Laboratory Exercise: Gas Chromatography

Introduction
Few methods of chemical analysis are truly specific. Often, the compound of interest is mixed in with a few similar compounds. In general, chromatography is a laboratory technique used for the separation of mixtures. The mixture is dissolved in a mobile phase which is passed through an immobile, stationary phase. The phases are chosen such that components of the sample with differing solubilities settle in phase. A soluble sample will take more time to pass through the column, while a non-soluble sample will pass through the sample quickly. High performance liquid chromatography (HPLC) and gas chromatography use narrow tubes packed with a stationary phase where the mobile phase is forced. This process is called elution.

As components of a mixture move through a gas chromatography column, they reach a detector. The components reach the detector at varying times due to differences in the partitioning between mobile and stationary phases. The detector sends a signal to the chart recorder which results in a peak on the chart paper. The area of the peak is proportional to the number of molecules generating the signal. In this laboratory, we will be observing the results of a mixture of CO, CH₄, CO₂, and ethylene that have passed through a gas chromatography unit.

Objectives
- Understand the theory behind gas chromatography.
- Understand basic data analysis of gas chromatography.

Required Hardware
- Students should have MATLAB R2009a loaded on their computer.
Procedure

- Students should work individually on the laboratory.

Part I Importing data
1. Most gas chromatography equipment allows data to be exported as Excel files. Import the PyrolysisData.xls off Blackboard and into MATLAB for data analysis.
2. In the workspace browser, create the variable 'Time' from the first column of the data, and the variable 'Response' from the second column.
3. Create a line plot with Time on the x-axis and Response on the y-axis. Label the x-axis as "Time (sec)" and the y-axis as "Response (mV)". Add a scatter plot to the same graph for a better understanding of where the peaks and valleys are. Refer to the video tutorial on visualizing data referenced in the introduction laboratory exercise for assistance.

Part II Calculating the area under each peak in MATLAB
To interpret the data, you will need to find the area under the peaks. One way to do this is through the TRAPZ function. Given the x and y values, this function computes the trapezoidal approximation of the area. Type in the command prompt "doc trapz" to see the function's documentation entry and examples.

1. Notice that the gas chromatography data is calibrated at y=0.3. It is necessary to exclude area under this value. This calibration can be fixed for all points through the following command:
   \[
   \text{>> Response = Response} - \min(\text{Response})
   \]

2. First, find where the minimum y-values along the total graph occur. The locations of the minimum values are needed in order to know where each peak begins and ends. This can be done in MATLAB using the ‘findpeaks’ function.
   \[
   \text{>> [minvalues, indices] = findpeaks(-Response);}
   \]
   % the findpeaks function finds the locations of the maximum values (peaks) throughout the data plot.
   % By inserting the negative value in front of 'Response' we are prompting for the locations of the minimum values (valleys).
   \[
   \text{>> indices}
   \]
   % outputs the indices of each minimum value.

3. The eight indices outputted are the sides of each of the 4 areas to be found. To visually see where these indices are on the plot, re-plot the
data as you did in Part 1 Step 3. Hold the figure data, and add the following prompt:

```matlab
>> plot(time(indices),response(indices),'ro')
% plots the location of the relative minimums with a red circle
```

4. Once you know where to be taking the areas along the x-axis, the area of each peak can be found using the ‘trapz’ function.

```matlab
>> peak1_area = trapz(time(indices(1:2)),response(indices(1:2)));
```

5. Repeat the command in Step 4 and find peak2_area, peak3_area, and peak4_area. Be sure to use the proper indices for each.

Once you find area under each of the curves, you can find the percentage of CO, CH₄, CO₂, and ethylene.
(Note: The lightest molecule passes through the column first and will appear the earliest on the graph. Simply calculate the molecular weights of each to determine which peak is which compound.)

Part III Noisy Data

In the chromatography data you have been working with, the only minimum value locations outputted were the bases of the peaks that were needed in calculating the areas. In a real chromatography analysis situation, noise affects the response data and the lines are not as smooth.

1. Make sure the given data.mat variable is saved into your MATLAB directory. This is the data you have previously been working with, but with the noise factor added in. Run the ‘real_life.m’ file provided to you. What differences do you see with this data? What makes this data more difficult to work with? Explain how one might create a code in MATLAB to find the area under the four curves when the noise factor is an issue (You do not have to actually write out MATLAB code. Write in complete sentences how the coding used in Part 2 could be altered to achieve the desired results).

Results section should include:
1. The two plots created in MATLAB: Part 1 Step 3 and Part 2 Step 3.
2. Table containing the percentages of each compound expelled from the pyrolysis unit.

Discussion Questions:
1. How would the curve of an ideal gas chromatography results appear? Tall and narrow? Short and wide? Why?
2. What do the ‘retention factor’ and ‘selectivity factor’ each describe? (answer: retention factor: migration rate of compound through the column; selectivity factor: relative separation between two species).
3. What are some common mobile phases?
4. Write 5-10 sentences answering the questions in Part III Noisy Data.
Appendix should include:
1. A published MATLAB Word document containing all coding used to find the area under the curves. The coding used to create the plot in Part I should also be in this published document. Refer to the appropriate tutorial for a refresher on publishing.
2. Molecular weight calculations.