Errors & Statistics

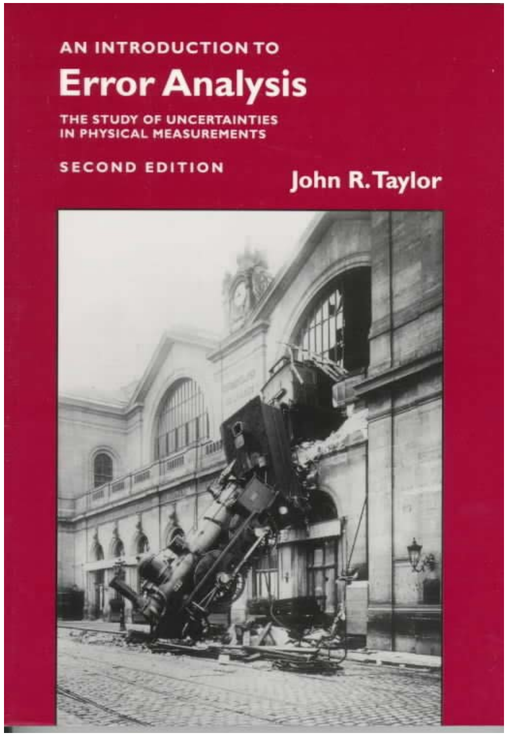
Prof. Matthias Möbius

mobiusm@tcd.ie

Module information

Recommended textbooks:

- *"An introduction to error analysis",* John R. Taylor
- *"Data reduction and error analysis for the physical sciences"*, Phillip R. Bevington



Lecture notes available on blackboard

Problem set on Mastering Physics

Syllabus

- Systematic and random errors
- Stating results and uncertainties, significant figures.
- Plotting data. Use of semi-log and log-log plots.
- Histograms and Probability densities.
- Discrete and continuous distributions such as binomial, Poisson, Gaussian. Moments of a distribution.
- Estimation of mean and standard deviation in a measurement.
- Error propagation and transforming variables in probability distributions.
- Linear and non-linear regression analysis, Method of least squares, Goodness of fit (Chi squared).
- Analysing plotting data using Logger Pro.

Error/statistical analysis is crucial!

- Science: Experimental Measurements underpin all physical theories. New theories and models need to be verified. Do the models fit the data within experimental error? Require statistics for error analysis.

- Physics: Statistical mechanics and Quantum mechanics require knowledge of probability theory.

- Medicine: Efficiency and side effects of new drugs need to be tested through clinical trials such as double blind studies.

