Good advice for writing reports

Hans Harhoff Andersen*

Christian Kraglund Andersen†

Department of Physics and Astronomy, Aarhus University, Denmark

October 23, 2011

Abstract

This is our personal collection of good advice and pointers to writing a report which can be approved. We wrote it inspired by the reports we were reading from the Introductory Mechanics course at IFA, Aarhus University fall 2011.

1   Important things if you want your reports to be approved

- *Always* remember to put units on all figures, tables and numbers in your report.

- *Always* remember to read your report after you print it. We very often get reports with missing appendices, figures, equations or completely meaningless sentences.

- Remember to take a lot of data. You need a lot of data to make it easier to spot the few data points that are weird (see section 8.1). If you only take one measurement you have no way of knowing whether your setup is stable and/or consistent until you try to analyze your data. Always try to measure the same thing several times to see if you get the same result!

- *Save* your data. Save all your data and in the rawest form you can. If you make a mistake somewhere in your data analysis you have to be able to go back and redo your calculations correctly. You probably prefer not to take your data again.

- Read the problem assignment closely. What have you been asked to do? If you neglect to answer a question that was explicitly asked in the assignment your report will *not* be approved.

- Whenever you do a measurement, calculate some value or derive an expression think about it for a second. Does it make sense? Does it make sense to say for instance *"we measured a $g=(11.2 \pm 0.1)\, \mathrm{m\, s^{-2}}$ which is too high. This could be due to wind resistance"*?

---

*hha07@phys.au.dk
†ctc@phys.au.dk

- Make nice figures and tables. Make sure that all axes have *correct* labels ie. parameter names and units and if you have more than one data series and/or fit plotted, use a legend to distinguish them. It is *crucial* that the figures are actually legible eg. text not too small, data easily visible. See sections 4 and 5 for details.

2   How to write an abstract

A good abstract contains a very concise description of what has been measured eg. *"we measured the gravitational acceleration of the earth using a pendulum"*. The abstract should not contain your personal feelings regarding the experiment or an explanation how your did relatively trivial things like calibrating, fitting and so on.

The abstract should also contain your results eg. *"We measured $g = (9.79 \pm 0.05)\,\mathrm{m\,s^{-2}}$."* and if applicable should contain your conclusion as to whether your results match the expected value. It should contain a *short* description of the reason for your experiment not working if it was not working eg. *"the measured gravitational acceleration was too low because we were neglecting wind resistance on a paper aeroplane which was not appropriate"*.

3   How to write a conclusion

The conclusion should *also* contain a rehash of all your results and your, well uh..., conclusions about the experiment. Do *not* write stuff like *"the experiment was fun"*. If you feel a need to write such things do it in the introduction, a footnote or make an acknowledgment section (or come talk to us :-) we like to hear it).

4   Fitting

An important part of data analysis is to be able to fit data and interpret the result. The most important thing to keep in mind is that if the fit does not *"go through"* the data points then your fit is bad. This does not mean that the fit, should touch every point, but you need to catch the general trend in the fit. So it is legitimate to have outliers (points that are way off), but if you systematically have points that deviate from your fit-curve then something is wrong.

Another important point is that you only have to fit *to the relevant points*. Often the data covers a larger time interval than the actual experiment eg. a camera could be recording while you are preparing to let go of an object and perhaps a little while after. The parts of the data that are *not* a part of the physics which you want to investigate should *not* be fitted; see figures 1 and 2 for examples of how to do it and how not to do it.

5   Figures

Once you have your data and perhaps a nice fit you need to put your figure in your report. Your figure should be placed near, but not necessarily right next to the text discussing it. If your figures are getting too small it is perfectly legal to put them on pages by themselves as appendices.

Your figures should have a caption describing what is shown in the figure. Your figures should also be numbered so that you can reference specific figures in your text.

