Peter Wessels

Zum Umgang mit gemessenen Daten im Physikunterricht.

Drei Gründe gibt es dafür, sich mit dem Thema zu befassen:
- Unterrichtsaspekte haben sich verschoben,
- Automatisierte Messtechniken finden zunehmend Anwendung, auch in der Schule,
- unter dem Kürzel GUM gibt es einen Vorschlag zur Änderung der Begriffe beim Auswerten von Messungen.

Unverzichtbarer Bestandteil des Physikunterrichtes sind noch immer reale Experimente, entweder vom Lehrer vorgeführt, oder von Klassen durchgeführt. Sehr häufig sind es Experimente mit quantitativem Charakter. Aus den durch Messungen gewonnenen Daten sollen im anschließenden Unterricht Schlüsse gezogen werden.

Während der kommentierenden Diskussion von Messungen werden von den Beteiligten – Lehrern, wie Lernenden – mitunter Formulierungen verwendet, die eher problematisch erscheinen, wie etwa

- das sind gute Ergebnisse,
- ungefähr ist das wohl richtig,
- das sieht gut aus, aber das dritte Wertepaar ist falsch,
- genauer können wir das in der Schule nicht,
- wir haben nur alte Geräte. Unser Fehler ist also 0,7 m/s² .
- Die Länge der Strecke habe ich zu 9,55 m gemessen. Der Messfehler beträgt etwa 1%.

Aus meiner Sicht sind solche Teile der Unterrichtspraxis zwar nicht tolerabel, aber verständlich.- Wir versuchen in der Schule zu oft das abzubilden, was wir im Studium gelernt haben. Und da ist es Faktum, dass im Praktikum gezeigt wird, dass viele Einzelmessungen zu einem „besseren“ Ergebnis führen. Die dabei verwendeten statistischen Verfahren benötigen aber viele Einzelmessungen.
Im Physikpraktikum der Universität stehen etwa 3 Zeitstunden für ein Experiment zur Verfügung. In der Schule – SI – wäre das ein qualifizierter Prozentsatz der Unterrichtszeit eines Schulhalbjahres. Also muss Schule anders arbeiten.
Von Bedeutung ist sicherlich auch, dass auch in den Hochschulen sehr variabel mit Begriffen gearbeitet wird. Die folgende Begriffsliste hat S. Heinicke, Aus Fehlern wird man klug, Diss. 2011 zusammengestellt:

(Mess-)Fehler	(Mess-)Unsicherheit	Fehlergrenzen
systematische Fehler	zufällige Schwankungen	zufällige Messfehler
Rauschen	Präzision	Abweichung
Standardabweichung	mittlerer (absoluter) Fehler	Genauigkeit
(mittlere) absolute Unsicherheit	relative Messunsicherheit	relativer Fehler
vermeidbare Fehler	nicht vermeidbare Unsicherheit	falsche Ablesung
exakt		

In dieser Arbeit wurde ebenfalls untersucht, wie diese Begriffe in verschiedenen Hochschullehrbüchern verwendet werden. - Von einheitlicher Verwendung der Begriffe kann nicht gesprochen werden.

2008 wurde von einem Joint Committee for Guides in Metrology ein „Guide to the expression of uncertainty in measurements“ (GUM) vorgestellt, der zu international einheitliche Begriffen und Verfahren führen soll. - Die Vorschläge sind bisher noch nicht in deutsches Recht übernommen. Trotzdem ist die Diskussion für die Schule sinnvoll, denn die in den beiden folgenden Kästen vorgestellten Begriffe bergen Chancen für den Unterricht.

„**Messabweichung**: **Messwert** minus einen **Referenzwert**.

Anmerkung 1: Der Begriff ‚Messabweichung' kann verwendet werden,

a) wenn es einen einzigen Referenzwert gibt ... oder wenn ein vereinbarter Wert vorliegt oder
b) wenn angenommen wird, dass eine Messgröße durch einen einzigen wahren Wert oder eine Menge von wahren Werten von vernachlässigbarer Spannweite dargestellt wird.

Anmerkung 2: Messabweichung sollte nicht mit Fehler verwechselt werden."

Info 4-2: Definition der (Mess-)Abweichung nach VIM (deutschsprachige Übertragung).