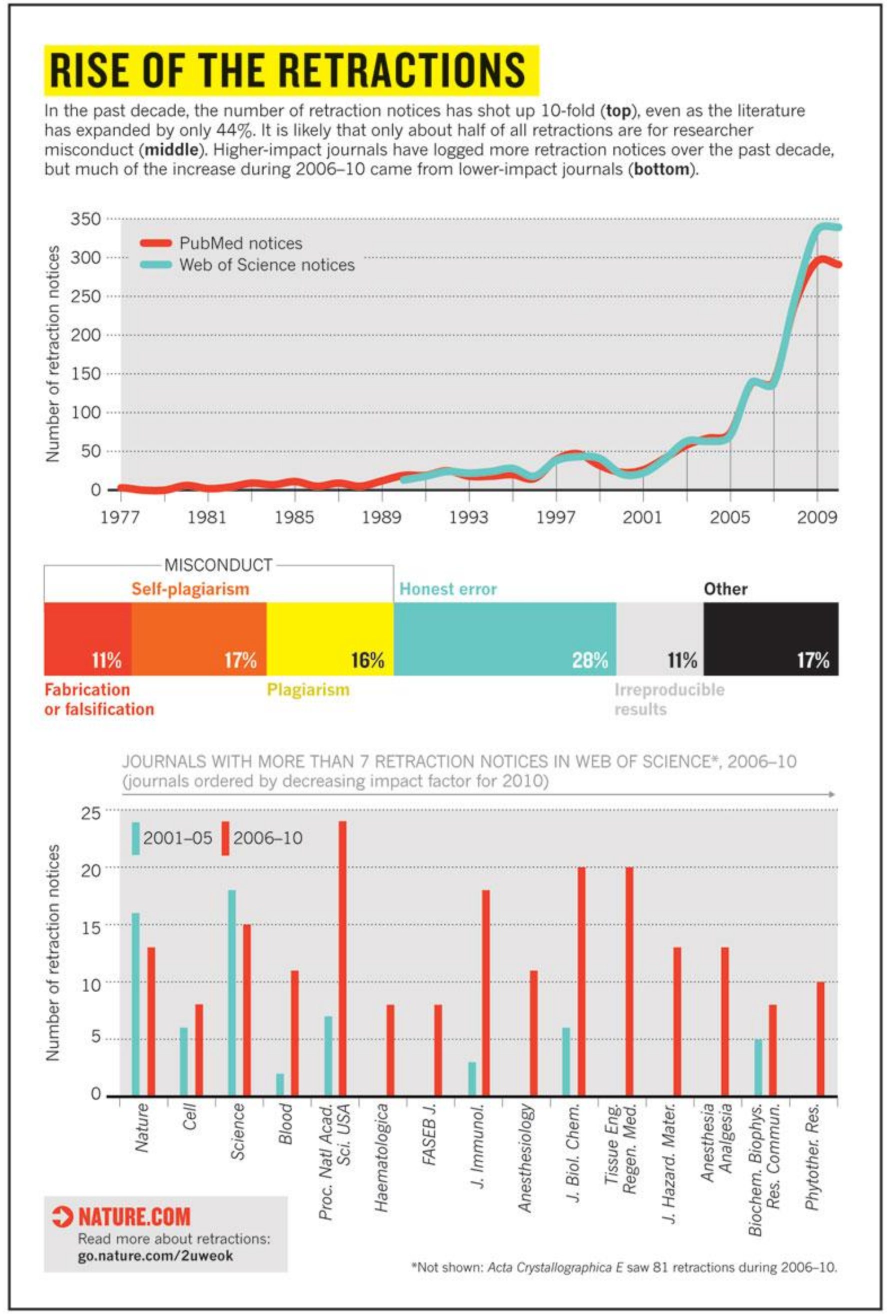
- Society: Policy decisions are (hopefully) often based on data that measure trends such as economic growth, jobless numbers, housing prices etc. In Ireland: Central statistics office (CSO).

Statistical analysis allows us to separate random events from real trends in the data.

It is the foundation of the scientific method !

The consequences of poor statistical analysis

- Number of retractions has risen dramatically.
- The journal “Nature” recently hired their own statistician to check validity of submitted papers.
- Submission of raw data increasingly required.
- Even though science is self-correcting, flawed papers can harm careers, slow scientific progress and bring a whole field of research into disrepute (e.g. Hendrik Schön affair).