Make sure that all data points are plotted as points, *not* lines, and that it is clear which are plots of data and which are theoretical or fits. The text should be large enough to read comfortably and the data points should be small enough to see the variations in them. Adding errorbars to the data is a really good idea, but is not always required.
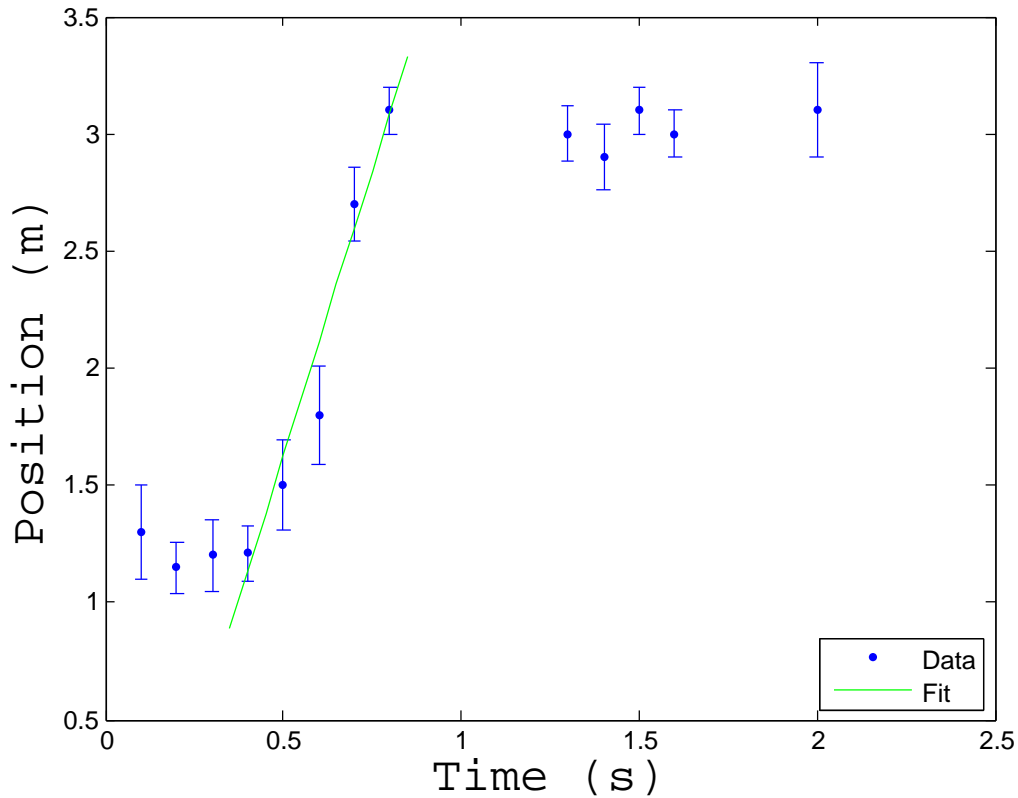
Figure 1: The good figure has a nice caption, parameter names and units/dimensions on the axes and a fit that actually traces the important data points. Notice that the y-axis does not go to 0.

Table 1: The caption should describe what is listed in the table and from which experiment the data is taken. It should also have any notes on the data eg. *"for the* (66.5 ± 0.4) kg *weight the collision was off-center so* ΔP *has a significantly larger error."*

mass[kg]	$\Delta P[\mathrm{kg\,m\,s^{-1}}]$
66.5 ± 0.4	12 ± 24
55.3 ± 0.3	18.2 ± 0.5
22.80 ± 0.12	2.2 ± 2.5

We have included three figures to illustrate the good(Figure 1), the bad(Figure 2) and the ugly(Figure 3) way of displaying a figure. Remember that make sure that the report can be understood even if printed in black and white.

