Introduction

Infrared spectroscopy has traditionally been an important qualitative analytical tool. Most chemistry students have heard of the fingerprint region of the infrared spectrum and its utility for the positive identification of pure compounds. The application of infrared spectroscopy to functional group identification in organic and inorganic molecular systems is widespread. However, there has not been nearly as many quantitative applications. This is due in part to the very rich but complex infrared spectra, often consisting of the superposition of more than one absorption band at any given wavelength.

In the early 1980's, a number of instrument manufacturers marketed relatively low cost Fourier Transform Infrared (FTIR) spectrometers. This development opened up new arenas for quantitative infrared spectroscopy. The reason is that FTIR is typically a more sensitive instrument than the dispersive infrared spectrometer, it can acquire spectra rapidly, and by its very nature, data are acquired and digitized. Once data are in the digital domain, it is possible to apply a variety of analysis algorithms to the rich spectral information, affording both quantitative and qualitative chemical information not previously accessible.

One such data treatment procedure is called factor analysis, which has been applied to the analysis of coal, bauxite, manganese dioxide ore, and for the determination of Cetaine index in diesel fuel\(^1\) and other systems\(^2,3\) using infrared spectra. Remarkable analyses are performed with relative ease compared with traditional analytical methods.

Factor analysis is a mathematical process which breaks the infrared spectrum down into a number of subspectra with unique weighing factors, analogous to the Taylor series representation of a mathematical function. The first member of the series (or first factor spectrum) has the greatest weight and best represents the function (or spectrum in the case of FTIR). The analysis is performed by multiple linear regression correlations of spectra with sample properties of interest, such as chemical composition, or physical properties. The values of the properties used in the correlation are determined by independent methods. This is the training step of the analysis. Once this step is completed unknown samples may be evaluated for those properties.

A number of factor analysis methods are applied to FTIR spectra, the most common one being partial least squares analysis (PLS). A complete discussion of this process is beyond the scope of this experiment and is not necessary since two excellent
reviews of the topic are in the literature$^{2,3}$. It is recommended that these sources be consulted for a better understanding of the details of the method.

The Midac FTIR instrument with Spectra Calc software is supplied with a partial least squares analysis program and a more sophisticated version is available from Galactic Industries. The program is called PLSQUANT and is accessed through the arithmetic menu, using the following keystrokes:

**F2 A D** then move the curser to PLSQUANT and press **ENTER**

In this experiment the concentration of hexane and cyclohexane are calculated in an unknown solution mixture. A series of training spectra of known composition are first acquired and then read into the PLSQUANT program as standards. The program is provided data on the composition of the training solutions, the number of training spectra which will be used, and other information. This is described in the Spectra Calc manual in the arithmetic applications guide. Once the training spectral information is provided to the program, the computation is commenced and the spectra are correlated with the chemical compositions. At the completion of calibration, unknowns are then determined.

**Equipment, Chemicals, and Supplies**

1. Midac FTIR spectrometer or equivalent. Set the collect menu to 10 scans and 4 wavenumber resolution. Acquire all spectra in absorbance mode from the collect menu.

2. Cell: This discussion pertains to a 0.1 mm pathlength cell with potassium bromide windows. Any appropriate cell may be used, however, if the pathlength is changed, the concentrations should be adjusted accordingly.

3. Reagent grade hexane and cyclohexane. Freon 113 (1,1,2-trichlorotrifluoroethane, Aldrich Chemical) was used as a solvent but any solvent with a window in the 3050 to 2800 cm$^{-1}$ region of the spectrum may be used as long as hexane and cyclohexane are soluble.

**Procedure**

Two stock solutions of hexane and cyclohexane are prepared by weighing to the nearest mg ca. 0.15 grams of each compound into separate 10 ml volumetric flasks and diluting to the mark with Freon 113. Use a Pasture pipet to deliver the hexane or cyclohexane to the bottom of the flask without transferring any to the walls on the neck of the flask. This will slow evaporation so an accurate weighing can be made. Dilute the hexane or cyclohexane immediately after weighing to avoid evaporative loss. Random aliquots of each stock solution are then transferred to a 2 ml screw capped vial which is weighed between each transfer (density of Freon 113 = 1.55 corrected for
0.15 g of hexane in 10 ml). Prepare five solutions in all; more if time permits. Solutions of 4 - 11 mg/ml in each component were used to produce the calibration which was used for the prediction shown in Figure 1. It is recommended that the solutions be analyzed as soon after preparation as possible to avoid evaporation. If the samples must be stored, do so in a refrigerator.

Begin by running a reference spectrum on the cell containing solvent only. The reference may be repeated between each sample analysis and will probably improve the analysis, but it is not essential that this be done. Then transfer the first calibration solution to the clean cell quickly to avoid evaporation loss. Stopper the cell and acquire the spectrum, saving it to disk.