Comment in Nature 2012

C. Glenn Begley & Lee M. Ellis

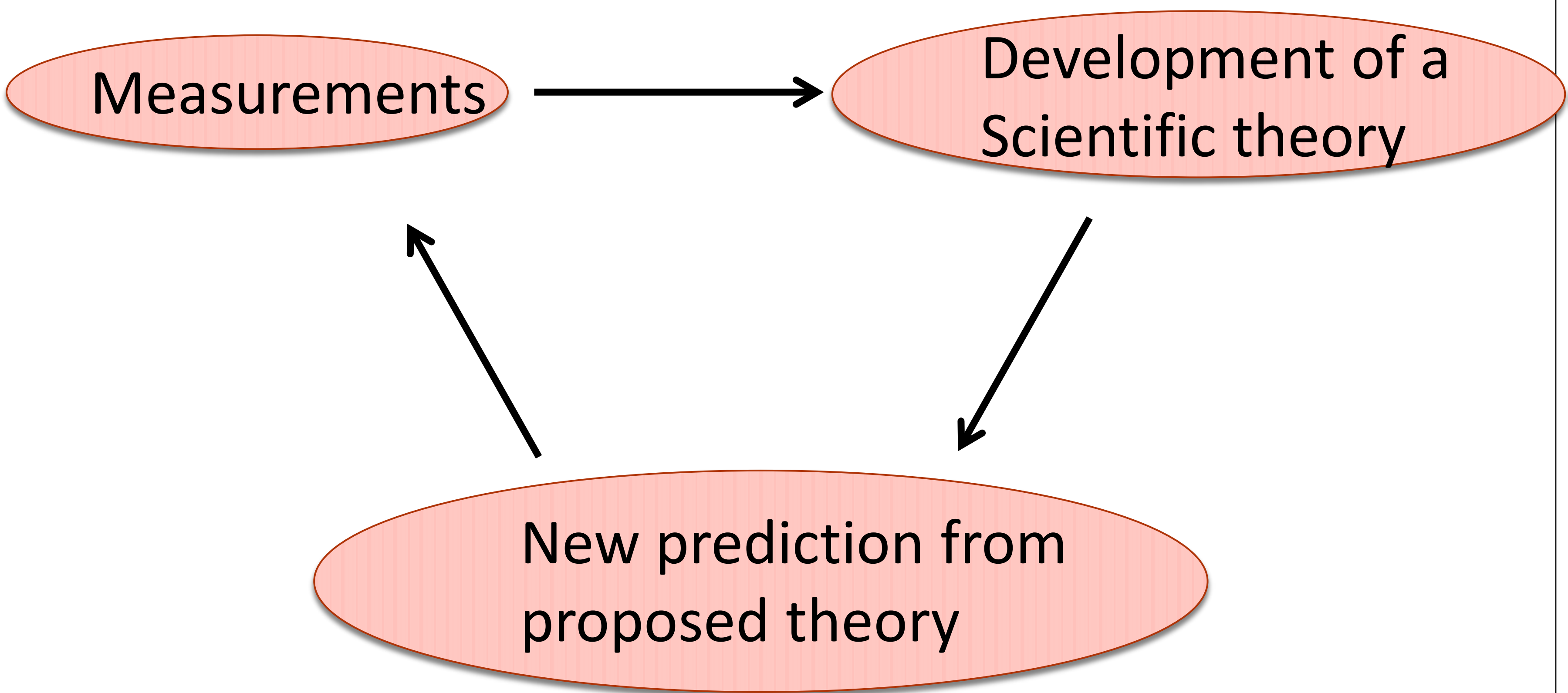
Over the past decade, before pursuing a particular line of research, scientists (including C.G.B.) in the haematology and oncology department at the biotechnology firm Amgen in Thousand Oaks, California, tried to confirm published findings related to that work. Fifty-three papers were deemed 'landmark' studies (see 'Reproducibility of research findings'). It was acknowledged from the outset that some of the data might not hold up, because papers were deliberately selected that described something completely new, such as fresh approaches to targeting cancers or alternative clinical uses for existing therapeutics. Nevertheless, scientific findings were confirmed in only 6 (11%) cases. Even knowing the limitations of preclinical research, this was a shocking result.

…

Unfortunately, Amgen's findings are consistent with those of others in industry. A team at Bayer HealthCare in Germany last year reported[4] that only about 25% of published preclinical studies could be validated to the point at which projects could continue.

