechnenden Größe nach ihren Komponenten durchgeführt, aber eine quadratische Summe der jeweiligen Komponenten multipliziert mit dem Quadrat der Standardabweichungen  $\sigma_{x_i}$  gebildet. [1] demonstriert, aus welchem Grund dies zur Bestimmung des fortgepflanzten Fehlers zulässig und sinnvoll ist. Es gilt stets

$$\sigma_y \leq \Delta y, \quad (13)$$

d.h. die statistische Unsicherheit, die aus der Gaußschen Fehlerfortpflanzung resultiert, wird stets kleiner oder gleich dem Größtfehler nach Gleichung 11 sein.

### Beispiel

Erneut soll aus den Messgrößen  $x$  und  $t$  eine Geschwindigkeit  $v = \frac{x}{t}$  bestimmt werden. Gesucht ist die Unsicherheit  $\sigma_v$ . Mittelwerte und Standardabweichungen wurden berechnet zu:

$$\begin{aligned} \bar{x} &= 6,2 \text{ cm} & \sigma_x &= 0,1 \text{ cm} \\ \bar{t} &= 3,1 \text{ s} & \sigma_t &= 0,1 \text{ s} \end{aligned}$$

Die partiellen Ableitungen zur Bestimmung der Unsicherheit der Geschwindigkeit bestimmen sich zu

$$\begin{aligned} \frac{\partial v}{\partial x} &= \frac{1}{t}, \\ \frac{\partial v}{\partial t} &= -\frac{x}{t^2}, \end{aligned}$$

womit für die Unsicherheit der Geschwindigkeit folgt:

$$\begin{aligned} \sigma_y &= \sqrt{\left( \frac{1}{\bar{t}} \right)^2 \cdot \sigma_x^2 + \left( \frac{\bar{x}}{\bar{t}^2} \right)^2 \cdot \sigma_t^2} \\ &= 0,07 \frac{\text{cm}}{\text{s}} \approx 0,1 \frac{\text{cm}}{\text{s}} \end{aligned}$$

Die statistische Unsicherheit der Geschwindigkeit ist demnach kleiner oder, aufgerundet, gleich dem Größtfehler, womit Gleichung 13 erfüllt ist.

### 3.3 Lineare Regression

Erwartet man einen linearen Zusammenhang zwischen zwei Messgrößen, ist es sinnvoll, eine Schaar von Messwertpaaren grafisch darzustellen und eine sog. Regressions- oder Ausgleichsgerade als Bestwert der Analyse durch die Messwerte zu legen. Da ein linearer Zusammenhang eindeutig zu erkennen ist, werden andere Abhängigkeiten, z.B. quadratische oder exponentielle, durch entsprechende mathematische Umkehroperationen (Radizieren, Logarithmieren) in einen linearen Zusammenhang gebracht und die Messwerte entsprechend dargestellt. Die Berechnung einer Ausgleichsgeraden dient also einerseits der Ermittlung des Proportionalitätsfaktors (der Steigung), andererseits auch dem Nachweis, dass grundsätzlich ein linearer (oder anderer) Zusammenhang existiert. Durch den Einsatz von Analyse-Software ist es allerdings zunehmend einfacher geworden, eine beliebige Zielfunktion direkt an einen Messdatensatz anzupassen („fitten“) und die Anpassparameter als Ergebnis anzugeben. Es macht allerdings auch dabei nur Sinn, eine physikalisch motivierte Zielfunktion zu verwenden - und **nicht** ein beliebiges Polynom, mit dem man einen „Elefanten“ anpassen kann! Im Folgenden wird die lineare Regression mittels des grafischen Verfahrens beschrieben sowie der mathematische Zugang dargestellt, der den meisten „Fit“-Programmen zugrunde liegt.

#### 3.3.1 Grafisches Verfahren

Es bestehe ein linearer Zusammenhang  $y = a_0 + a_1 \cdot x$  zwischen den Messgrößen  $x$  und  $y$ , gesucht sind  $a_0$  und  $a_1$ . Mittels eines grafischen Verfahrens können nun die beiden gesuchten Größen bestimmt werden. Dazu werden die gemessenen Wertepaare in ein Diagramm eingetragen und bei bekannter Unsicherheit mit sogenannten **Fehlerbalken** versehen. Sind beide Messgrößen  $x$  und  $y$  mit Unsicherheiten behaftet, müssen diese durch Fehlerkreuze dargestellt werden. Durch die streuenden Messpunkte legt man mit Hilfe eines durchsichtigen Lineals nach Augenmaß eine **Ausgleichsgerade** so, dass etwa gleich viele Messpunkte oberhalb und unterhalb der Geraden liegen. Aus der **Steigung** der Geraden bestimmt man  $a_1$ , aus dem Schnittpunkt mit der Ordinate ( $x = 0$ ) erhält man  $a_0$ . Für die Berechnung der Steigung wähle man zwei Punkte in möglichst großem Abstand auf der Ausgleichsgeraden (nicht zwingend Messpunkte verwenden!). Der **Fehler der Steigung** folgt in grober Abschätzung aus den mit den Messpunkten zu vereinbarenden extremen Geradensteigungen.

