

EXCEL EXERCISE AND ACCELERATION DUE TO GRAVITY

Objective: To learn how to use the Excel spreadsheet to record your data, calculate values and make graphs. To analyze the data from the Acceleration Due to Gravity Experiment by means of a computer and a spreadsheet program and to learn the treatment of experimental errors.

Part 1: EXCEL EXERCISE

EXERCISE 1. Entering Data and Calculations in Excel

Fifteen students of a Physics Class are observing a pendulum that is swinging back and forth. Each of them was holding a stopwatch and each one timed the pendulum as it makes 10 swings back and forth. The time recorded by each student is shown below.

Measurements in seconds:

20.5

21.2

19.9

20.6

18.9

21.1

20.6

20.3

19.8

20.0

19.6

20.4

21.5

20.2

20.8

a) **Enter the data in column A of Excel:** Using the top row of first column (cell A1), type in *Time for 10 swings, (sec)* to identify the data and its unit. You may need to widen your column. On the second row (cell A2), start typing in the data.

b) **Calculate the period T of the pendulum using Excel, where $T = (\text{Time for 10 swings})/10$:** On the top row of Column B (cell B1), type in *Period T, (sec)*. Equations in Excel always start with the “=” sign. On the second row of column B (cell B2), input the equation $=A2/10$ and hit ‘Enter’. You just divided your first time measurement (20.5 sec) by the number of swings (10) which gave you a period $T = 2.05$ seconds.

Now, Excel makes it easy for you to calculate the rest of the periods. Go back to cell B2. Note the little square on the bottom right corner of the cell. With your mouse, left click on that square and drag it down to the cell beside the last data entry, i.e., B16, then release. This translates the equation for the period T to the selected cells and instantly calculates the rest of the periods.

USING BUILT IN FUNCTIONS IN EXCEL

Excel has built-in functions that can be used to obtain values like the AVERAGE and STANDARD DEVIATION for numbers in a column of cells. These functions can be invoked by using the form $=FUNCTION(starting\ cell:ending\ cell)$ where the starting and ending cells define the range of data to be used. This range can be selected by left clicking on the starting cell and dragging it down to the ending cell.

c) **Get the average of the period T:** The average value of the period T can be obtained by adding all the values for T and dividing the sum by the number of T's added.

$$T_{ave} = (T_1 + T_2 + \dots + T_N) / N$$

This can be done easily in Excel by using the built-in function AVERAGE. Go to cell B18 and type in $=average(startingcell:ending\ cell)$ e.g. $=average(B2:B16)$ then hit 'Enter'. This gives you the average period T_{ave} . Label it by typing in *AVERAGE PERIOD* in the cell to its left.

d) **Get the standard deviation of the period T:** The standard deviation σ is given by the equation

$$\sigma = \{((T_1 - T_{ave})^2 + (T_2 - T_{ave})^2 + \dots + (T_N - T_{ave})^2) / (N - 1)\}^{1/2}$$

This can be a lengthy calculation but is done easily in Excel by using the built-in function STDEV. Go to cell B19 and type in $=stdev(startingcell:ending\ cell)$ e.g. $=stdev(B2:B16)$ then hit 'Enter'. This gives the standard deviation σ which is the uncertainty in the measurement of the period T. Label it by typing in *STANDARD DEVIATION* in the cell to its left.

FORMATTING CELL ENTRIES

There are several adjustments like centering, adjusting the number formats and many others that can be done to improve the appearance of your data tables. Some simple examples are shown below.

e) **Aligning your entries:** You may note at this time that all your numbers are staying on the right edge of your cells. A little bit of formatting may improve your data table. Try to center your entries by highlighting all the used cells (click on the first cell A1 and drag it down to the last cell used) and do the following:

Click on *Format – Cells – Alignment* – select 'Center' for Horizontal – OK

f) **Setting the number of decimal places for your entries:** Some calculations in Excel turn out with a lot of decimal places. Sometimes we need to round off and drop the excess decimal places to make our result more meaningful and realistic. For example, the stopwatch used may only measure up to the hundredths of a second, so it may not mean much to have an average or standard deviation that is up to 8 decimal places. Set the average and the standard deviation to two decimal places by highlighting their cells and do the following:

Click on *Format – Cells – Number – Number* – select '2' for *Decimal Places* – OK

See how much better your data table looks? Type in your name and your partner's name somewhere on the sheet then print a copy of this table for each student in the group.

GRAPHING DATA IN EXCEL

An easy way to graph in Excel is by using its Chart Wizard. When using this method you have to arrange the data in adjacent columns with the data on the left column plotted on the x or horizontal axis and the data on the right column plotted on the y or vertical axis.