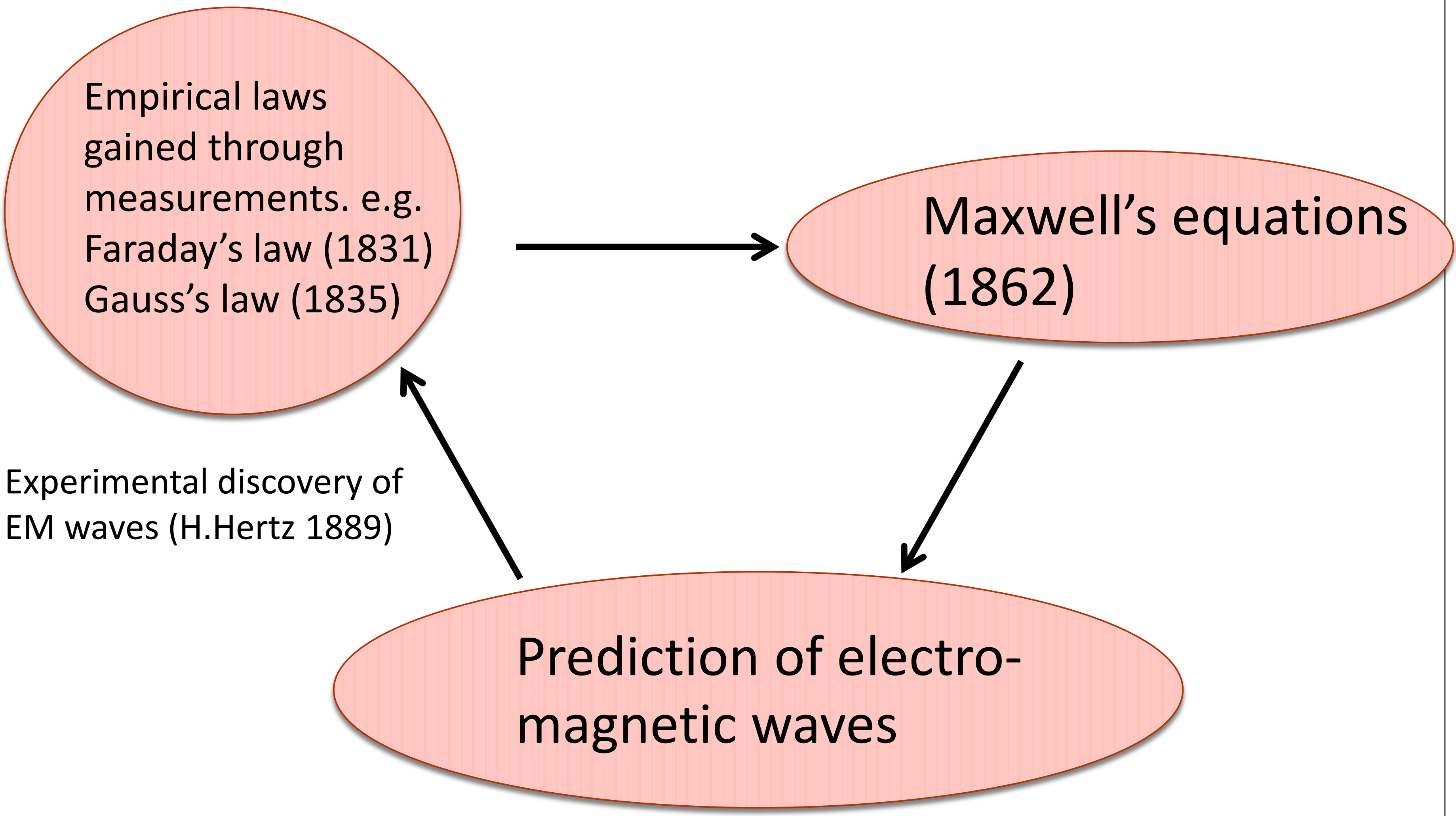
Freedman et al: Overall, the team found that poor materials made the largest contribution to reproducibility problems, at 36%, followed by study design at 28% and data analysis at 26%.

How does Science (typically) progress?



