Noise

In this laboratory exercise we will determine the Signal-to-Noise (S/N) ratio for an IR spectrum of Air using a Thermo Nicolet Avatar 360 Fourier Transform Infrared Spectrometer. In particular, we will determine how the S/N ratio depends on the instrument's signal intensity and resolution. We will also examine how to improve the S/N ratio by using Ensemble Averaging as well as Signal Smoothing.

Visualization of the Output of a Noise Generator
https://github.com/mikolalysenko/spatial-noise

Instrumentation is now ubiquitous in the chemical laboratory. By the late 1800’s, chemists realized that the chemical composition of a sample influenced its physical properties and began to develop instrumental methods for measuring these properties. Limited initially to spectroscopic methods, analytical chemists now have a wide variety of instrumentation at their disposal with which to characterize their samples. Currently, electronic analytical instrumentation is capable of detecting minute amounts of even the most difficult analytes.
In broad strokes, we can imagine that an analytical instrument will stimulate the sample to be characterized and then measure the sample's response to this stimulus. Because of the inevitable Noise, the response typically must be filtered to remove as much of the noise as possible. The resulting signal can then be amplified and recorded for further analysis. Of course there are many variations on this basic outline, but the overall scheme is pretty much the same from instrument to instrument; with chromatographic instrumentation being a major exception.

Over the course of the semester, we will have an opportunity to examine several specific instruments and their use in the analytical laboratory. For now we want to focus on the generic features of analytical instrumentation as outlined above. In this exercise, we will examine the Noise superimposed on instrumental Signals. We will not consider the sources of this noise; that will come later when we examine particular instrumentation. Instead we will consider how the signal-to-noise ratio can be influenced by instrumental parameters, such as signal strength and resolution. Of course, we desire this ratio be very large. However, the cost in time and other resources may prohibit this ideal. We may instead have to settle for a signal-to-noise ratio that is simply acceptable.

For our purposes, Noise (N) is defined as the standard deviation $\sigma$ of a Signal's measured values and the Signal (S) is the average $\bar{x}$ of those measurements. (Note: I am using $\sigma$ to represent the standard deviation of the sample distribution rather than the typical $s$. This is done to avoid confusion between $S$ and $s$.) The Signal-to-Noise ratio (S/N) is then the ratio of $\bar{x}$ to $\sigma$.

$$S/N = \frac{\bar{x}}{\sigma}$$  \hspace{1cm} (Eq. 1)
This is simply the inverse of the Relative Standard Deviation of the signal;

\[
S/N = 1 / (\bar{x}/\sigma) = 1/RSD
\]

(Eq. 2)

This does assume the signal is relatively constant over the period for which these measurements are made.

If the noise is Normally Distributed and we select a confidence level of 99%, then all but 1% of the signal measurements will lie between $+2.5\sigma$ and $-2.5\sigma$ of the mean $\bar{x}$. This means that if our Maximum and Minimum signal values can be assumed to be good estimates for $\bar{x} \pm 2.5\sigma$, then $\sigma$ can be approximated according to:

\[
\sigma \sim (\text{Max} - \text{Min}) / 5
\]

This provides us with a method for quickly estimating the S/N ratio for those cases where precision is not an issue.

\[
S/N \sim 5 \bar{x} / (\text{Max} - \text{Min})
\]

As noted by Skoog, *et. al.*:

In most measurements, the average strength of the noise $N$ is constant and independent of the magnitude of the signal $S$. Thus, the effect of noise on the relative error of a measurement becomes greater and greater as the quantity being measured decreases in magnitude. For this reason, the *signal-to-noise ratio* (S/N) is a much more useful figure of merit than noise alone for describing the quality of an analytical method or the performance of an instrument.

The upshot is, all other factors being equal, a better S/N ratio allows us to detect weaker analyte signals.

Now to a couple of methods for increasing the S/N ratio digitally, post data collection. First is Ensemble Averaging, or scan Coaddition. Here several data scans are performed by the instrument and the resulting data is averaged point by point.
The Coadded Signal \((S_x)\) is determined by averaging the individual signals \((S_i)\) obtained for each of the \(n\) scans at each point \(x\) along the scan:

\[
S_x = \frac{\sum S_i}{n} = \frac{s_{sum}}{n} \quad \text{(Eq. 3)}
\]

where \(s_{sum}\) represents the signal of the summed scans.