#### 3.3.2 Mathematische Vorgehensweise ohne Betrachtung individueller Unsicherheiten

Erneut wird ein linearer Zusammenhang  $y = a_0 + a_1 \cdot x$  zwischen den gemessenen Datenpunkten  $(x_1, y_1), \dots, (x_n, y_n)$  vermutet. [1] führt aus, wie aus diesen Datenpunkten mit

Hilfe der Methode der kleinsten Quadrate, dem sogenannten **least square fitting**, auf die zu ermittelnden Koeffizienten  $a_0$  und  $a_1$  geschlossen werden kann. Hierzu wird die Summe der Quadrate der Abweichungen zwischen den Messwerten  $y_i$  und dem von der Ausgleichsgeraden vorgegebenem Werten  $y(x_i)$  betrachtet :

$$\chi^2 = \sum_{i=1}^n (y_i - y(x_i))^2 \quad (14)$$

Grundidee ist dabei, dass die Ausgleichsgerade so gewählt wird, dass  $\chi^2$  ein Minimum anstrebt. Wir wollen das an dieser Stelle nicht in der Ausführlichkeit durchführen sondern direkt zum Ergebnis kommen.  $a_0$  bestimmt sich zu

$$a_0 = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{\Delta}, \quad (15)$$

sowie  $a_1$  zu

$$a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\Delta}. \quad (16)$$

Dabei ist

$$\Delta = n \sum x_i^2 - (\sum x_i)^2 \quad (17)$$

und  $n$  die Anzahl der Datenpunkte. Alle Summen laufen von  $i = 1, \dots, n$ . Für die Bestimmung der Koeffizienten  $a_0$  und  $a_1$  ist hierbei angenommen worden, dass alle Messwerte keine individuellen Unsicherheiten  $\sigma_{y_i}$  haben, d. h. jeder Messwert hat zwar eine Unsicherheit, die anhand der Formel

$$\sigma_y = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (y_i - a_0 - a_1 \cdot x_i)^2} \quad (18)$$

berechnet wird, diese ist aber für alle Messwerte gleich und insbesondere für die Bestimmung von Steigung und Ordinatenschnittpunkt nicht relevant! Dies ist ein vereinfachter Fall. Die aus Gleichung 18 bestimmte „Unsicherheit“ fließt dagegen in die Fehlerfortpflanzung zur Bestimmung der Unsicherheit der beiden Koeffizienten  $a_0$  und  $a_1$  ein:

$$\sigma_{a_0} = \sqrt{\frac{\sigma_y^2}{\Delta} \sum x_i^2} \quad (19)$$

sowie

$$\sigma_{a_1} = \sqrt{\frac{\sigma_y^2}{\Delta}} n. \quad (20)$$

### 3.3.3 Mathematische Vorgehensweise mit Betrachtung individueller Unsicherheiten

Der oben besprochene Fall der linearen Regression zog nicht in Betracht, dass die Messwerte individuelle Unsicherheiten haben können. Dies ist z. B. schon dann der Fall, wenn jeder der Messpunkte, die in einem linearen Zusammenhang stehen sollen, in Wirklichkeit der Mittelwert einer Reihe von Messwerten ist. Dann hat jeder Punkt eine Standardabweichung  $\sigma_{y_i}$ , die jeweils unterschiedlich sein wird. Die Folge ist, dass diese Standardabweichungen sowohl in die Bestimmung der Koeffizienten als auch der Unsicherheit der Koeffizienten einfließen wird. Die individuelle Bedeutung der Unsicherheiten der Messpunkte wird durch Gewichte

$$w_i = \frac{1}{\sigma_{y_i}^2} \quad (21)$$

festgelegt. Im Folgenden ergeben sich die Koeffizienten  $a_0$  und  $a_1$  zu

$$a_0 = \frac{\sum w_i x_i^2 \sum w_i y_i - \sum w_i x_i \sum w_i x_i y_i}{\Delta}, \quad (22)$$

sowie

$$a_1 = \frac{\sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i}{\Delta}. \quad (23)$$