EXERCISE 2. Graphing Data:

A cart is moving and its speed was increasing constantly. The velocity of the cart was measured at different times and recorded as shown below. Using Sheet 2 in Excel, enter the following data with time t in column A and velocity V in column B.

t (sec)	V (m/sec)
1	5.1
2	7.8
3	11.2
4	13.9
5	16.7
6	20.1
7	22.9
8	26.7
9	30.3
10	33.1

a) **Make a Velocity vs. Time graph using Excel's Chart Wizard with t plotted on the x axis and V plotted on the y axis:** To plot your data, highlight the t and V columns and click on 'Chart Wizard' icon which is the symbol with the red, yellow and blue vertical bars on Excel's Toolbar. On the Chart Wizard's window, do the following:

select *XY Scatter* for Chart Type and *Scatter without lines* for Chart Subtype – NEXT

The next window displays the chart type you selected, click NEXT, and this bring you a window where you can put the title of your graph and label your x and y axes. Type in '*Velocity vs Time Graph*' for Chart Title, '*Time, t (sec)*' for the x-axis label and '*Velocity, V (m/sec)*' for the y-axis label. Note that in labeling your axes, it is important to put in the variable name, its symbol and its units all the time. Then hit NEXT.

The next window asks where you want your graph to be. If you select '*As an object in Sheet2*', the graph will be put in the sheet where your data is. You can reposition and resize your graph on the sheet by clicking on it and dragging its corners. If you select '*As a new sheet*', the graph will be in a separate sheet by itself. For this exercise, select '*As a new sheet*', then click FINISH.

FINDING THE BEST FIT LINE – USING LINEAR REGRESSION AND THE METHOD OF LEAST SQUARES FITTING

In many experiments, the values of a variable Y are measured using different values of another variable X. The prime objective is then to find the equation that relates these two variables. This is done by plotting the results of the two measurements and finding the line that best fits through the data. This means of predicting the value of Y for a given X is called the regression of Y on X.

The Method of Least Squares Fitting is commonly used to find the best fit straight line. The criterion for fitting the line is based on the idea that the difference between the data and the fitted line must be minimized. The most accepted practice to do this is by choosing the best fit line such that the following sum is a minimum:

$$[y_1(\text{measured}) - y_1(\text{best fit})]^2 + [y_2(\text{measured}) - y_2(\text{best fit})]^2 + \\ [y_3(\text{measured}) - y_3(\text{best fit})]^2 + \dots + [y_N(\text{measured}) - y_N(\text{best fit})]^2 .$$

This computation can be easily done (and subsequently, the best fit line is determined) by using a computer program like Excel as shown in activity (b) below.

b) Plotting the best fit line and getting the equation of the line from Excel: Go back to your chart in Excel. The best fit line and the equation of the line $y = mx + b$ relating variables y and x can be obtained from Excel by doing the following: On the chart, right click on one of the dots of your graph. In the window that pops out,

click on ‘*Add Trendline*’ – select ‘*Linear*’ under Type –

– select ‘*Display Equation on Chart*’ under Options – OK

This will draw the best fit line and display the equation of the line with the form $y = mx + b$.

Click on the title of your graph and position the cursor at the end of the title. Type in your first name and your partner’s first name and enclose them with a parenthesis. Print out this graph with the equation on it and its corresponding data table, for each student in the group.

It is also customary to replace the variables y and x with the physical quantities that are plotted in the respective axes. In this exercise, replace y with v (for velocity) and x with t (for time). Now, write down the final form of the equation using v and t. In this final equation form, what does the slope m represent? The intercept b represents? Write your answers on your printout.

PLOTTING USING LINEAR REGRESSION IN EXCEL

So far a quick and easy way of plotting the variables (i.e. Chart Wizard) and obtaining the equation from Excel had been utilized in this exercise. Oftentimes, your experiment requires a more in depth look at the plotted data – is the difference between the measured y values and the best fit y values (called “residuals”) reasonably consistent? Are there any obvious outliers or bad data points? Also there are associated uncertainties or errors in the slope m and intercept b obtained in the equation for the best fit line and these uncertainties sometimes need to be

addressed in the analysis of your experiment. In the next activity (c), the Linear Regression feature of Excel will be utilized to plot the Velocity vs. Time data.

c) **Plot Velocity and Time using Excel's Linear Regression.** To plot using Linear Regression go to *Tools – Data Analysis – Regression*. In the window that comes out, click on *Input Range Y* – go to the column for Velocity, click on the top cell (i.e. the column title Velocity V (cm/sec)) and drag the mouse down to the last data cell. This defines the values to be plotted on the y axis. Click on *Input Range X* – go to the column for Time and click the top cell and drag the mouse to the last data cell defining the values to be plotted on the x-axis.