„**Messunsicherheit**: nichtnegativer Parameter, der die Streuung der **Werte** kennzeichnet, die der **Messgröße** auf der Grundlage der benutzten Informationen beigeordnet ist.

Anmerkung 1: Die Messunsicherheit schließt Komponenten ein, die sich aus systematischen Effekten ergeben ... sowie die **Eigenunsicherheit** [Komponente, die aus der begrenzten Definition und Kenntnis über eine Messgröße ergibt, vgl. VIM 2.27] ... manchmal [auch nicht korrigierte] systematische Effekte ..."

Info 4-1: Definition der (Mess-)Unsicherheit nach VIM (deutschsprachige Übertragung).

Die Anmerkung zum Begriff Messunsicherheit ermöglicht und fordert eine Diskussion auf physikalischer Basis.

Wie viele Messungen sind erforderlich, um zu einem akzeptablen Ergebnis zu kommen? - Diese Frage ist nicht leicht zu beantworten. Soll eine Konstante, etwa die Fallbeschleunigung an der Erdoberfläche, bestimmt werden, so genügt im Prinzip, vorausgesetzt das Fallgesetz ist bekannt, genau eine Messung, wenn ich den überzeugt bin, dass ich fehlerfrei arbeite. Offensichtlich ist das keine unbedingt sinnvolle Aussage!

Sucht die Aufgabenstellung nach einem unbekannten funktionalen Zusammenhang, dann ist das prinzipiell schwierig. Um das zu verstehen, erinnern wir uns was wir sehr häufig in der Physik machen.
Wir betreiben sehr häufig unvollständige Induktion, denn wir führen einige Messungen durch und argumentieren dann, dass „das schon so weiter geht".
Zunächst eine Erinnerung an die „unvollständige Induktion":
Der „Physiker - Beweis" dafür, dass alle Zahlen Primzahlen seien, muss hier nicht wiederholt werden.
Die Behauptung aber, dass stets gilt $2^n \leq n^2$ soll überprüft werden. Man sucht einen Anfang, z.B. n =2. Für n=3 gilt das ebenso wie für n=4. Ist das ein Beweis? Nein, schon für n=5 wird das kla

Von Euler stammt die Vermutung, dass die Gleichung $p=n^2-n+41$ für n Primzahlen liefert. Das gilt für n = 1 . . .40. Nicht aber für n = 41.

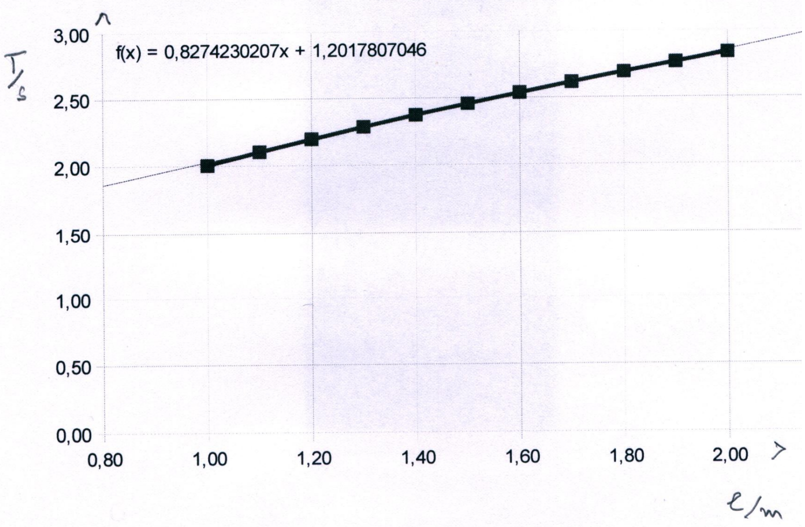
Der Zusammenhang zur Physik ist klar: Wenn wir ein theoretisch vorhergesagtes Gesetz überprüfen wollen, liefert auch eine große Anzahl von Messungen nicht notwendig einen „Beweis“.

Dazu zwei Beispiele. Für die Entladung eines Kondensators gilt $U(t)=U_0\cdot(1-e^{\frac{-t}{R\cdot C}})$ Bei genauerer Untersuchung zeigt sich, dass Depolarisationseffekte des Isoliermaterials zu Abweichungen dieses Gesetzes führen .
Das U-I-Diagramm eines NTC – Widerstandes ist exakt linear, aber nur solange die Temperatur konstant bleibt.