Dabei ist hier

$$\Delta = \sum w_i \sum w_i x_i^2 - \left( \sum w_i x_i \right)^2 \quad (24)$$

Alle Summen laufen von  $i = 1, \dots, n$ . Um die Unsicherheiten auf die Koeffizienten zu bestimmen, muss erst einmal  $\sigma_y$  berechnet werden, welche leicht abgewandelt zur Gleichung 18 ist:

$$\sigma_y = \sqrt{\frac{1}{n-2} \sum_{i=1}^n \left( \frac{y_i - a_0 - a_1 \cdot x_i}{\sigma_{y_i}} \right)^2} \quad (25)$$

Die Unsicherheiten auf die Koeffizienten ergeben sich dann zu:

$$\sigma_{a_0} = \sqrt{\frac{\sum w_i x_i^2}{\Delta}} \sigma_y^2 \quad (26)$$



sowie

$$\sigma_{a_1} = \sqrt{\frac{\sum w_i}{\Delta}} \sigma_y^2. \quad (27)$$

Auch diese Bestimmung der Koeffizienten und Unsicherheiten der Koeffizienten ist ein Spezialfall, da hier keine Unsicherheiten  $\sigma_{x_i}$  betrachtet werden. Diese Unsicherheiten für die  $x$ -Komponente der Messpunkte sind mit Sicherheit vorhanden, werden im Rahmen des physikalischen Anfängerpraktikums allerdings nicht betrachtet.

### 3.3.4 Beispiel zur linearen Regression

In diesem Beispiel soll die mathematische Methode für die Anpassung einer linearen Funktion an Messwerte, die zur Bestimmung der Austrittsarbeit einer Röhrendiode aufgenommen wurden, gezeigt werden. Es werden individuelle Unsicherheiten  $\sigma_{y_i}$  angenommen, d. h. es alle Messpunkte sind mit einer Unsicherheit in der  $y$ -Komponente versehen. Diese Unsicherheit drückt sich als Fehlerbalken im zu zeichnenden Diagramm (Abbildung 3) aus. Im Praktikum sollen stets Fehlerbalken in Diagramme eingezeichnet werden!

Um die Austrittsarbeit der Röhrendiode bestimmen zu können, wird die sogenannte Richardson-Gleichung verwendet, die einen exponentiellen Zusammenhang zwischen dem Sättigungsstrom der Diode  $I_S$ , der Austrittsarbeit  $E_\varphi$  sowie der Temperatur  $T$  zeigt:

$$I_S(x) = AT^2 \exp\left(-\frac{E_\varphi x}{k_B T}\right).$$

Durch Logarithmieren wird dieser exponentielle Zusammenhang linearisiert und man erhält

$$\ln(I_S/T^2) = \left(-\frac{E_\varphi x}{k_B T}\right) + \ln A. \quad (28)$$

Damit ist die Struktur eines linearen Funktion  $y = a_0 + a_1 \cdot x$  ersichtlich, im vorliegenden Fall ist

$$\begin{aligned} a_0 &= \ln A, \\ a_1 &= E_\varphi, \end{aligned}$$

womit die Geradensteigung unmittelbar die Austrittsarbeit definiert. Die folgende Tabelle zeigt Daten, die an einer Röhrendiode aufgenommen wurden, und mit denen die Methode der kleinsten Quadrate angewandt werden kann:

$T/(K)$	1389	1425	1473	1483	1519
$I_S/(mA)$	0,35	0,67	1,3	1,6	2,71
$\frac{I_S}{T^2}/(10^{-6} mA/K^2)$	0,181	0,330	0,600	0,728	1,174
$\frac{x}{k_B T}/(1/eV^{-1})$	8,35	8,14	7,88	7,82	7,64
$\sigma_{\frac{I_S}{T^2}}/(10^{-6} mA/K^2)$	0,007	0,017	0,028	0,036	0,041