As the individual signals \(S_i\) themselves are embedded in noise we have:

\[
S_{sum} = n \bar{x}_i \quad \text{(Eq. 4)}
\]

Then, for the "summed" spectrum:

\[
\sigma^2_{sum} = \sum \sigma_i^2 = n \sigma^2 \quad \text{(Eq. 5)}
\]

where \(\sigma_i^2\) is the variance associated with \(S_i\) and is assumed to be the same for each \(S_i\); given by \(\sigma\), the scan’s noise. Thus, the S/N ratio for the coadded spectrum is given by:

\[
\frac{S_x}{N_x} = \frac{(s_{sum}/n)}{(\sigma_{sum}/n)} = \frac{\bar{x}_i}{\sqrt{n} \sigma/n} = \sqrt{n} \frac{(\bar{x}_i/\sigma)}{\sqrt{n}} = \sqrt{n} (S/N) \quad \text{(Eq. 6)}
\]

This means that for the coadded signal:

\[
S_x/N_x \sim n^{1/2} \quad \text{(Eq. 7)}
\]

Thus, provided the data acquisition time for obtaining a scan is not too great, coadding multiple scans can improve the signal to noise ratio, but only at the rate of the square root of the number of scans.

Assuming the signal varies only slowly with time, Boxcar Averaging can smooth a data scan and improve the S/N ratio. This technique is as illustrated below. Data points are averaged within each boxcar and the resulting average is taken to be the new data point within the boxcar. As noted by Skoog et al: "...detail is lost by boxcar averaging, and its utility is limited for complex signals that change rapidly as a function of time. Boxcar averaging is of considerable importance, however, for square-wave or repetitive outputs where only the average amplitude is important."

Basically, considerable resolution is lost as a result of boxcar averaging.
As an example, consider a 3-point boxcar average of the following raw data:

\{5, 12, 17, 18, 21, 22, 25, 30, 35, 26, 25, 21, 15, 13, 18\}

First the data is divided into boxcars:

\[[5, 12, 17] [18, 21, 22] [25, 30, 35] [26, 25, 21] [15, 12, 18]\]

Then the averaging is performed:

\[(5 + 12 + 17) / 3 = 11.3\]
\[(18 + 21 + 22) / 3 = 20.3\]
\[(25 + 30 + 35) / 3 = 30.0\]
\[(26 + 25 + 21) / 3 = 24.0\]
\[(15 + 13 + 18) / 3 = 15.3\]

This "smoothed" data can then be reported to the user:


Note the loss of resolution.

A Moving Average Smooth is an improvement on this scheme as it provides for a one-to-one replacement of old data points with new ones. This method of data smoothing averages the data within a "bracket" and then moves the "bracket" one data point down the scan. Each "bracket" contains an odd number of data points and few data points at the ends of the scan are lost or are averaged over fewer data points.

This is a very common technique. However, if the "bracket" sizes are too large, significant signal distortion will set in.

As an example of this smoothing routine, let's perform a 3-point moving average scan on the above data:
The smoothed data is then reported as:

\[ \{5.0, 8.5, 15.7, 18.7, 20.3, 22.7, \ldots\} \]

Data smoothing of this nature is similar to passing the data through a low-pass filter. Signal smoothing will not improve the accuracy of an analyte determination and it is lossy.

Finally, to the specifics of this laboratory exercise. We will determine the S/N ratio for the Thermo Nicolet Avatar 360 Fourier Transform Infrared Spectrometer, considering the effects of spectrometer resolution, signal intensity and number of coadded scans.

Conceptually, an IR spectrometer operates such that IR radiation of intensity \( I_0 \) is passed through a sample and the resulting signal \( I \) is measured using an appropriate detector.