Es soll nun die Schwingungsdauer eines Fadenpendels bei kleinen Amplituden in Abhängigkeit von der Pendellänge untersucht werden. Es werden 10 Messungen bei unterschiedlichen Pendellängen durchgeführt. Wie vielfach üblich werden die Messwerte mit Unterstützung einer Tabellenkalkulationssoftware ausgewertet. Nachfolgend die Daten und die graphische Auswertung:

Fadenpendel

l/m	T/s
1,00	2,01
1,10	2,10
1,20	2,20
1,30	2,29
1,40	2,37
1,50	2,46
1,60	2,54
1,70	2,62
1,80	2,69
1,90	2,77
2,00	2,84

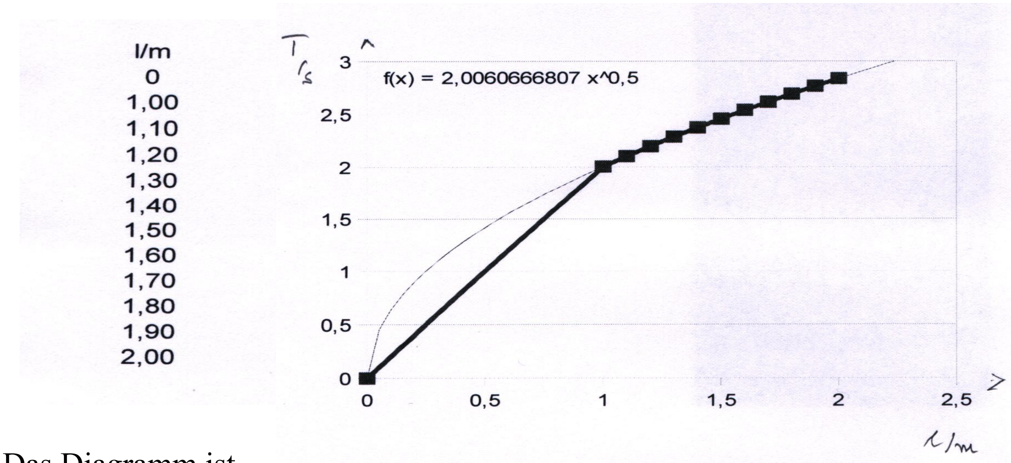


Das Ergebnis verblüfft. Die Regression ergibt eine Gerade. Ersetzt man x und y durch die physikalischen Größen, so lautet das Ergebnis $T(l)=0{,}83\cdot l+1{,}20$. Bedenkt man weiter, dass das absolute Glied die Einheit Sekunde und die Steigung die Einheit Sekunde/Meter haben muss, so erhält man zwar eine Gleichung, die richtig ist, wenn man die Einheiten betrachtet. Sie muss aber falsch sein, denn bei der Pendellänge 0 soll die Periode T = 1,20s sein. Das aber ist physikalischer Unsinn, denn mit der Pendellänge 0 gibt es keine Schwingungsperiode. Und ein linearer Zusammenhang als Ergebnis ?

Es wäre möglich, dass Messfehler zu diesem Ergebnis geführt haben. Dazu muss ich aber fairerweise sagen, dass die Ergebnisse nicht gemessen, sondern berechnet sind. Das System liefert also tatsächlich einen linearen Zusammenhang.

Die Zahl der Messungen reicht also nicht aus. Es muss vielmehr eine hinreichend große Variation erreicht werden. Wird ein nichtlinearer Zusammenhang vermutet, so ist sicherlich es sicherlich sinnvoll, dass die zu verändernde Größe mindestens verdreifacht oder vervierfacht, nicht wie in diesem Beispiel lediglich verdoppelt wird.

Zur Überprüfung der TabKal wird jetzt die Tabelle um das Wertepaar (0;0) ergänzt. Die nächsten beiden Bilder zeigen das Ergebnis.