Scientific progress: Interplay between experiment and theory

Example: Electromagnetism

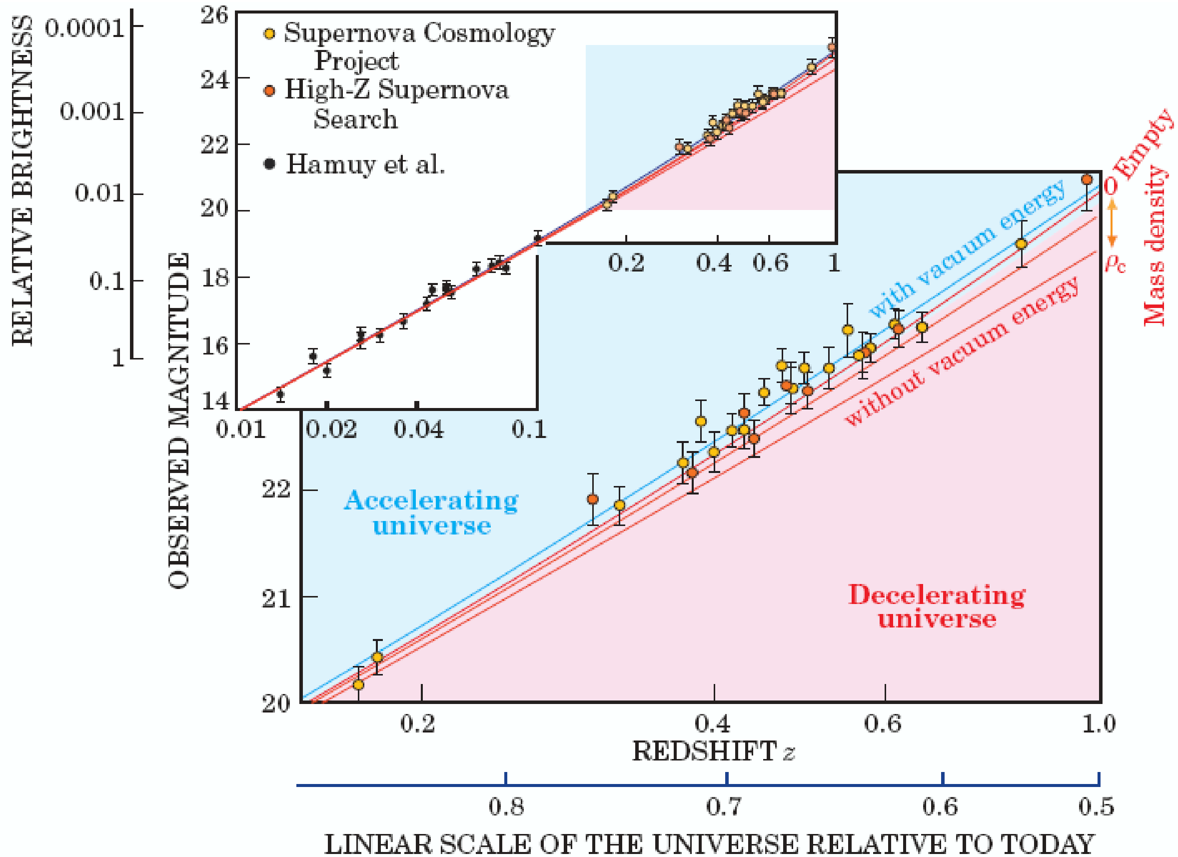


Heinrich Hertz on the discovery of EM waves:

It's of no use whatsoever[…] this is just an experiment that proves Maestro Maxwell was right—we just have these mysterious electromagnetic waves that we cannot see with the naked eye. But they are there.

Expansion of the universe (Noble price 2011)

Figure 3. Observed magnitude versus redshift is plotted for well-measured distant[12,13] and (in the inset) nearby[7] type Ia supernovae. For clarity, measurements at the same redshift are combined. At redshifts beyond $z = 0.1$ (distances greater than about 10^9 light-years), the cosmological predictions (indicated by the curves) begin to diverge, depending on the assumed cosmic densities of mass and vacuum energy. The red curves represent models with zero vacuum energy and mass densities ranging from the critical density ρ_c down to zero (an empty cosmos). The best fit (blue line) assumes a mass density of about $\rho_c/3$ plus a vacuum energy density twice that large—implying an accelerating cosmic expansion.



For this claim need to have a good estimate of error bars! Line fit (blue line) weights the data points according to their error bars.

Higgs Boson – Noble prize 2013

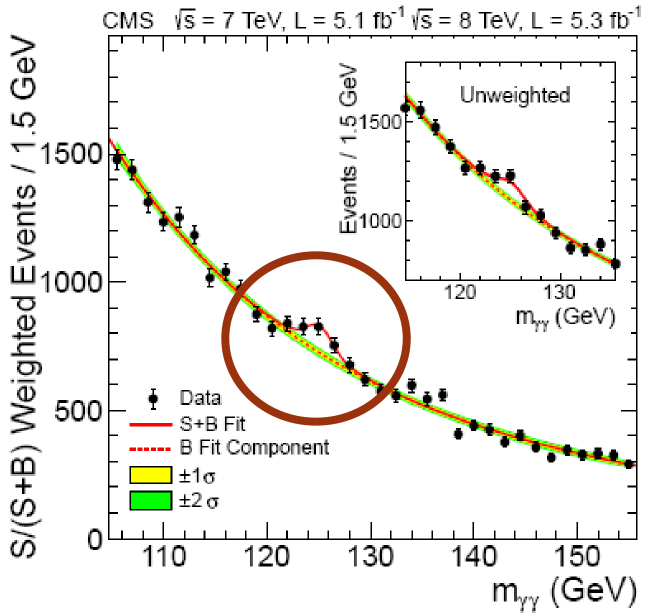


Figure 13: The diphoton invariant-mass distribution for the 7 and 8 TeV data sets (points), with each event weighted by the predicted $S/(S+B)$ ratio of its event class. The solid and dotted lines give the results of the signal-plus-background and background-only fit, respectively. The light and dark bands represent the ±1 and ±2 standard deviation uncertainties respectively on the background estimate. The inset shows the corresponding unweighted invariant-mass distribution around $m_{\gamma\gamma} = 125$ GeV.

