Key

LAB #3

Noise Reduction Strategies

PART I: Principle Types Of Noise

A. Thermal, Johnson or Constant Noise - noise produced by resistors

Noise has its origin in different parts of an instrument. For example, **Thermal**, **Johnson** or **Constant** noise occurs in the electronic components of the instrument such as in resistors. Even in the absence of a current, thermal vibrations and collisions of charge carriers produces random voltage fluctuations given by:

$$V_{rms} = \sqrt{4\,k\,T\,R\,B}$$

where
- Vrms = root mean square voltage
- k = Boltzmann's constant (1.38×10^{-23} J/K)
- T = absolute Kelvin temperature
- R = resistance in Ohms
- B = signal bandwidth

Johnson noise has a "white" frequency spectrum, that is, all frequencies are present in equal amplitudes. Notice that Johnson noise is not dependent on a particular frequency, just on the range of frequencies present (the bandwidth).

B. Shot Noise - Noise from current producing devices

Another common type of instrumental noise is **Shot Noise**. Shot noise occurs in devices which measure the number of discrete units or charges passing across a potential. A photomultiplier tube (PMT) is a good example of an electronic device which produces shot noise. PMT's will be discussed in more detail later on in the course. For now, look at Figure 1. A PMT works on the photoelectric effect. Light incident upon an Na/K/Sb alloy (photoemissive cathode in the figure) will photoeject electrons. The electrons are accelerated through a series of positively charged electrodes called "dynodes". The dynodes also photoeject electrons which results in a chain reaction. The chain reaction continues until one incident photon can result in several million flowing out of the PMT. A PMT produces lots of shot noise if you aren't careful! The shot noise current is given by:

$$i_{rms} = \sqrt{2\,I\,e\,B}$$

where
- i_{rms} = root mean square shot noise current
- I = average current over time
- e = charge on the electron (1.6×10^{-19} C)
- B = frequency bandwidth

Like Johnson noise, shot noise has a white frequency spectrum, but there is one important difference. Since shot noise depends on the average current coming from the device, it will increase as the square root of the magnitude of the signal. Johnson noise does not have this property since the signal voltage can increase across a resistor without changing the resistance.

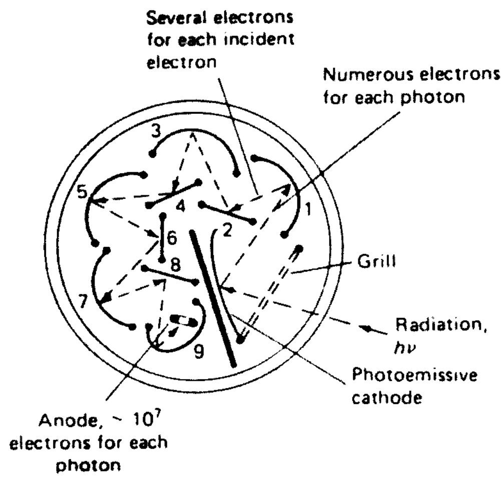


Figure 1: Configuration of a Photomultiplier Tube

C. Flicker or 1/f Noise - Low frequency noise from lamps

A third common type of noise is FLICKER or 1/f NOISE (pronounced "one over f- noise"). The source of flicker noise is unknown, although it interesting to note that it is easy to produce using the concepts of fractals and nonlinear dynamics. The slow variation in light output from a lamp is an example of 1/f noise. At frequencies less than 100-300 Hz flicker noise starts to become noticeable. 1/f noise can be removed by modulating the signal at high frequencies by using a LOCK-IN AMPLIFIER. We will discuss lock-in amplifiers in more detail in lecture when we cover Atomic Spectroscopic methods.