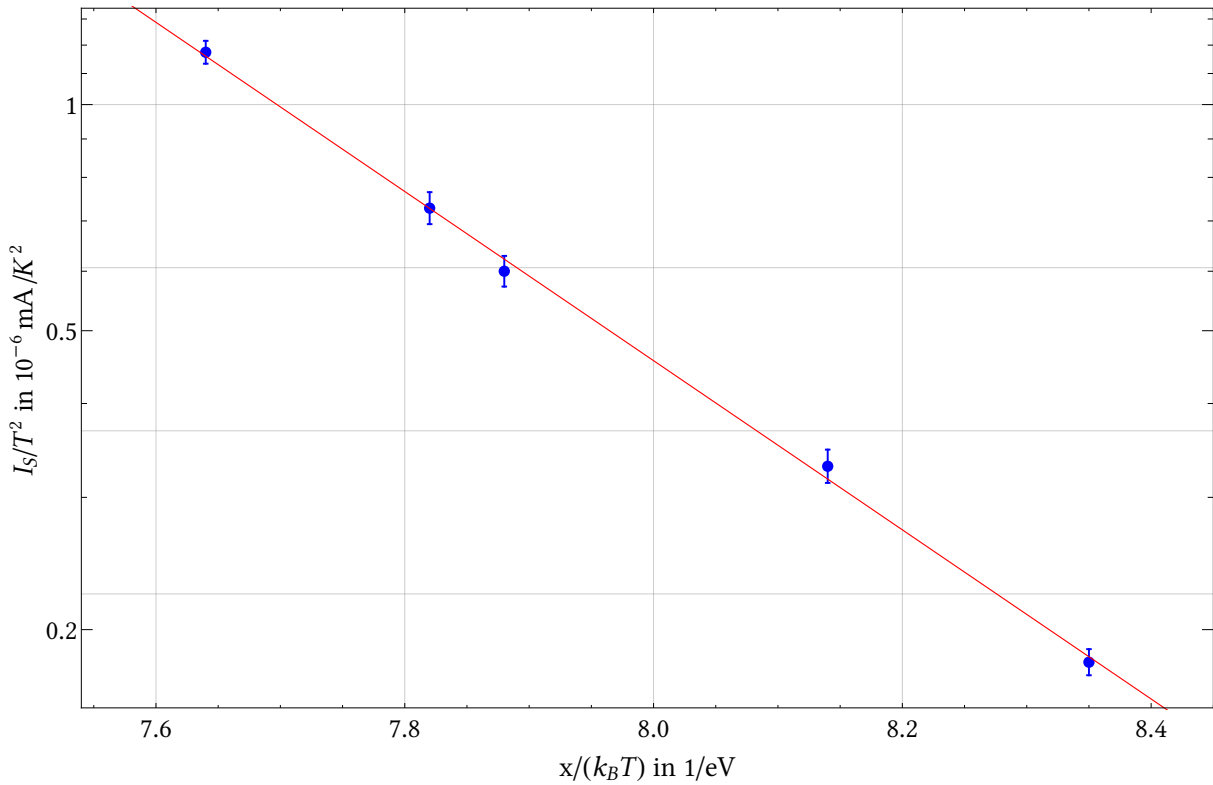


Abbildung 3: Richardson-Gleichung:  $I_S/T^2$  logarithmiert als Funktion von  $1/(k_B T)$  für eine Röhrendiode aufgetragen. Über die Steigung der angepassten linearen Funktion wird die Austrittsarbeit der Röhrendiode ermittelt.

Aus diesen Daten ergeben sich nach Gleichungen 22 und 23 unter Berücksichtigung der jeweiligen Unsicherheiten auf jeden Datenpunkt

$$a_0 = 19,96 \cdot 10^{-6} \text{ mA/K}^2,$$

$$a_1 = -2,59 \text{ eV}.$$

Unter Anwendung von Gleichungen 25, 26 und 27 erhält man die Unsicherheiten auf die

y-Werte und im Folgenden für den Achsenabschnitt und die Steigung:

$$\sigma_{a_0} = 0,50 \cdot 10^{-6} \text{ mA/K}^2,$$

$$\sigma_{a_1} = 0,06 \text{ eV}.$$

Damit ist Ausgleichsrechnung abgeschlossen. Für die Austrittsarbeit der Röhrendiode erhält man damit

$$E_{\varphi} = (2,59 \pm 0,06) \text{ eV}.$$

Abbildung 3 visualisiert die ermittelte Ausgleichsgerade und die Datenpunkte.

## 4 Gauß- und Poissonverteilungen

### 4.1 Gaußverteilung

Ein wesentlicher Bestandteil der gesamten vorherigen Argumentation und der Methoden zur Fehlerrechnung ist, dass die ermittelten Datenpunkte, aus denen Mittelwerte, Standardabweichungen und Fehlerfortpflanzungen berechnet werden, normal- oder gaußverteilt sind. Die Gaußverteilung ist gegeben durch

$$G_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right). \quad (29)$$

Die Verteilung bildet die bekannte Glockenkurve um einen Mittelwert  $\mu$  und weist ei-