Discovery of new particles require 5 sigma confidence levels (more on that later in the course).

Statistics tell us if it is noise or real.

What are errors?

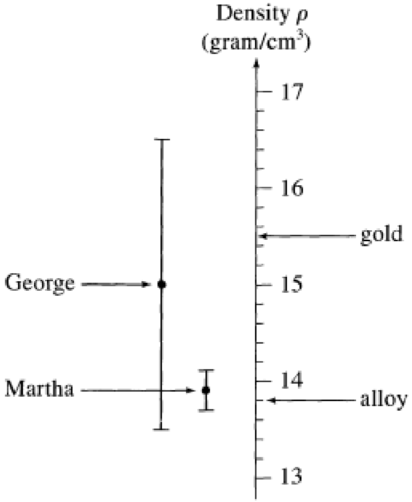
Errors ≠ Mistake/Blunder

Errors = uncertainty

<u>There is no such thing as a perfect measurement.</u>

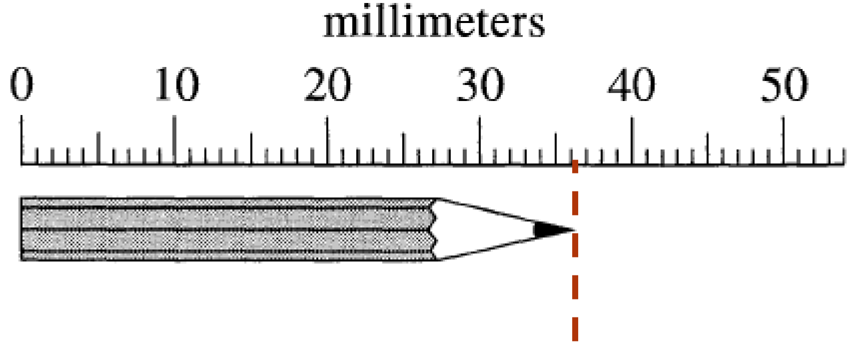
Every measurement has an uncertainty associated with it due to limitations of the measuring tools. Not reporting the uncertainty renders the measurement <u>useless.</u>

<u>Example:</u> Measure density of crown using Archimedes law



Sources of uncertainty

Example: Measuring the length of a pencil with a ruler



In general the tip of the pencil will not line up with the mm marking on the ruler.

Length L=36mm as measured to the closest marking.

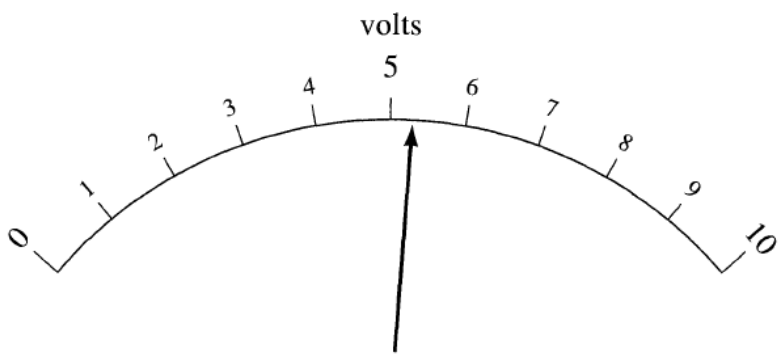
If not stated with any uncertainty, this means

35.5mm < L < 36.5mm

Always better to state uncertainty explicitly: L=(36.0 $\pm$ 0.5) mm

Sources of uncertainty

- Even if you use more advanced measurement techniques, there will always be a finite uncertainty. E.g. light interferometry is limited by the wavelength of light.
- Can use interpolation to improve meter readings:



- Closest marking 5.0 Volts, conservative estimate is (5.0 ± 0.5) Volts
- Using interpolation (by eye), best estimate (5.3 ± 0.1) Volts
- Can get additional uncertainty from parallax effects.