Exercise #1 – Investigating the Properties of Constant Noise vs. Shot Noise

Run the program Microsoft Excel on the computer. After the program starts, enter the file **noise.xls**. This spreadsheet will produce signals with constant (Johnson) and shot noise. Under the column labeled "Johnson Noise" there are three parameters T,R, and F (for Kelvin temperature, resistance in Ohms, and frequency bandwidth). The parameters are set at 298 K, 50 Ohms, and 10 Mhz. Noise produced under these conditions has been added to a Gaussian peak shape typical of chromatographic separations. A plot of this noisy signal is shown in columns E through I.

Estimate the noise signal and peak signal level in the Johnson noise data table below using the definitions given in Figure 2. Then calculate the Signal to Noise ratio.

Move the "cell pointer" to the temperature parameter so that 298 is highlighted. Enter a value of 100. What happened to the magnitude of the noise? Do the same process with a temperature of 600 K. Enter the results in the data table below. Now change the resistance to 1000 Ohms while keeping T at 298 K. Record your observations. Finally, change the frequency bandwidth from 10 to 100 MHz while keeping the value for T at 298 K and R at 50 ohms. Enter your results in the table.

Look at the bottom left-hand corner of the spreadsheet. Press the tab labeled "Shot Noise". The new sheet displays a graph of shot noise on the same Gaussian peak. Notice how the magnitude of the noise at the top of the peak is greater than on the wings. This is typical of shot noise. Vary the parameters I and F and record the results in the Shot Noise Data Table.

Johnson Noise Data Table

	noise signal level	**peak signal level**	**S/N**
T=298 K R=50 Ω F=10 MHz	0.1	1	10
T=100 K R=50 Ω F=10 MHz	0.05	1	20
T=600 K R=50 Ω F=10 MHz	0.15	1	6.7
T=298 K R=1000 Ω F=10 MHz	0.6	1	1.7
T=298 K R=50 Ω F=100 MHz	0.25	1	4

Shot Noise Data Table

	signal at top of peak	**noise at top of peak**	**S/N**
I= 1 mA F= 10 MHz	1	0.15	6.7
I=100 mA F= 10 MHz	1	2	0.5
I= 1 mA F=100 MHz	1	0.6	1.7

PART II: Signal Averaging

Imagine two oscilloscope traces of a noisy square wave signal in time. At a given point in time, let's say that the noise has fluctuated positively in the first oscilloscope trace, and negatively in the next (Figure 2a). This will happen in reality because of the symmetric, random nature of noise. If we averaged these two signals together, the positive and negative fluctuations should partially cancel out leaving us with a smoother signal. This is the basis for "signal averaging". Many modern **digital oscilloscopes** allow the user to take many sequential samples of a signal, add them together and store the result in memory. These instruments are similar to the analog oscilloscope used in the electronics lab, but with added memory and capability.