Das Diagramm ist zunächst genau so gezeichnet, wie man es erwarten kann. Der Bereich, für den keine Daten vorliegen, wurde linearisiert.
Fordert man dann eine „potentielle Regression“, so entsteht der etwas dünn gezeichnete Graph. Die Gleichung, lautet, übersetzt, $T(l)=2{,}01\cdot\sqrt{l}$. Damit die Gleichung die richtigen Einheiten bekommt, muß die Konstante die Einheit $\frac{s}{\sqrt{l}}$ haben. Wir erhalten $2\cdot\frac{\pi}{\sqrt{g}}=2{,}01$.
Drei Gesichtspunkte sind hier festzuhalten:
Die in der TabKal eingebaute Mathematik ist offenbar sehr gut.
Mit Lernenden sollte man sehr vorsichtig sein, um nicht zu Ergebnissen zu kommen, die fast keinen Lerneffekt haben.
Einfach ein Wertepaar zu ergänzen braucht eine ebenso gute Begründung, wie einfach ein Wertepaar das „nicht passt“ wegzulassen!

Wir haben etwas zur Zahl und Auswahl von Messungen gesagt, wenn es um die Untersuchung eines Gesetzes geht.
Wenn es aber um die Bestimmung eines Wertes geht, sei es die Länge einer Strecke, die Verlängerung einer bestimmten Feder unter Einwirkung einer bestimmten Kraft, oder wenn die Dichte eines Metalles zu messen ist – wie geht man dann vor?

Das Wiederholen von Messungen ist – nicht nur unter der oben genannten zeitlichen Problematik zunehmend schwieriger zu vermitteln, da sich die verwendeten Techniken geändert haben. - Nehmen wir als Beispiel die Zeitmessung. Es ist klar, dass Starten und Stoppen nach akustischem oder optischen Signal ein Vorgang ist, der sehr unterschiedlich verlaufen kann.
Wir messen heute aber in vielen Fällen automatisiert (Lichtschranken, Filmen).
Einsichtig ist auch, dass beim Ablesen von Drehspulinstrumenten Fehler durch Parallaxe auftreten können. Wir verwenden heute aber fast ausschließlich Digitalinstrumente.
Da müssen andere Fragen gestellt werden. Wie sinnvoll ist die Angabe der fünf Stellen, die das digitale Multimeter anzeigt? Was weiß man über die Reproduzierbarkeit der Messungen? Wie stabil sind diese Geräte über lange Zeiten?
Messungen mit Cassy und ähnlichen Systemen suggerieren den Lernenden eine höhere Genauigkeit als Messungen, die von Hand aufgenommen werden. Zu diskutieren ist bei solchen Systemen die Reliabilität der intern verwendeten Verarbeitungsroutinen. Einige der damit verbundenen Fragen lassen sich durch Wiederholen von Messungen nicht klären.

Das ändert nichts an der Tatsache, dass irgendwann gelernt werden muss, dass jede Messung mit einer Unsicherheit behaftet ist. Gute Erfahrungen wurden gemacht, wenn das Thema recht früh in die Lerngruppen gebracht wird.

Versuch 1:
Auf der Wandtafel ist eine Kreidestrich (50cm $<l<$80 cm) gezeichnet. Zum Messen steht eine Holzlatte, die in Dezimeter geteilt ist zur Verfügung. Die Strecke soll auf cm genau gemessen werden.
Es wird sofort klar, dass hier geschätzt werden muss. Die einzelnen Messungen werden erst zusammen getragen, wenn alle gemessen haben.
Es wird deutlich: Die Messungen sind unterschiedlich. Es gibt Unsicherheiten. Vielleicht ist der Mittelwert ja am nächsten beim wahren Wert. Diese Idee kann dann mit einem üblichen Lineal überprüft werden. Wenn der „Mittelwert“ als Begriff akzepiert wurde, dann kann die mittlere Unsicherheit als weiterer Begriff eingeführt werden.
Als Ergebnis können wir begründet annehmen, dass der Mittelwert besser ist als ein beliebiger Einzelwert.

Das Experiment kann zu der Frage führen, ob die unterschiedlichen Einzelergebnisse dadurch verursacht sind, dass verschiedene Personen gemessen haben. Dieser Frage soll jetzt nachgegangen werden.

Versuch 2:
Eine Person hält ein Lineal (30 cm) am oberen Ende fest. Eine zweite Person hält ihre Hand so in Höhe der 0-Markierung um das Lineal, dass sie es noch nicht berührt, aber mit einem kurzen Griff zwischen zwei Fingern zu fassen bekäme.
Der Versuch wird mehrfach wiederholt. Es ergeben sich unterschiedliche Resultate. Die Ergebnisse verschiedener Personen können sich deutlich unterscheiden.