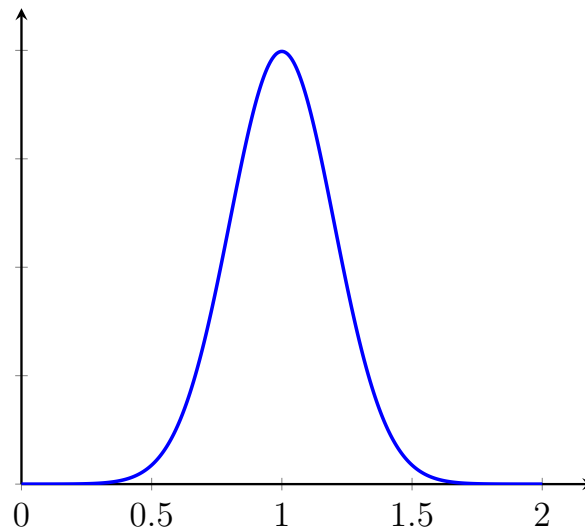


Abbildung 4: Gaußsche Verteilung um einen Mittelwert  $\mu = 1$  mit Breite  $\sigma = 0.2$ .

ne Breite von einer Standardabweichung  $\sigma$  auf. Eine Beispielverteilung ist in Abbildung 4 dargestellt. In der Abbildung 5 sind Gaußverteilte Zufallswerte in Form zweier Histogramme gezeichnet, in einem Fall 10 Werte, im anderen Fall 1000 Werte, die sich um einen Mittelwert von  $\mu = 1$  mit Standardabweichung  $\sigma = 0.1$  befinden. Ebenfalls sind dazugehörige Gaußverteilungen eingezeichnet. Damit kann man gut sehen, dass Gaußverteilte Zufalls- oder Messwerte durch die entsprechende Gaußverteilung abgebildet werden können, und überdies ist auch gut zu sehen, dass je mehr Daten, in diesem Fall zufällig ausgewürfelte, vorhanden sind, die Daten immer besser einer Gaußverteilung entsprechen. Darauf sind wir bereits in Abschnitt 2.3 eingegangen, in dem erwähnt wurde, dass eigentlich für wenige Gaußverteilt angenommene Messwerte ein Korrekturfaktor nötig ist, wenn eine Standardabweichung angegeben werden soll. Dieser Korrekturfaktor  $t$  wird bei vielen

Messwerten hinfällig.

Mittels der Gaußverteilung lässt sich nun die Wahrscheinlichkeit  $p$  ausrechnen, mit der ein Messwert  $x_i$  um den Mittelwert  $\mu$  zu finden sein wird. Die Wahrscheinlichkeit ist gegeben als die Fläche unter der Gaußfunktion in einem Intervall  $\mu \pm \sigma$ :

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu-\sigma}^{\mu+\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx. \quad (30)$$

Durch Substitution erhält man das bekannte Gaußsche Fehlerintegral

$$p(z) = \frac{1}{\sqrt{2\pi}} \int_{-t}^t \exp\left(-\frac{z^2}{2}\right) dz, \quad (31)$$

wobei  $t$  Vielfache der Standardabweichung  $\sigma$  sind. Die Evaluierung dieses Integrals liefert nun die gewünschten Wahrscheinlichkeiten, dass ein Messwert innerhalb von  $t\sigma$  um den Mittelwert herum zu finden sein wird (siehe Tabelle 1).

$t$	1,0	2,0	3,0	4,0	5,0
$p$ in %	68,3	95,4	99,7	99,99	99,999 9

Tabelle 1: Wahrscheinlichkeit, innerhalb derer ein gaußverteilt angenommener Messwert in einem Intervall von  $\pm t\sigma$  um einen Mittelwert  $\mu$  anzutreffen sein wird.

Aus Tabelle 1 lässt sich also entnehmen, dass mit einer Wahrscheinlichkeit von 68,3 % ein Messwert innerhalb einer Standardabweichung um den Mittelwert liegen wird. Mit einer Wahrscheinlichkeit von 99,7 % wird er dagegen in der dreifachen Standardabweichung vorzufinden sein. Im Umkehrschluss heißt dies, dass ein Wert, der außerhalb der dreifachen Standardabweichung liegt, nur mit einer Wahrscheinlichkeit von 0,3 % vorzufinden sein wird! Dies ist also sehr unwahrscheinlich. Nehmen wir hier noch einmal das Beispiel vom Beginn aus Kapitel 1: Es wurde angenommen, dass die Elementarladung, vermeintlich sehr präzise, gemessen wurde zu  $e_{\text{prakt}} = (1,6000 \pm 0,0005) \cdot 10^{-19}$  C. Der Literaturwert, mit der gleichen Anzahl gültiger Stellen wie der Mittelwert dargestellt, ist  $e = 1,6022 \cdot 10^{-19}$  C. Die Differenz der beiden Werte ist also  $\delta e = 2,2 \cdot 10^{-22}$  C. Die Standardabweichung der Messung ist aber  $\sigma_e = 5 \cdot 10^{-23}$  C, der ermittelte Wert ist demnach von dem Literaturwert  $4.4\sigma$  entfernt. Es ist also äußerst unwahrscheinlich, dass Literaturwert und Messergebnis zusammen passen, auch wenn bei dieser hypothetischen Messung eine sehr hohe Präzision erreicht worden ist. In diesem Fall ist der Unterschied zwischen theoretischem und praktischem Wert signifikant und es müsste untersucht werden, ob nicht weitere Unsicherheiten vorhanden sind. Wäre der ermittelte Wert innerhalb eines  $1\sigma$ - oder auch noch  $2\sigma$ -Intervalls passend zum erwarteten Literaturwert, dann wäre der