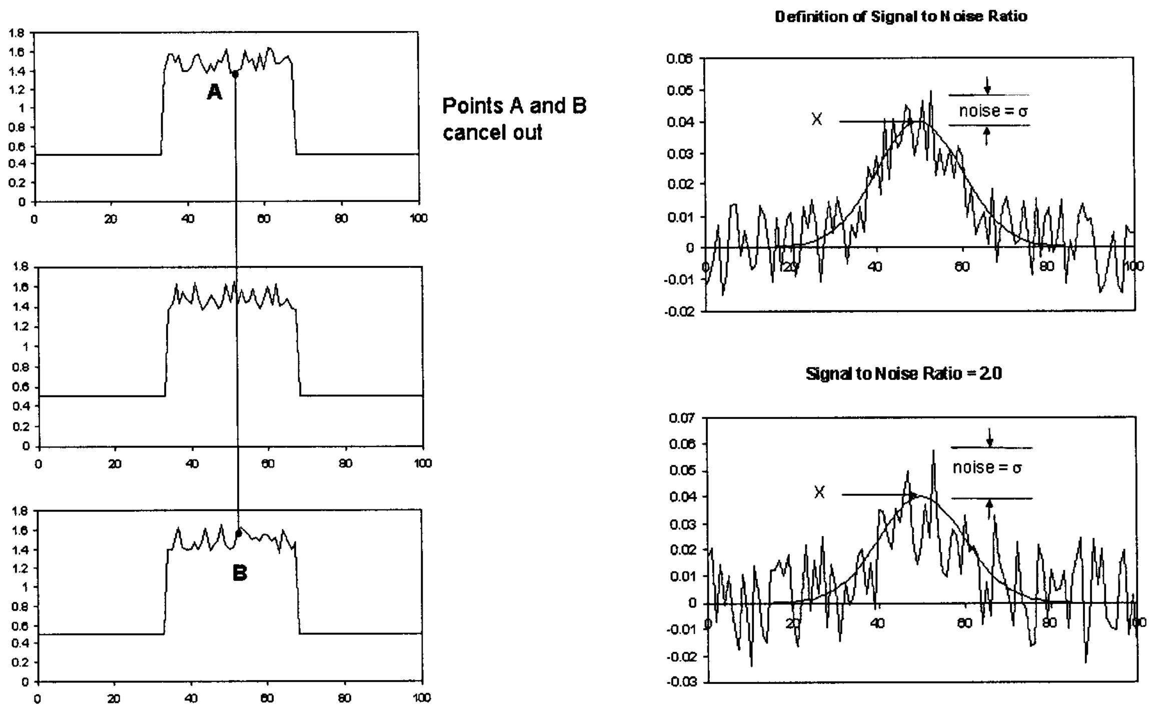


Figure 2: (a) cancellation of random noise when added, (b) definition of the S/N

After we average a signal N times, we observe an improvement in the S/N of the waveform (Figure 2b). The noise magnitude will be decreased by a factor of $1/\sqrt{N}$. This can be shown as follows. The standard deviation of the signal is a measure of the spread in the measurements, and therefore is equal to the noise level in the signal. The standard deviation σ for a <u>large</u> data set with average <x> is equal to -

$$\sqrt{\frac{\sum_{i=1}^{N} (x_i - <x>)^2}{N}}$$

The S/N is therefore -

$$\frac{S}{N} = \frac{x_i}{\sqrt{\frac{\sum_{i=1}^{N} (x_i - <x>)^2}{N}}} = \frac{\sqrt{N} x_i}{\sqrt{\sum_{i=1}^{N} (x_i - <x>)^2}}$$

And so the S/N increases with the square root of the number of averages N. For example, if you want to increase the S/N ratio by a factor of 10, you must average the signal 100 times. For extremely noisy signals such as astronomical observations, you may have to average tens of thousands of signals together to obtain a satisfactory S/N.

Exercise #2 – Improvement in the Signal to Noise Ratio by Averaging

Once again, run the Microsoft Excel on the computer. Input the file **average.xls** into the spreadsheet. Hold down the **CONTROL** key while pressing the "o" key (not the 0 key!). Let go of the keys. A graph of a noisy, unaveraged signal will appear along with the underlying pure sine wave. Estimate the noise magnitude where the sine wave is at a maximum, and enter it in the Signal Averaging Data Table below. Now press the CONTROL key followed by the "f" key. Displayed on the screen will be the original clean signal plus the noisy signal averaged 4 times. Enter the noise magnitude in the data table. Now hold down the CONTROL key while pressing the "e" key. Now you see the data averaged 8 times. Continue this procedure by sequentially pressing the CONTROL + "s" (16 averages), CONTROL + "t" (24 averages), and finally CONTROL + "h" (32 averages). You can view any of the graphs at any time by clicking on them (after "wading through" the stack of graphs piling up in the spreadsheet).