Note in this case that we have included the column titles in the range. Click on *Labels* specifying that the range includes the data labels. Also click on *Line Fit Plots – Residual Plots – New Worksheet Ply* – OK.

This gives two plots and some tables of statistical values. The Line Fit Plot shows the graph of velocity and time. The actual values of velocity are plotted in dark blue while the best fit values are in pink. (Note the plot titles include the variable plotted along the x axis, i.e. Time, t (sec) Line Fit Plot). The Residual Plot show how each of the actual values of velocity differed from the best fit values. The residuals (difference between the measured and best fit values) are shown in the “Residual Output”. The table which has “Coefficients” and “Standard Error” columns gives us the values of the intercept b and the slope m and their associated uncertainties.

Click on the Residual Plot and drag it beside the Residual Output Table. Click on the Line Fit Plot and reposition it below the Residual Output Table. Resize the Line Fit Plot to make it big. Right click on a best fit point (pink) and *Add a Trendline* which is *Linear* and *Display Equation* as done in activity (b). Go back to the top of the page and again, type in your names at a convenient location.

d) **Adjusting the page before printing.** To avoid having tables and graphs cut off, check out the Page Break Preview before printing. To do this, click on an empty cell then go to *View – Page Break Preview*. Read the message window that pops out and click OK. The solid blue line defines the whole range to be printed. The dashed blue lines define the boundaries of each page. If any dashed blue line cuts through a plot or data table, adjust the page size by clicking on the dashed blue line then dragging it out. Reposition the Line Fit Plot so that it is entirely on the second page. Make sure the tables and plots are all contained in two pages before printing a copy for each group member.

e) **Finding the slope and intercept and their uncertainties.** On your print out, circle the number labeled “Intercept” under “Coefficients”. This is the value of the intercept b in the equation $y = mx + b$. Circle the number labeled “Time t (sec)” under “Coefficients”. This is the value of the slope m. The uncertainties for the slope and the intercept are the corresponding numbers beside each of them under the column “Standard Error”. Identify and encircle each of these uncertainties on the print out.

f) **Inspecting the residuals of the data plotted.** Look at the Residuals Plot. Write a comment on your print out on whether the points are equally scattered above and below the best fit line. Are there stray points or obvious outliers or possibly a point which was entered erroneously?

PLOTTING ERROR BARS

Error bars on a graph describe the uncertainty in the measurements. The error bars can represent how confident one is in saying that a plotted point truly represents the value of the quantity being studied. One way to represent the error bar in a measurement is the standard deviation. However, when the averages or means (each taken from several measurements) are used to represent the values plotted in the graph, the standard deviation of the mean (also called standard error) is used to represent the error bar for each point. **The standard deviation of the mean is calculated by dividing the standard deviation by the square root of number of measurements N that make up the mean.**

i) **Plotting errors bars using Excel.** Consider the following data where the velocity of sound was determined by using a resonance experiment conducted at 3 different temperatures (10°, 20° and 30°). Making at least five trials for each case, the results are shown as follows:

Temperature (C)	Velocity of Sound (m/sec), 5 trials					Average Velocity	Standard Deviation	Standard Deviation of the Mean
10	334.8	333.8	339.5	335.2	339.7			
20	345.2	343.8	340.3	342.6	344.9			
30	350.8	346.5	341.6	351.2	345.8			

Use the knowledge you have already learned – enter the data into Excel, then calculate the average velocity, standard deviation and the standard deviation of the mean for each set of trials (adjust to two decimal places). The standard deviation of the mean (or standard error) can be calculated by entering the equation ($=\text{cell no. of the Standard Deviation}/\text{sqrt}(N)$) where $N = 5$ for this data set.

Highlight the 3 values of Temperature. Press the “Ctrl” button and highlight the 3 values of Average Velocity. Click on the Chart Wizard and plot these two quantities by using the *xy scatter option*, label accordingly and place “*as object in sheet 1*”.

Resize the graph if it is too small and position it under the data table.

The values of the Standard Deviation of the Mean will be used to put error bars on each of the plotted points. Click on one of the plotted points – click on “*Format Data Series – Y Error Bars – Both – Custom*”. Click on “+” and highlight the 3 values of the Standard Deviation of the Mean to specify the values for the error bar. Click on “-” and also highlight the 3 values of the Standard Deviation of the Mean – then click *OK*. The errors bars should now be seen on the plotted points.