Was wir hier gemessen haben nennt man die Reaktionszeit. Diese kann bei verschiedenen Personen verschieden sein. Sie ist aber auch bei der gleichen Person abhängig von Konzentration, Müdigkeit etc.
Die Bedeutung für den Straßenverkehr darf nicht übergangen werden.

Wenn auch nach Ursachen für die „lange Leitung“ gefragt wird, kann man die Zeit, die ein Lineal zum Passieren einer Markierung braucht, mit einer Softwarestoppuhr messen.

Versuch 3:
Wird diese letzte Variation des Experimentes mit Linealen unterschiedliche Länge durchgeführt, so kommt man automatische zu Frage, wer am besten gemessen hat. Sicherlich liegt die Annahme nahe, dass die Unsicherheit möglichst klein sein muss. Es liegt aber auch die Idee nahe, dass der Betrag der gemessenen Größe von Bedeutung sein könnte.
Um der Idee zu folgen, werden deutlich unterschiedliche Strecken vermessen: Länge des Klassenraums, Länge der Wandtafel, Länge eines Tisches.
Die Diskussion der Ergebnisse führt dazu, dass der Quotient aus der Meßunsicherheit und der gemessenen Größe, also die relative Unsicherheit, ein geeignetes Maß für die Güte einer Messung ist.

Bei passender Gelegenheit kann dann die Fortpflanzung der relativen Unsicherheit betrachtet werden.

Versuch 4:
Die Fläche eines Rechtecks ist zu vermessen. Das ist nichts anderes als der zweite teil in Versuch 3. Bei der Auswertung wird ein Rechteck gezeichnet mit den Seitenlängen $a+\Delta a$ und $b+\Delta b$. Aus der Zeichnung liest man ab: $F=a\cdot b$ und $F_f=(a+\Delta a)\cdot(b+\Delta b)$. Sei die Unsicherheit ΔF , dann gilt $\Delta f=b\cdot\Delta a+A\cdot\Delta b$. Die relative Unsicherheit ist dann $\Delta\frac{F}{F}=\Delta\frac{a}{a}+\Delta\frac{b}{b}$.

Noch nicht berücksichtigt ist, dass die relativen Abweichungen auch negativ sein können. Formale Lösungen dafür sind in Bezug auf die Lerngruppe zu geben.

Man könnte meinen, wir wären wieder da, wo wir schon immer waren. Das ist nicht der Fall, denn die Experimente können parallel in Gruppen bearbeitet werden. Ziel ist, dass die Lernenden zu einer realistischen Einschätzung der eigenen Möglichkeiten und der durch die verwendeten Instrumente bzw. Versuchsanordnungen kommen.

Arbeitet man in der Elektrizitätslehre mit analogen Instrumenten, so wird man auch hier Zwischenwerte schätzen müssen, den Blickwinkel beachten müssen und auf die Güteklasse des Instruments zu achten haben.

Wie ist es bei der Verwendung digitaler Messgeräte für Strom und Spannung? Sie zeigen doch ein eindeutiges Ergebnis, beim Ablesen gibt es keine Fehler, jedenfalls im Idealfall.
Häufig sind jedoch die letzten Stellen der Anzeige nicht stabil.

Versuch 5:
Mehrere digitale Messinstrumente werden entweder als Spannungsmesser parallel geschaltet, oder als Strommesser mit einem Verbraucher in Reihe geschaltet. Ergänzend kann es interessant sein, zusätzlich ein analoges Instrument zu verwenden.

Hier sind offene Diskussionen, je nach Art und Qualität der verwendeten Instrumente zu erwarten.

Beim Übergang zu automatischer Registrierung und Auswertung von Daten erscheint es unverzichtbar mindestens von einem ausgewählten Experiment mehrere Datensätze aufzunehmen und zu speichern, um vergleichen zu können.

Heinicke, S. (2011). Aus Fehlern wird man klug. Eine genetisch-didaktische Rekonstruktion des ”Messfehlers”. PhD thesis, Universität Oldenburg.

Joint Committee for Guides in Metrology, editor (2010; corrected version). Evaluation of measurement data – Guide to the expression of uncertainty in measurement. Bureau International des Poids et Mesures, Paris.