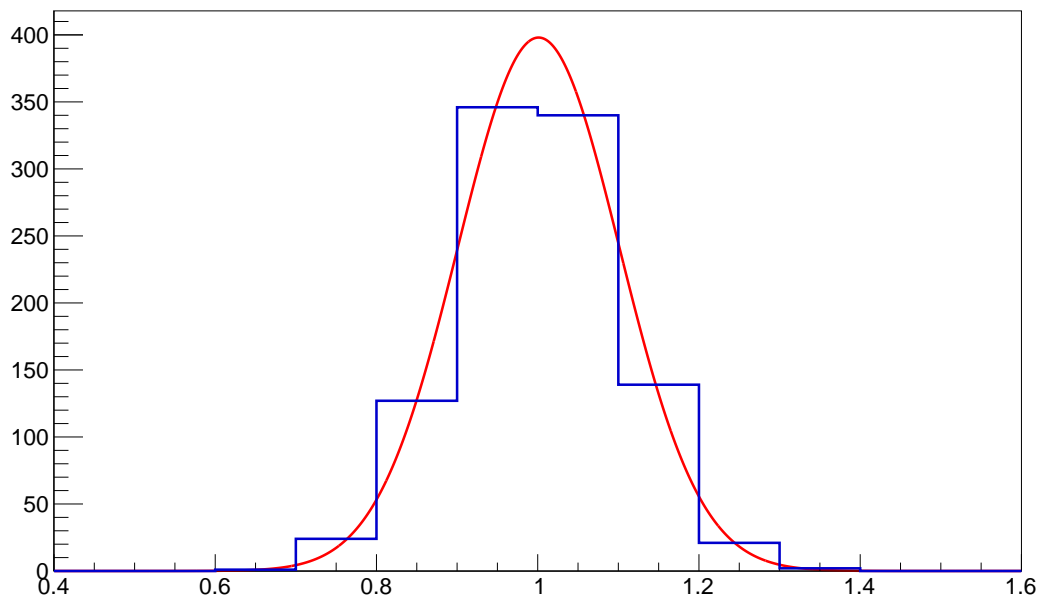
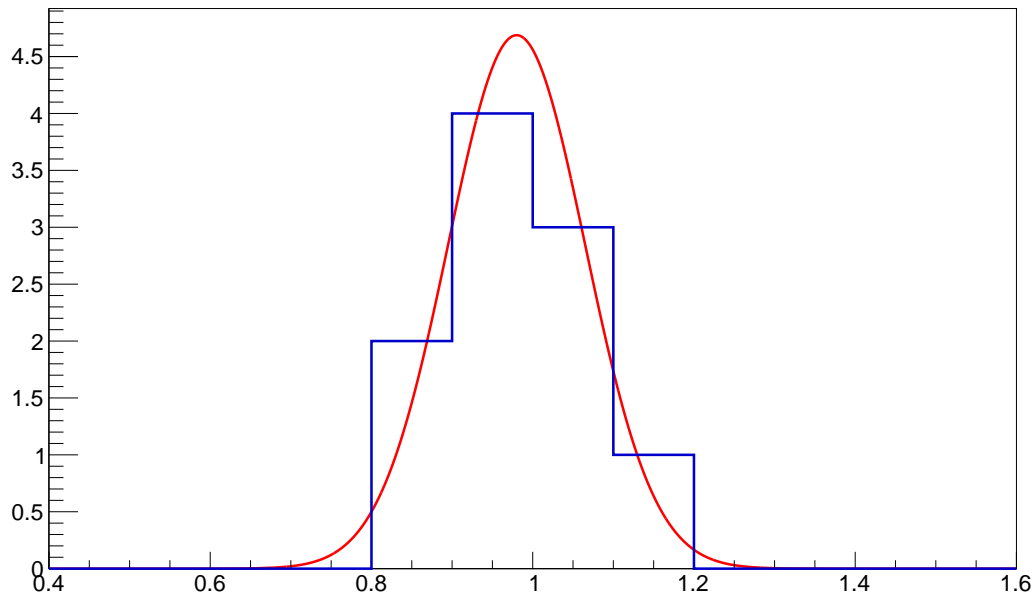


