Brief notes on the linear fits

1. Each measured or calculated value must be represented with its uncertainty.

2. Any xy-diagram with data points and a function of best fit y(x) must show error bars for both the x- and the y-values. The error bars for the y-values are either the random (statistical) uncertainty or the reading uncertainty, whichever is greater. The error bars for the x-value represent the reading uncertainty. Specify in a caption the uncertainty value you have used for the error bars for x and y.

3. Finding the line of best fit by drawing a trend line and then calculating the coefficient of determination R^2 in Excel is not a sufficient analysis for this course because it doesn't account for the uncertainty of the measurements.

4. Excel's least squares linear fit can be used if the data points are shown with the error bars, and the uncertainty for both the slope and the y-intercept are shown in the diagram or given in the caption.

5. The more scientific presentation of the linear regression in Excel is possible with the LINEST function in Excel. The calculation returns the slope and the y-intercept with their uncertainties and the y- dimension's error bars as the "random uncertainties" s(y) in the y-values.

6. The LINEST or any other technique based on the least squares method does not account for uncertainties in x- values, because it cannot account for reading errors in both dimensions. Thus the x-values are treated as the ones defined without random uncertainty. This is a result of the routine procedure of the least square fitting: you are minimizing the sum of squares of differences (deviations) between the experiment value of y (not x!) and the value calculated for same x using the best-fit function.

7. To be accurate with LINEST, compare the reading error in y and the value of variance of the y-values s(y). The biggest between the two must be taken as the final uncertainty of the y-value.

8. You can use Matlab to do the fitting by entering error bars for both dimensions along with the polynomial power and the zero estimate of the slope and the y-intercept. To find the zero estimate, use Excel. Matlab's linear fit returns the slope and the y-intercept with their uncertainties, the chi-squared criterion of the

goodness of fit (which should be close to the number of degrees of freedom), and the graph of residuals vs variable values (the distribution of residuals about zero justifies the correctness of the choice of the fitting function, e.g. the correct choice of the power of the polynomial).

9. You can also use your own Python code to find the chi-squared criterion of the goodness of fit using the least squares technique. Present your result as the best-fit line function including the parameters of the function with their units and uncertainties, and the line of the best fit superimposed on the set of data points with their error bars. The description of how the uncertainties in the parameters and the error bars were calculated is required. The formulae for the slope and the y-intercept along with the formulae for calculations of their uncertainties can be found in the document called 'Uncertainties for linear fits' posted on Quercus.