Sources of error

Reading a dial/ruler is just one of many sources of error.

Example: Measuring the period of a pendulum with a stop watch.

Since we don't know the reaction time a priori need to <u>repeat the measurement several times in order to estimate the uncertainty.</u>

- e.g. 4 measurements: 2.3s, 2.4s, 2.5s, 2.4s
- Best estimate of period = Average = 2.4s
- What is the uncertainty? Based on the range of measurements (2.3-2.5s), uncertainty is about 0.1s.

(Discuss later in the course how to extract uncertainty from repeated measurements using the standard deviation)

Random vs systematic errors

So far we discussed sources of errors that are random. When reading a dial or measuring with a stopwatch, you are equally likely to under- or overestimate. Other sources may be random fluctuations in temperature, humidity etc.

Another type of error are systematic errors. They can arise from faulty equipment or external influences that have not been accounted for. Systematic errors are not random.

Example: Stopwatch runs consistently fast by 5%. All measurements will then consistently be 5% too long.

Systematic errors

- Systematic errors cannot be reduced by repeating the experiments.
- Hard to detect systematic errors. (Not done in JF labs)
- Can minimise them by measuring known references. E.g. checking the calibration of a meters by accepted standards.
- Systematic error should always be smaller than desired precision.
- Experienced experimenters need to anticipate sources of systematic errors (equipment, tides, etc.).
- High precision experiments can be sensitive to unexpected outside events. e.g. TGV leaving Geneva station affected measurements at the particle accelerator at CERN.
- Infamous recent example: 2011 it was claimed that Neutrinos move faster than light (see Wikipedia entry on “Faster-than-light neutrino anomaly”)

How to report errors

Errors can be stated in the following ways

$1.7\ Volts$ (not recommended)

$1.70\ Volts\ \pm 0.05\ Volts$

$(1.70\ \pm 0.05)\ Volts$ (recommended)

When using variables use greek delta: $x \pm \Delta x$ or $x \pm \delta x$

How many significant figures do we need to show?

Significant figures

The number of significant figures correspond to the measurement resolution.

Significant figures = number of digits
– number of leading zeros
(– number of trailing zeros in integers)

e.g.	$0.00014 = 1.4 \times 10^{-4}$	two significant figures
	2.80×10^{-4}	three significant figures
	3000	<u>one or four</u> significant figures

Last example depends on context.

When converting units additional zeros are not significant. e.g. measure 20 mm with ruler - 2 significant figures. Converting to 20000 microns does not increase measurement resolution – still 2 significant figures.

Significant figures and uncertainty

Rule for stating uncertainties
Experimental uncertainties should (almost) always be rounded to one significant figure.

Note: For error propagation calculations can keep one additional figure – more on that later.

Example:

Incorrect measured $g = (9.82 \pm 0.02385)$ m/s^2

Uncertainty cannot have higher resolution than the stated result.

Correct measured $g = (9.82 \pm 0.02)$ m/s^2

Significant figures and uncertainty

Example:

Incorrect measured speed= (6051.78 ± 30) m/s

Stated result cannot have higher resolution than the uncertainty.

Correct measured speed= (6050 ± 30) m/s

Order of magnitude 10^1

Rule for stating results
Last significant figure in stated results should be of the same order of magnitude as the uncertainty.

Significant figures and uncertainty

If the answer 92.81 has an uncertainty of 0.3 then round

$$92.8 \pm 0.3$$

If the uncertainty is 3 then 93 ± 3

If the uncertainty is 30 then 90 ± 30

For very small/large numbers use scientific notation and put result and uncertainty in the same form:

$$(1.61 \pm 0.05) \times 10^{-19} \text{ Coloumbs}$$

and not $(1.61 \times 10^{-19} \pm 5 \times 10^{-21})$ Coloumbs

Significant figures and uncertainty

Exceptions to the rules of stating results and uncertainties:

1) If the leading digit in the uncertainty is 1 or 2, then keeping 2 significant figures is better.