6 Tables

For tables it also holds that they should have captions and be easy to read. In general it is best to put information in tables if the same type of values are repeated more than, say, twice. A nice table should look something like table 1 or table 2.
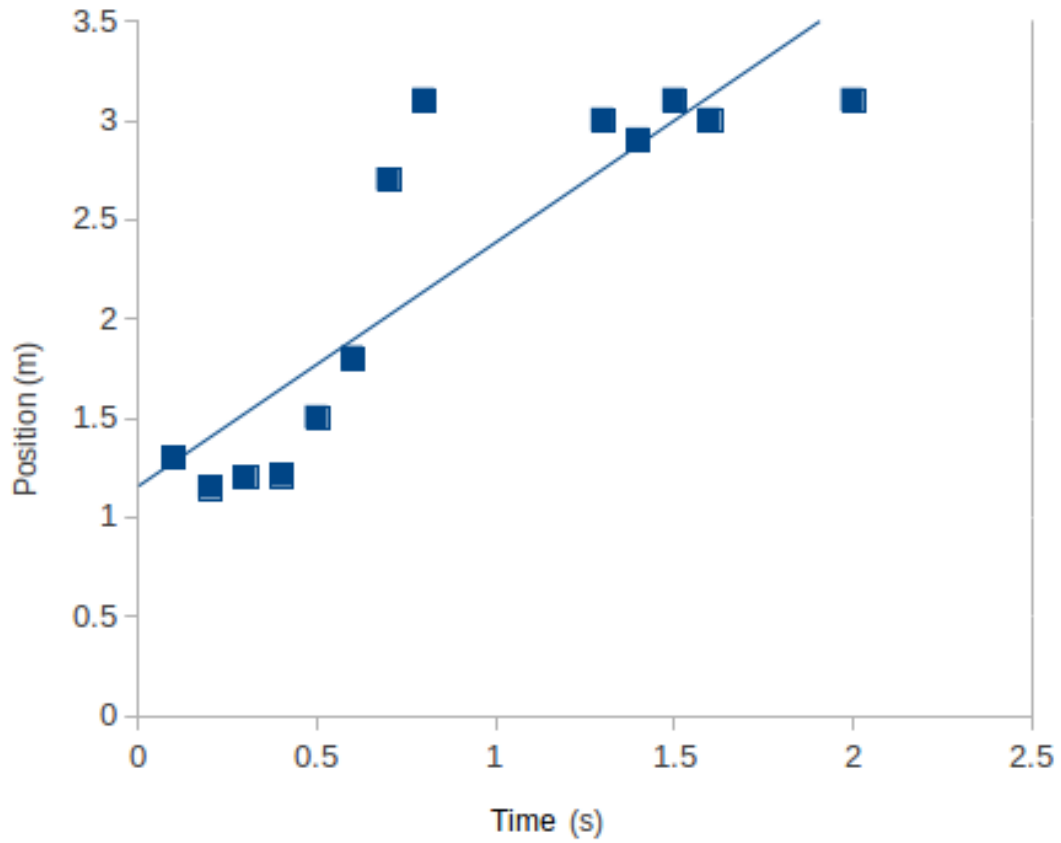
Figure 2: The bad figure has a fit that is *wrong* because the fit does not follow the data points. That means that the fit is not valid and cannot tell us anything about the data.

Table 2: Comparison of data to theoretical value.

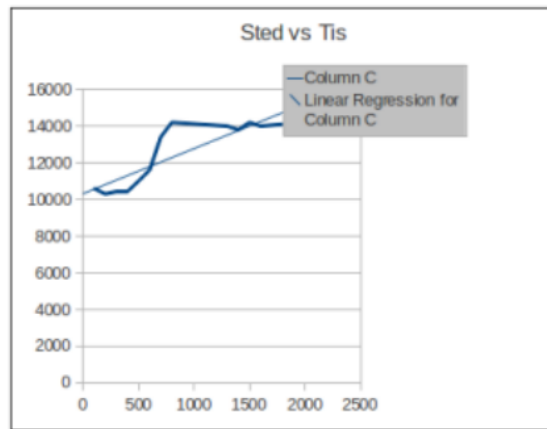|	Speed (m/s)
Mesurement 1	4.8 ± 1.1
Mesurement 2	4.2 ± 0.8
Mesurement 3	4.4 ± 0.9
Mesurement 4	3.9 ± 0.7
Mesurement 5	4.7 ± 0.6
Theoretical	4.3

Figure 3: The ugly figure is grainy, too small, does not show the data points, has a fit that is incorrect and a legend that covers important data. There is no labeling of the axes and the units that have been used are definitely not the natural units to use for the data. A lot of the figure space is wasted on white space and the legend does not give any relevant information. Both axes goes to zero which hides details of the figure. The title is uncessary and incorrect; the information that you want to put in the title often belongs in the caption.

7 Errors and uncertainties

You always need to distinguish between systematic errors[1] and uncertainties. Systematic errors are errors in you experiment that changes the result in a specific manner, while uncertainties are errors, that in the mean are zero. To give an example let us assume that you have a watch and you need to stop it by the time a stone hits the floor. If you are generally sloppy with your measurements and sometimes stop too soon and some times too late, then you will get an uncertainty in your result. On the other hand, if you are generally too slow and always stop the watch too late, you would create a systematic error.

A very important thing to remember is that every time you have a systematic error, you need to consider what this does to your final result. If you measure a time and you are slow with the watch, it cannot explain why you get a time that is shorter than the theoretically calculated time.

8 Uncertainties

Any parameter you use or measure has two parts. The first part is the value of the parameter; the second part is the uncertainty of the parameter. For instance we could measure a voltage $V = (4.3 \pm 0.3)\,\text{V}$ which means that our meter reads 4.3 and either it is flickering from 4.0 to 4.6 *or* the meter is rated to a precision to $0.3\,\text{V}$[2]

---

[1]Fejlkilder in Danish.