Plot each S/N versus the square root of N. Hand in a copy of this graph along with this lab. Ask me for help if you do not know how to do this.

attached

Signal Averaging Data Table

# averages	Signal*	Standard Deviation (noise)	S/N (y-axis)	square root of the number of averages (x-axis)
1	1	0.56	1.8	1
4	1	0.27	3.8	2
8	1	0.20	4.9	2.83
16	1	0.15	6.5	4
24	1	0.12	8.2	4.9
32	1	0.10	9.7	5.66

*The signal magnitude at the maximum of the sine wave is equal to one in all examples

actual values
from spreadsheet
(you estimated from graph)

Exercise #3 – Digitizing Oscilloscopes

You have seen signal averaging work in theory, now it's time to see it with a real, noisy signal. Connect a 1 KHz sine wave to the white noise generating box, and then connect the output of the white noise generator to channel one of the Hewlett-Packard digitizing oscilloscope. Using a BNC T-connector, input the signal into channel 2 as well.

The digitizing oscilloscope is controlled using the keypad buttons on the front panel. There is a line of Menu buttons on the top of the panel. Press the **Status** button which is the second button from the left. Using the **Next [ ]**, **Prev [ ]**, cursor arrow keys, and the numeric keypad, enter the values below into the Status screen. After the scope is finished averaging, plot out a graph of the screen and hand it in with the lab.

Channel--Timebase---------------------

	Input 1	Input 2		
			Sampling @ 200 kHz	
Range	4.0 V	4.0 V	Mode [Auto]	
Offset	0.000 V	0.000 V	Range 5.00 ms	
Label			Delay 0.000 s	
Coupling	[ac] [1:1]	[ac] [1:1]	Reference	[Left]
Store Mode	[Normal]	[Ave] [256]	Alias Test	[Disabled]
Auto Scale	[Enabled]	[Enabled]	Auto Scale	[Period]

Trigger---

Source [Chan 1]		Auto Scale	[Disabled]
Level 650.0 mV	[+ Slope]	Store Mode	[Ave] [256]
Range [Track]	4.0 V	Coupling	[ac] [1:1]
Label			

PART III: Fourier Transforms and the Convolution Theorem

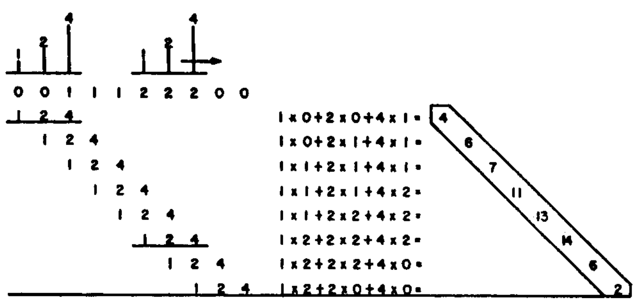
Sometimes an instrument will "spread out", broaden, or otherwise alter a signal because of its "instrument response". This process can be illustrated as follows:

$$S_i(t) \Rightarrow \boxed{R(t)} \Rightarrow S_o(t)$$

An input signal $S_i(t)$ is combined or "convoluted" with the instrument response function f(t), producing an output signal $S_o(t)$. Convolution is mathematically defined as -

$$S_o(t) = \int S_i(t)\, R(t-x)dx = S_i(t) * R(t)$$

where the integration limits are from minus to plus infinity and the asterisk denotes "convolution". This effect of the convolution integral is best explained using an example as shown in Figure 3. When two functions are convoluted together, the first function is reversed in time and then multiplied point by point times the second function. The delay "x" between data points is incremented and the process repeated.



