PHYS 160 - Homework #8

Data Analysis

This exercise is to practice data analysis. You are to analyze a data set, y versus x, where x measures the distance from the interface between two media. Download the data file *potential.csv* . The data set has been placed in two columns, x and y,

- **(Plot1)** Plot y versus x. Label the plot with a title. Note how the data separates in two sub-domains, $x = [0, 2]$ and $x = [2, 5]$.
- Perform a global *linear* fit. Record the equation and the R^2 statistics for this fitted line.
- **(Plot2 and Plot3)** Re-plot the data in the two sub-domains. Label each plot with a title.
- Perform a *linear* fit in each sub-domain. Record the equations and the R^2 statistics for each fitted lines.
- Perform an *exponential* fit in each sub-domain. Record the equations and the R^2 statistics for each fitted lines.
- Comment on the quality of these fits.
- **(Plot4)** Plot the data together with the linear and exponential fits in the sub-domain regions *in one graph.*

Area under the curves

- Numerically calculate the area under the data via the trapezoidal approach. However be careful that the data set is *NOT equally spaced* in x. Use a modified trapezoidal rule, namely

$$area = \sum_{strips} \frac{1}{2}(y_{i+1} + y_i)(x_{i+1} - x_i)$$

  to calculate the area.
- The areas under the fitted curves are approximations to the area above. Calculate the areas under the three fitted curves, i.e., the global linear, sub-domain linear and sub-domain exponential fits. You may perform the integrals analytically by hand or use MAPLE.
- Calculate the percent error in each of these three areas as compared to the area calculated based on the data.
- Comment on the accuracy of the three areas calculated based on the three fitted lines. Over the each sub-domain, which fit was the best?