[2]An equivalent, but less common, way of writing this is 4.3(3) V, where the numbers in the parentheses indicates the uncertainty of the last digits of the value.

8.1 How to get uncertainties

To get uncertainties on your numbers you can estimate them (educated guesses) or look up the accuracy of the instruments which you are using. In this course *educated guesses are always enough*, but if you really have no clue how to guess, an alternative method is to calculate the uncertainties more systematically.

If you know the accuracy of each measurement you can repeat the measurement several times while noting down the result of the measurements with enough digits that the spread is visible. Then use the following equations to determine the mean, μ, and uncertainty of the mean, S_n, of your measurements of the variable x.

$$\mu = \frac{1}{N} \sum_{i=1}^{n} x_i \tag{1}$$

$$S_n = \frac{\sigma}{\sqrt{n}} \tag{2}$$

where x_i are the individual measurements and σ is the uncertainty or accuracy of the individual measurements. You can then write your result as $\mu \pm S_n = \mu \pm \sigma/\sqrt{n}$.

If you do not know the accuracy of your individual measurements you can still estimate the uncertainty of the mean value by calculating

$$s_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mu)^2} \tag{3}$$

where μ is the mean value of x and n is the number of data points. Notice that if you only have a single data point then this expression is not defined. This makes sense since you cannot estimate the uncertainty with only one data point. This estimate of s_n can then be substituted for σ in eq. (2) to find S_n. Note that s_n is the estimated precision of each individual measurement and S_n is the estimated precision of the mean.

To calculate the mean, spread and other statistical parameters you can use the built-in functions in most programs like *Excel, Matlab* or *Scipy* or you can do it yourself using the expressions given here.

8.2 Propagation of error

To get the error/uncertainty of a result derived from uncertain measurements you need to look at propagation of error. In high school it was sufficient to look at the number of significant digits of your measurement data and then you just assumed that you would get the same number of significant digits on your result. This is in a way the same as measuring a relative error, in percent say, of a variable, a, and then saying that the derived value $f(a)$ has the same relative error. This holds to some extent in simple cases. This is the absolute minimum you should do to get errors on your results.

A better way is to use the law of propagation of uncertainty/error[1]. This explains how to get the error of a parameter which is a function of one or more uncertain parameters. For the simple case where the errors on the input values are not correlated (ie. they are independent) the expression simplifies to (for the case of only two variables)

$$\sigma^2(f(a,b)) = \left(\frac{\partial f}{\partial a}\right)^2 \sigma^2(a) + \left(\frac{\partial f}{\partial b}\right)^2 \sigma^2(b) \tag{4}$$

where σ is the uncertainty, spread, error or standard deviation. Using equation 4 will give you a good estimate of your uncertainties.

8.3   Significant digits and uncertainties of uncertainties

When writing your result with uncertainties it is important to use a sensible number of digits for both the result and the uncertainty. If you have calculated your value and uncertainty to 1310 and 43 respectively then the proper way to write it is $(131 \pm 4) \times 10^1$ with the exception of uncertainties which start with a 1 eg. 562 ± 12.

8.4   Tools and further reading

There are several tool available to help you with calculations of uncertainties. The Matlab script *camas.m* calculates the uncertainty of the value of a function when given an expression of the function as well as the uncertainties on the input variables [2]. *mean etc.xlsx* is an Excel spreadsheet which can calculate the mean, standard deviation etc. for some input data. The files *camas.m* and *mean etc.xlsx* are available via the link in [2].

For more information and the reasoning behind the equations given here see [4]. For more advanced readers [3], which is used in *Advanced Data Analysis*, is also strongly recommended.

Bibliography

[1] Wikipedia, Propagation of uncertainty, http://en.wikipedia.org/wiki/Propagation_of_uncertainty

[2] camas.m is a Matlab script and a such requires Matlab. For instructions on how to use it look inside the file. The camas file can be found here: http://aula.au.dk/main/document/document.php?cidReq=EKSPF06&curdirpath=%2Fusikkerheder

[3] R.J. Barlow, Statistics - A Guide to the Use of Statistical Methods in the Physical Sciences, Wiley, 1989.

[4] Helge Knudsen, Præcision, Usikkerhed og Sandsynlighed - Et kort kompendium, 2008 (in Danish)