The sample's Transmittance (%T) is then determined using:

\[ \% \text{T} = \frac{I_0}{I} \times 100 \]  

(Eq. 8)

This is done at each wavenumber \( \tilde{\nu} \) for which IR absorbance data is required. \%T vs. \( \tilde{\nu} \) data is then the spectrum of the sample. The fact that our instrument is a Fourier Transform device does not alter this basic conceptual scheme but it does make its description less intuitive. We will deal with what it means to be a Fourier Transform instrument in a later laboratory exercise.
Because our instrument is a single beam instrument, we must first obtain a Background spectrum ($I_o$ vs. $\bar{\nu}$) and then independently determine the Sample spectrum ($I$ vs. $\bar{\nu}$) in order to obtain a %T vs. $\bar{\nu}$ spectrum.

If nothing is in the sample cell when the Sample spectrum is acquired, then the detector should always record a Transmittance spectrum of 100%. Because of instrumental noise, this will not be the case. Deviations from 100% transmittance are the considered to be the instrument's noise. Thus, we can determine the S/N ratio for our IR spectrometer according to (Eq. 1):

$$S/N = \frac{100}{\text{root-mean-square noise}}$$  

(Eq. 9)

We will first determine S/N for a basic single scan signal of fairly high resolution. Then we will vary the instrument's resolution to determine the effect this has on its S/N ratio. We will also increase the number of Coadded scans to confirm (Eq. 7). Finally, we will examine the effect of the instruments signal strength on the S/N ratio. If we examine the Background spectrum, we will notice the signal strength is not constant over the frequency range of the instrument. In fact, the instrument will have its greatest signal strength over the 2200-2000 cm$^{-1}$ range. So, we will compare the S/N ratio for this frequency range to the S/N ratio outside this range.

At this point, you should be wondering, what are the sources of the Noise for an FTIR. This issue will be taken up as a topic of discussion in the lecture. However, you should always be aware of the noise inherent in signals provided by any instrument and how the S/N ratio of the instrument may be affecting your analyses.
Procedure

The laboratory instructor will review how to use the Thermo Nicolet Avatar 360 Fourier Transform Infrared Spectrometer. In particular, you will need to know how to obtain a background spectrum ($I_o$ vs. $\tilde{\nu}$) and a sample spectrum ($I$ vs. $\tilde{\nu}$) for a specified frequency ($\tilde{\nu}$) range. For each spectrum, you should be able to adjust the number of scans performed and the resolution of the spectrum.

Because the software for our spectrometer is not equipped to handle the calculations needed, you will need to be able to export your data to a portable data storage device. This will allow you to analyze your spectral data outside the laboratory.

General Procedure for Collecting a Spectrum

Here we will go through the basics for how to collect a spectrum with a given resolution and over a given frequency range. The specific values we will use in this example will be those needed for the first exercise.

1. Turn on the Nicolet Avator 360 FTIR spectrometer.

2. Turn on the computer and log-in. (Username and Password is physically attached to the computer's case.)

3. Launch the OMNIC 32 software package.
4. *Click* on "Collect" to view its drop-down list of options and *select* "Experimental Set-Up".

5. On the "Set-Up" page:

![Experiment Setup Interface](image)

Set the options as follows; for our first experiment:

- No. of Scans = 1
- Resolution = 0.5
- Final format = %Transmittance
- Correction = None
- Save automatically = checked
- Save Interferograms = checked
- Base Name = "Your Initials"
- Background Handling = "Collect background before every sample"
- Experiment Title = "Noise"

Now "Save" the settings and then *click* on "OK".

Set the "Spectral Range" to "2200" to "2000".

"Save" the settings and then click on "OK"

7. "Exit" the "Set-Up" window.

8. Make sure that no sample is in place and the "Viewing Window" is closed.

9. Collect a Background spectrum and save it as a csv text (*.csv) file in the "411" folder. (You should create your own sub-folder and save your spectra in this folder. These subfolders will be periodically cleared away.)

10. Collect a Sample spectrum and save it as a csv text (*.csv) file in the "411" folder.
**Simple S/N Calculations and Smoothing**

Here we will collect both a "Background" spectrum and a "Sample" spectrum. We will take a single scan for each case; using a fairly high resolution. We will limit ourselves to a frequency range of 2200-2000 cm\(^{-1}\). This is the frequency range over which the instrument delivers its maximal power to the sample. We will then use Moving Average and Boxcar smoothing procedures to reduce the Noise.