Why? Rounding $\Delta x=0.14$ to 0.1 is a 30% reduction of the error, while rounding $\Delta x=0.54$ to $\Delta x=0.5$ is just 7%

2) Any answer and uncertainties used in subsequent calculations such as error propagation should carry at least one more significant to avoid rounding errors.

Fractional uncertainties

Uncertainties can also be stated as fractional uncertainties.

e.g. $x = 21 \pm 1$ and $y = 0.21 \pm 0.01$ have the same fractional uncertainty:

$$\frac{\Delta x}{x} = \frac{\Delta y}{y} = \frac{1}{21} = \frac{0.01}{0.21} = 0.05 = 5\%$$

Fractional uncertainties decrease with the number of significant figures in the answer.

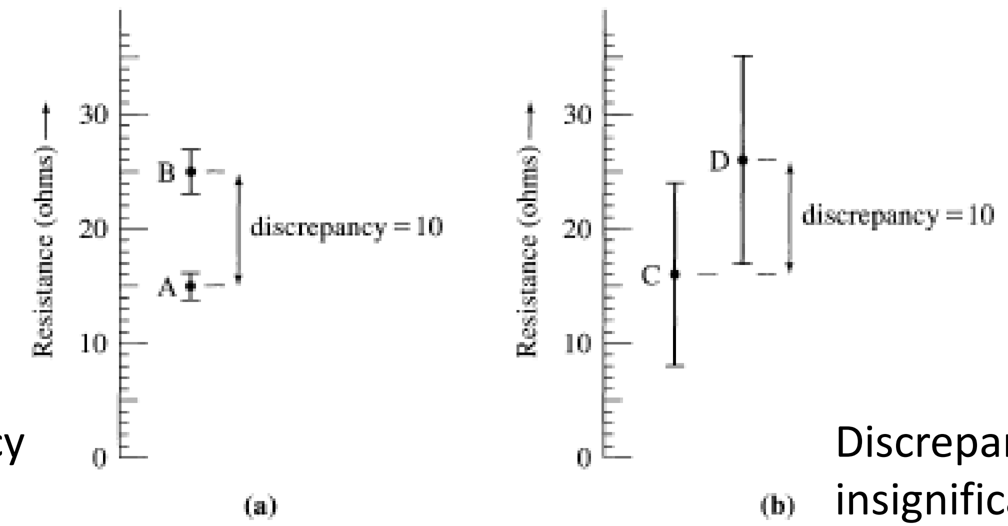
e.g. 21.0 ± 0.1 corresponds to a 0.5% uncertainty.

Comparing measurements

Typically, measurements are compared with one another or a prediction/reference value

e.g. student A and B measure resistance separately.

Do their measurements agree?



Discrepancy significant

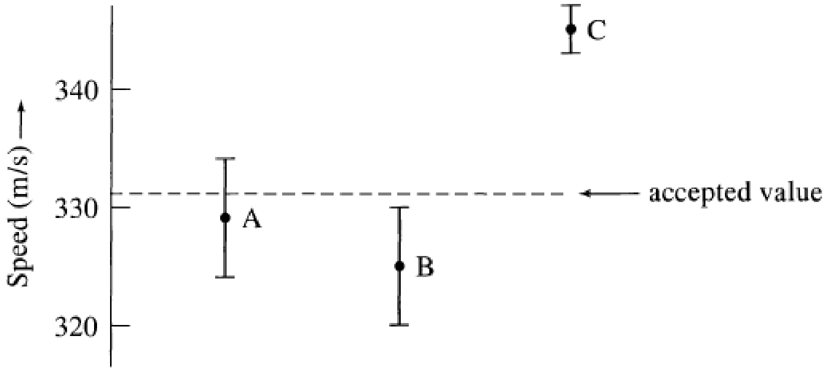
Discrepancy insignificant

Comparing with an accepted value

Often measurements are compared to a reference / accepted value that has been measured to a high accuracy.

Example: accepted value of speed of sound = 331 m/s

Students A,B and C compare their measurement with the accepted value



Student A: $v = 329 \pm 5\ m/s$ lies within the margin of error

Student B: $v = 325 \pm 5\ m/s$ consistent with accepted value

Student C: $v = 345 \pm 2\ m/s$ discrepancy = 14 m/s

= 7 x uncertainty

Something went wrong!