The longer error bars show that there is a wider range in that particular value, hence less confidence in that mean value representing the real value of that quantity. Also note that the standard error grows smaller as the number of measurements N is increased, and the confidence in the mean value increases. If the error bars of one point overlaps with the error bars of another value, there is a lower likelihood that these two values differ significantly.

Type in your names on a convenient location on the sheet. Make sure that the table and plot are all in one page by checking the "Page Break Preview" before printing a copy for each member of the group.

TURN IN THE FOLLOWING TO GET CREDIT: the data table for Exercise 1, the data table for Exercise 2, the Excel graph with the corresponding equation of the line, the plot using the Linear regression, and the plot with error bars.

NOTE: Keep these exercises in mind as they will be greatly utilized in the data recording, calculations and graphing requirements of several lab experiments done in this course.

Part 2: ANALYZING THE ACCELERATION DUE TO GRAVITY DATA

I. Brief Comments on the Treatment of Experimental Errors

Experimental error or uncertainty is inherent in *any* experimental result at some level, however small. It is set by a combination of the design of the experiment, the quality of the apparatus and the skill of the experimenter. Generally it is possible to separate the sources of experimental errors into two categories: *random* and *systematic*.

A. Random Errors

The existence of random errors in a measurement can be inferred if repetition of the measurement does not give the same result each time. By definition, random errors are those that tend to average out upon repetitions of the measurement. For random errors then, the more repetitions of a measurement the less the uncertainty in the resulting average. From the mathematics of statistics it can be shown that the uncertainty due to random error in the average of N measurements decreases as \sqrt{N} when N is "large". Thus 400 measurements should give an average with half the uncertainty due to random error as compared to 100 measurements.

The first step in treating the random error in a large number N of repeated measurements is to calculate the average. The average is the desired result and it is the uncertainty of the average due to random errors which must then be determined.

If the N measurements are labeled $y_1, y_2, y_3, \dots, y_N$ the average as denoted by \bar{y} is defined as:

$$\bar{y} = \frac{1}{N} (y_1 + y_2 + y_3 + \dots + y_N) \quad (2b.1)$$

(The average is also sometimes called the mean in statistics.)

To establish the uncertainty in the average, \bar{y} the usual procedure is to first calculate what is called σ (sigma), the standard deviation of measurements. If it is believed that no

one measurement is more accurate than any of the others (as we will assume) then σ is defined by:

$$\sigma = \sqrt{\frac{1}{N-1} [(\bar{y} - y_1)^2 + (\bar{y} - y_2)^2 + (\bar{y} - y_3)^2 + \dots + (\bar{y} - y_N)^2]} \quad (2b.2)$$

The significance of σ is that any one additional y measurement has about a 2 in 3 chance of falling between $\pm \sigma$ of the average, \bar{y} . Statistical analysis further shows that the average, \bar{y} has about a 2 in 3 chance of falling within $\pm \sigma$ of the true value. Thus, after specifying the average of N repeated measurements, it is common to append $\pm \sigma$ as a measure of the uncertainty due to the randomness of the measurements.

Example

Suppose you have the following 5 measurements of the same quantity.

$$y_1 = 1.02$$

$$y_2 = 1.04$$

$$y_3 = 1.06 \quad \text{Then the average is}$$

$$y_4 = 1.01 \quad \bar{y} = (1.02 + 1.04 + 1.06 + 1.01 + 1.06) \\ = 1.0434$$

$$y_5 = 1.06$$

Note: One or two more decimal points must be kept in the calculations than are in the data until the calculations are complete.

The standard deviation of these measurements is:

$$\sigma = \sqrt{\frac{1}{5-1} [(1.0434-1.02)^2 + (1.0434-1.04)^2 + (1.0434-1.06)^2 + (1.0434-1.01)^2 + (1.0434-1.06)^2]} \\ = \sqrt{\frac{1}{4} [(0.0234)^2 + (0.0034)^2 + (-0.016)^2 + (0.0334)^2 + (-0.0134)^2]} = 0.0233.$$

The uncertainty in the average is $\sigma/\sqrt{N} = 0.0233/\sqrt{5} = 0.0104$. The result $\bar{y} \pm \sigma/\sqrt{N}$ is then 1.0434 ± 0.0104 . At this point some rounding off is appropriate and the result would be reported as 1.04 ± 0.01 or possibly as $1.04_3 \pm 0.01_0$.