1. Collect a background spectrum at 0.5 cm\(^{-1}\) resolution using 1 scan. Likewise, collect a Sample spectrum at 0.5 cm\(^{-1}\) resolution using 1 scan. Save the Sample spectrum to a *.csv file so that the S/N ratio can be determined over the 2200-2000 cm\(^{-1}\) at a later time.

**Effect of Resolution on S/N**

Here we will collect spectra using different instrumental resolutions. We will start with a fairly high resolution (0.5 cm\(^{-1}\)) spectrum and proceed to fairly low resolution spectra (32.0 cm\(^{-1}\)). We will then examine how the S/N ratio depends on the resolution. We will also examine how the S/N ratio is influenced by

2. Averaging 8 background and sample scans, obtain spectra at resolutions of:

   0.5 cm\(^{-1}\), 1.0 cm\(^{-1}\), 2.0 cm\(^{-1}\), 4.0 cm\(^{-1}\), 8.0 cm\(^{-1}\), 16.0 cm\(^{-1}\), 32.0 cm\(^{-1}\)

   You will need to run a Background spectrum for each spectral resolution. All of these spectra should cover the full frequency range of the instrument. This will allow us to examine how the S/N ratio is influenced by the power deliver to the sample. (So, you will need to go back into the "Bench" set-up and reset the Spectral Range to 3000-400 cm\(^{-1}\).) Save the Background spectrum at 0.5 cm\(^{-1}\) resolution to a *.csv file. (This will allow you to determine the instrument's signal strength over the entire frequency range of the instrument.) For each Sample spectrum, export your data to a *.csv file for further analysis.

**Effect of Coadding Scans on S/N**

Here we will examine the effect of scan coaddition on the S/N ratio.

3. Set the resolution at 4.0 cm\(^{-1}\); the resolution used for our first exercise. Collect background and sample spectra using the following number of scans:

   1 scan, 4 scans, 16 scans, 64 scans, 256 scans, 512 scans

   As usual, export the data for further analysis.
Data Analysis

Simple S/N Calculations and Smoothing

1. Using the "Sample" data from which the "Background" has already been subtracted, determine the S/N ratio over the 2200 cm\(^{-1}\) to 2000 cm\(^{-1}\) frequency range.

2. Perform a 5 point Moving Average Smooth on the above spectrum, plot the smoothed spectrum and re-calculate the S/N ratio. Recalling the unsmoothed spectrum each time, perform the following Moving Average smoothing routines:

9-point, 13-point, 17-point, 21-point, 25-point

Calculate the S/N ratio for each smoothed spectrum. (You do not need to plot each smoothed spectrum.) Plot the S/N ratio versus the number of points used in the spectral smoothing. Comment.

3. Perform a 5 point Boxcar average on the unsmoothed spectrum, plot the averaged spectrum and re-calculate the S/N ratio. Comment.

Effect of Resolution on S/N

4. Plot the "Background" spectrum from procedure step (2). This will allow you to determine \(I_o\) vs. \(\bar{v}\); and, hence, the regions of the spectrum where the signal strength is great and where it is weak.

5. Determine the S/N for each spectrum obtained at the different spectral resolutions (step (2) above) for the following three frequency ranges:

- 3000 - 2800 cm\(^{-1}\)
- 2200 - 2000 cm\(^{-1}\)
- 600 - 400 cm\(^{-1}\)

Recall, the spectrometer delivers its greatest power to the sample over the 2200 cm\(^{-1}\) to 2000 cm\(^{-1}\) range.

Plot the S/N ratio versus Spectral Resolution for each frequency range on the same graph. Examine the single beam signal intensities for each of the above frequency ranges. This can be done by examining the \(I_o\) vs. \(\bar{v}\) Background spectrum (#4 above). Comment on the effect of resolution and signal strength on the S/N ratio.
Effect of Coadding Scans on S/N

5. Determine the S/N ratio for the 2200 - 2000 cm\(^{-1}\) for each spectrum obtained in procedure step (3) above. Plot the S/N ratio versus Number of Coadded Scans. Fit the data to a "Power Function." Compare this functional form with that which is expected.

You should submit a Memo Style report of the above findings. Be sure to include the "Background" spectrum from procedure step (2). Also include all plots requested.
References