Record the solution parameters in the memo, and name the sample "samp1". Continue this procedure until "samp5" has been run. An unknown spectrum may then be acquired. Window the screen from 3100 to 2700:

F2 E ENTER L L 3200 ENTER R 2700 ENTER Q

then normalize all spectra to this window with alt F4. Go to window mode:

F2 E ENTER M W *or Alt P if this macro is installed*

All 6 spectra will be visible on the screen for comparison. See if you can tell which calibration spectrum most clearly matches your unknown. Make sure all five calibration spectra and the unknown are saved on disk, then delete all but samp1, which will be the only active spectrum.

Now run PLSQUANT:

F2 A D PLSQUANT ENTER

Type a name for the calibration file (less than nine characters). A file will be created with this name and a .CAL extension will be added automatically. Next choose CALIBRATE TRAINING SET. Enter 2 for the number of components, 5 for the number of calibration spectra, and select 1 for the number of spectral bands. Choose EVERY POINT which utilizes all the data in the band. Select the bands to be used in the analysis with the cursor or mouse by selecting P for point to frequency (alternatively these may be entered from the keyboard). Move the cursor to 3050 or as close as possible. Hold down the alt key and repeatedly press the left or right arrow to obtain fine control of the wavelength. Then select this band edge by pressing F9. Select the right band edge at 2800 in the same manner, but by pressing F10 to make the selection and then hit ENTER. Choose 3 for the number of dimensions. Larger or smaller dimensions may be tested later by rerunning PLSQUANT and selecting RECALIBRATE DIMENSIONS. Enter the concentrations of the two components when asked, letting hexane be component 1 and
cyclohexane component 2. Then from themenu select samp2 to be read into active memory and enter its component concentrations when asked. Continue this process until all spectra and their data have been entered, at which time the calibration calculations will commence.

Load the unknown spectrum into active memory. Then run PLSQUANT again, and select PREDICT UNKNOWN from the menu. Then choose DISPLAY to see the results of the prediction. Record the values in your laboratory notebook.

**Questions**

1. Discuss the most significant sources of error in this analysis. What measures may be taken to reduce these errors?

2. Did your estimate of which calibration spectrum most closely matched the unknown composition agree with the PLSQUANT prediction? What about the ratios or the two components, i.e., is the ratio of hexane to cyclohexane in your prediction and in the PLSQUANT calculation similar?

3. Discuss other analytical problems which may be difficult to determine by other methods, yet may yield to this approach. Hint: FTIR is able to analyze insoluble solids (e.g. polymers), liquids and gases.

4. Name two other analytical tools which may be used to perform this analysis. What are the time advantages of each method, assuming 10 samples per day are analyzed, and only periodic calibration is required for the FTIR method? What is the estimated accuracy of each method? What is the sensitivity of each method?

5. Beer’s law is the basis for the quantitative analysis here. Discuss this relationship with regard to the concentrations and pathlength used. What are the limits of concentration and pathlength?

6. What molecular phenomena are responsible for the absorption band(s) in hexane and cyclohexane between 3200 and 2700 wavenumbers? How many bands are involved for each compound in the region we are working, and what are the frequencies of the maxima for these peaks?

7. Discuss how the PLSQUANT program resolves the problem of overlapping bands when Beer’s law applies only to one absorption process at one wavelength. Refer to the literature articles cited.

8. Why are random aliquots of the two stock solutions used to prepare the calibration mixtures rather than taking serial dilutions of each stock solution and
combining these?

References:

**Notes to Instructors:**

Two stock solutions may be prepared by the laboratory instructor, and these dispensed to the student for subsequent dilution. The unknown is also prepared from these stock solutions and given to the student. This will reduce prediction errors in the analysis, since all data are referenced to the same source of solutions. The student may also prepare an unknown from these same solutions, and predict its composition. This prediction can be compared with the unknown given out by the laboratory instructor.

It is important that the unknown be within the range of concentrations of the calibration set. Furthermore, the narrower the range of calibration standards, the better will be the prediction. The more standards in the calibration set, the better will be the prediction.

A longer pathlength cell can be easily fabricated from a glass cylinder with one or two sidearms for sample introduction. Epoxy two KBr windows to the ends of the cylinder. The diameter of the glass cylinder should be at least 1 cm preferably larger--the Midac beam diameter is approximately 0.6 cm centered 2.5 inches from the floor of the sample compartment. Make sure the cell is centered in the beam--use the align function in the collect menu if necessary. A longer pathlength cell is used in the EPA method 418.1 for the determination of total petroleum hydrocarbons by infrared spectroscopy. This method is strictly a single absorbance maximum reading, not PLSQUANT, and it is applicable to concentrations in the low PPM range.

Macros may be written for all of the operations described in this experiment. This will reduce (but probably not eliminate) keystroke errors by the students. If it is of interest for the students to learn the software, then the keystrokes should be used. In our opinion, it would be useful for the students to write the macros themselves, since the real world of analytical chemistry is filled with macros.

The values of the concentrations used here keep the absorbencies below 0.7 units. It is likely the system will function up to significantly higher absorbance values but this was not tested.