It should be pointed out that the number 5 is stretching the concept of large N beyond reasonable bounds but that this is common practice when a large number of measurements are not available.

B. Systematic Errors

A systematic error is one which tends to repeat and thus create a shift in the average from the true value. Systematic errors may be provided by the experimenter, the apparatus or by poor design of the experiment. Because they are not revealed by repeated measurements, care must be exercised to investigate and account for all possible sources of systematic errors. This can be very difficult to do. Such errors sometimes remain unknown until other experimenters with other apparatus obtain convincing evidence that a previous result is off by more than the originally reported uncertainty.

II. Computer Analysis of Your Data for Acceleration Due to Gravity

Using Excel, enter in separate columns the measured distances ΔD in cm between the successive timing marks and the corresponding time t in seconds (i.e. $1/60$, $2/60$, $3/60$,...) obtained last week. Identify each column with the name of the quantity entered and the units they are in. Note that you made three separate runs in last week's experiment. (Each run can be entered in a separate Excel sheet.) Make sure to identify each run and that the mass of the cart m_2 , mass of the pulling weight m_1 , and the mass of the hanging weight m used to determine the friction force for each run are properly recorded and identified in your data table.

In another column, calculate the velocity v for each time interval by dividing ΔD by $1/60$ sec which is the length of each time interval corresponding to ΔD . Graph v versus t using the "Regression" feature of Excel which can be obtained from "Tools" – "Data Analysis" menu. The time t should be plotted on the x-axis and the measured distances ΔD on the y-axis. Make sure to click on the boxes for *Residuals*, *Residual Plots*, *Line Fit Plots* and *New Worksheet Ply* to produce the information and graphs that you need for analysis. Repeat the calculation of velocity v and the graphing instructions on this paragraph for the other 2 runs.

A. Interpreting the Results

Inspect the "Line Fit Plot" and check how well the measured values of the velocity v plotted on the y axis coincide with the "Predicted Y" on the computer fitted line. Inspect the "Residuals" and the "Residual Plot". Are your residuals reasonable or are there obvious outliers or data points that are possibly entered in error?

Go to your "Line Fit Plot". Add a trend line and display the equation of the best fit line for v versus t . Get the slope of this line (what does this slope represent?). Also determine the uncertainty in the slope by inspecting the resulting tables.

Repeat all of the above for each of the runs.

Tabulate the obtained values for acceleration a and its uncertainty. Calculate the value of the acceleration due to gravity g using the values of acceleration obtained from the Excel graphs. Tabulate these 3 values of g . Calculate the standard deviation σ of the three measurements of g as a measure of its uncertainty due to the randomness of the measurements.

B. Lab Report and Discussion

This section of your report should contain:

1. *Presentation of your results.* This includes printouts of your data tables and graphs from Excel. Make sure you complete a summary sheet which neatly summarizes all your results from your own hand graph analysis and those from this weeks computer analysis. Also, prepare a plot which shows (a) your three values of g determined last week by hand, (b) the three values of g you got today from the computer fits, with their estimated uncertainties shown as vertical error bars, and (c) the real value of $g = 9.81 \text{ m/s}^2$.

2. *Answers to the questions posed below (a to d).*

Questions:

a. By how much do each of your values for g differ from 980 cm/s^2 ? How does each difference compare with the corresponding uncertainty Δg due to the random errors in the data? Does it appear that there is a systematic error in either or both of your results?

b. If you worked carefully and your equipment worked properly, you may see the effect of a small systematic error inherent in the design of this experiment. In order to get the cart wheels to spin faster and faster as the cart accelerates, the table must exert a static friction force on the wheel rim in a direction opposite to the cart motion. This is *not* the friction force you measured when the cart was moving at constant velocity and *so it has not been accounted for*. A calculation of this retarding force on the cart based on the mass and geometry of the wheels and the physics of circular motion shows that *it should cause the measured value of g to come out about 2% low* for the values of m_1 , and m_2 used here.

What result do you obtain if your g values are each increased by 2% (about 20 cm/s^2) to account for the systematic errors above? Are your data accurate enough to see this effect? Do you still have systematic errors larger than the above?

c. List a few other possible sources of systematic errors that should be checked if time allowed. Hint: Think about some of the assumptions you likely have made about the equipment that may or may not be valid.

Because each run was made under different conditions, i.e., with different masses, they do not represent an exact repetition of a single measurement. Each run may be subject therefore, to different systematic errors. Until such errors are accounted for, averaging of your results is not strictly valid.

d. Compare the slopes you obtained last week with those generated by the computer analysis. Did you do a good job by hand?