Abbildung 5: Gaußverteilte Zufallswerte und die Gaußverteilungen, welche die Daten beschreiben. Oben sind zehn Zufallswerte histogrammiert worden und durch eine Gaußverteilung angepasst, unten eintausend derer. Je mehr Daten, in diesem Fall zufällig ausgewürfelte, desto besser entsprechen die Daten eine Gaußverteilung.

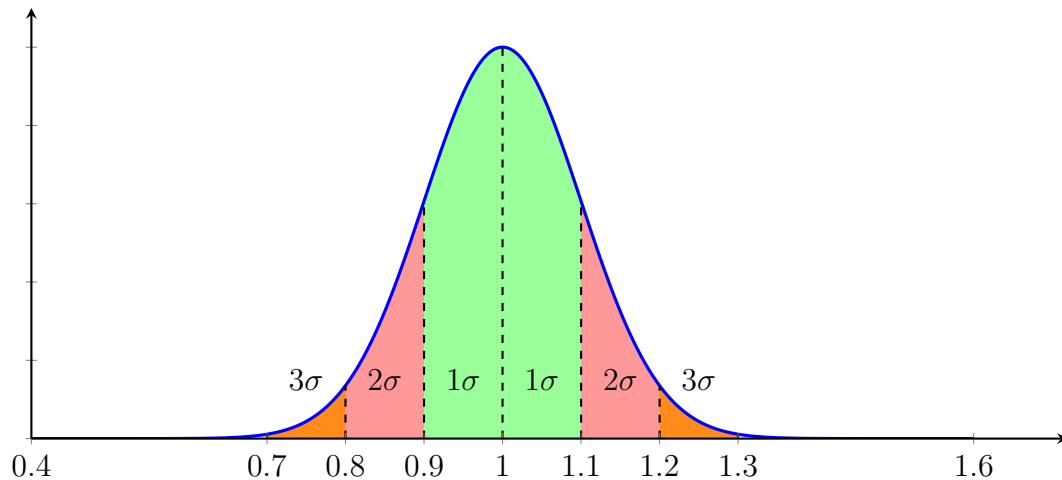


Abbildung 6: Gaußsche Verteilung um einen Mittelwert  $\mu = 1$  und Standardabweichung  $\sigma = 0,1$ .  $1\sigma$ -,  $2\sigma$ - und  $3\sigma$ -Intervalle sind farbcodiert in blau, rosa und orange markiert. Innerhalb des  $1\sigma$ -Intervalls werden 68,3% aller Messwerte um den Mittelwert zu erwarten sein, im  $2\sigma$ -Intervall 95,4% und im  $3\sigma$ -Intervall 99,7%.

Unterschied nicht signifikant sondern würde von den allermeisten Wissenschaftlern nicht weiter beachtet werden.

Es ist wichtig, sich noch einmal diese Konzepte klar zu machen: In der Regel gehen wir, besonders im Praktikum, von gaußverteilten Messwerten um einen Mittelwert aus. Mit der Standardabweichung der Messwerte können wir mit fest definierter Wahrscheinlichkeit sagen, in welchem Intervall um den Mittelwert sich die Messwerte (und weitere, hätten wir sie aufgenommen) befinden sollten. Sobald wir zusätzlich die Standardabweichung des Mittelwertes berechnet haben, können wir mit einer fest definierten Wahrscheinlichkeit sagen, in welchem Intervall um den Mittelwert herum der **wahre Wert** liegen sollte. Das alles liefert die Gaußverteilung. Sie ermöglicht es auch direkt herauszufinden, ob ein im Experiment bestimmter Mittelwert gut zu einem theoretischen Wert passt oder nicht. Andersherum - ob ein theoretischer Wert gut experimentell abgebildet werden kann - geht es natürlich auch.

## 4.2 Poisson-Verteilung

Die Poisson-Verteilung ist, im Gegensatz zur Gaußverteilung, eine diskrete Wahrscheinlichkeitsverteilung. Sie wird benutzt, wenn selten und unabhängig voneinander auftretende Ereignisse modelliert werden sollen, die dennoch im Mittel eine konstante Rate aufweisen.

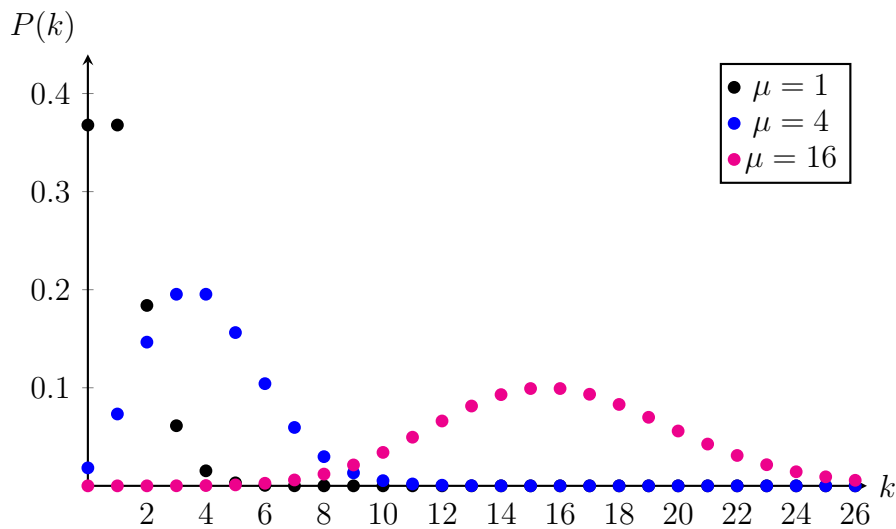


Abbildung 7: Verteilungen für drei verschiedene Parameter  $\mu$ . Je größer  $\mu$  wird, desto geringer wird die Asymmetrie der Verteilung und desto mehr ähnelt die Poisson-Verteilung einer Gaußverteilung.

Als Beispiele können radioaktive Zerfallsprozesse genommen werden, die im Mittel zwar konstant wirken, aber statistisch sind, also in kleinen Zeiträumen zufällig auftreten werden, mal mehr, mal weniger Ereignisse. Ein interessantes Beispiel aus dem Alltag ist die Vorhersage von geschossenen Toren in einem Fußballspiel. Im Mittel wird in jedem Spiel eine konstante, kleine Anzahl an Toren geschossen werden, wobei es auch Spiele geben wird, in denen keine und Spiele, in denen viele geschossen werden. Dies kann man mittels der Poisson-Statistik sehr gut modellieren.

Mathematisch ist die Verteilungsfunktion der Poisson-Statistik gegeben durch

$$p_{\mu}(k) = \frac{\mu^k}{k!} \exp(-\mu), \quad (32)$$

und in Abbildung 7 für drei verschiedene Parameter  $\mu$  als Funktion von  $k$  gezeichnet. Die Verteilungsfunktion weist eine Asymmetrie auf, die für höhere Parameter  $\mu$  verschwindet: Die Poisson-Verteilung geht dann in eine Gaußverteilung über. Die Verteilungsfunktion weist ein wichtiges Detail auf: Der Parameter  $\mu$  ist sowohl der Erwartungswert („Mittelwert“) der Verteilung, er ist aber zugleich auch die Varianz. Dies bedeutet, dass im Mittel  $\mu$  Ereignisse gezählt werden, und die Unsicherheit auf diese Ereignisse, die sich als die Quadratwurzel aus der Varianz ergibt, die Quadratwurzel des Erwartungswertes ist! Dies ist ein wichtiger Aspekt, denn er sagt unmittelbar, dass längere Messungen bei poissonverteilten Ereignissen zu einer kleineren Unsicherheit führen.

- Nehmen Sie an, es wären über einen Zeitraum von 5 min genau 100 Ereignisse bei-



spielsweise eines radioaktiven Zerfalls gemessen worden. Damit ergäbe sich eine Zählrate von

$$Z_{5 \text{ min},1} = 100 \pm \sqrt{100} = 100 \pm 10 ,$$

bzw. heruntergerechnet auf eine Minute

$$Z_{1 \text{ min},1} = 20 \pm 2 = 20 \pm 10\% .$$

- Messen Sie nun eine Stunde lang, und werden in diesem Zeitraum 1200 Ereignisse gemessen, dann erhält man eine Zählrate von

$$Z_{60 \text{ min},2} = 1200 \pm \sqrt{1200} = 1200 \pm 35 ,$$

bzw. ebenfalls auf eine Minute heruntergerechnet

$$Z_{1 \text{ min},2} = 20,0 \pm 0,6 = 20 \pm 3\% .$$

Damit ist gut erkennbar, dass die Unsicherheit mit längeren Messzeiten bei poissonverteilten Ereignissen erheblich abnimmt.