Convolution of sampled signals (1, 1, 1, 2, 2, 2) and (4, 2, 1). The first time series represents the input signal, the second the impulse response, and the eight-term series shown as a diagonal the output. Note reversal of second term (*From Peterson and Dobrin.*[4])

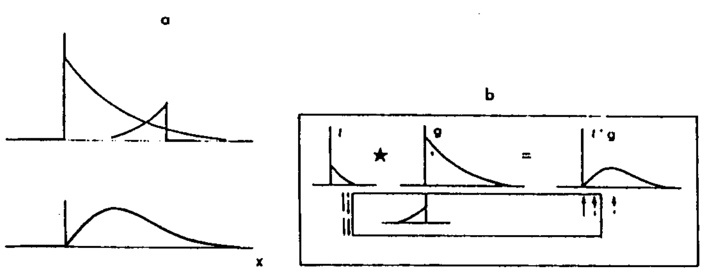


Figure 3

A good example of convolution of a signal by an instrument response function is given by the measurement of a short laser pulse using BNC cable, a photomultiplier tube, and an oscilloscope (Figure 4). The BNC cable has stray capacitance, which combines with the oscilloscope input termination resistance of 50   to form an RC filter. The instrument response of such a system is given by $e^{-t/RC}$. The temporal profile of a laser pulse is Gaussian in shape. The output signal $S_o(t)$ is the convolution of the Gaussian with the exponential. The net effect is to broaden the width of the measured laser pulse.

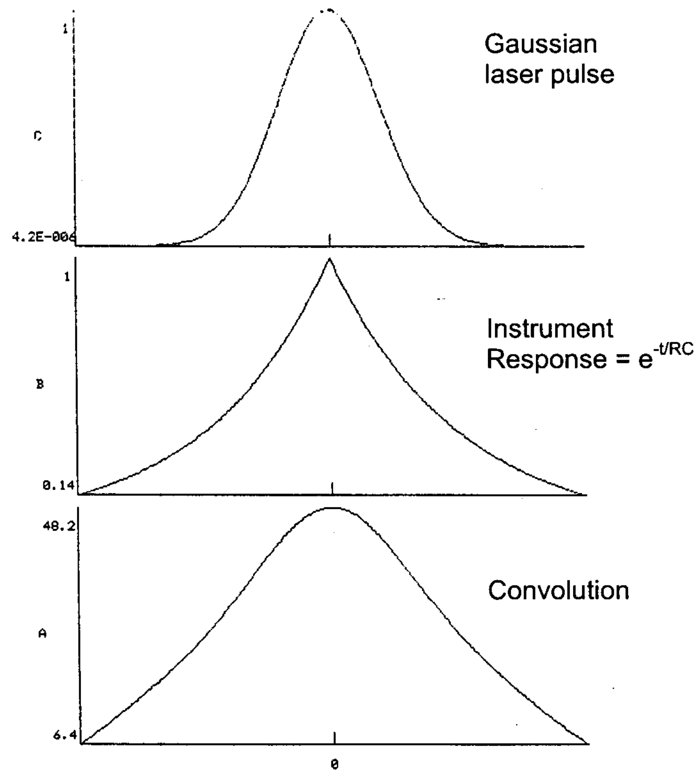


Figure 4

There is a relationship between the convolution of two functions with their Fourier transforms that is known as the Convolution Theorem. This very useful theorem can be stated as: "the convolution of two functions is equal to the product of their Fourier Transforms" -

$$S(t) * R(t) = S(\omega) \cdot R(\omega)$$

where the FT of $S(t) = S(\omega)$ and the FT of $R(t) = R(\omega)$. We will apply this theorem numerous times throughout the class. In the next section we will see how moving average and Savitsky-Golay filters are actually examples of convolution of weighting coefficients with the noisy signal.

PART IV: Finite Impulse Response (FIR) Filters

A Finite Impulse Response filter is defined over only a short interval around the data point to be smoothed. The filter is generally defined as -

$$X_n = \sum_{i=-m}^{+m} F_i x_i$$

where the interval around the point to be smoothed consists of 2m+1 data points, the F_i 's are filtering coefficients, the x_i are the data points, and X_n is the smoothed data point. The filter is illustrated in Figure 5. The filtering process consists of selecting a symmetrical interval about the data point to be smoothed, and sequentially multiplying each data point in the interval by it's corresponding smoothing coefficient.

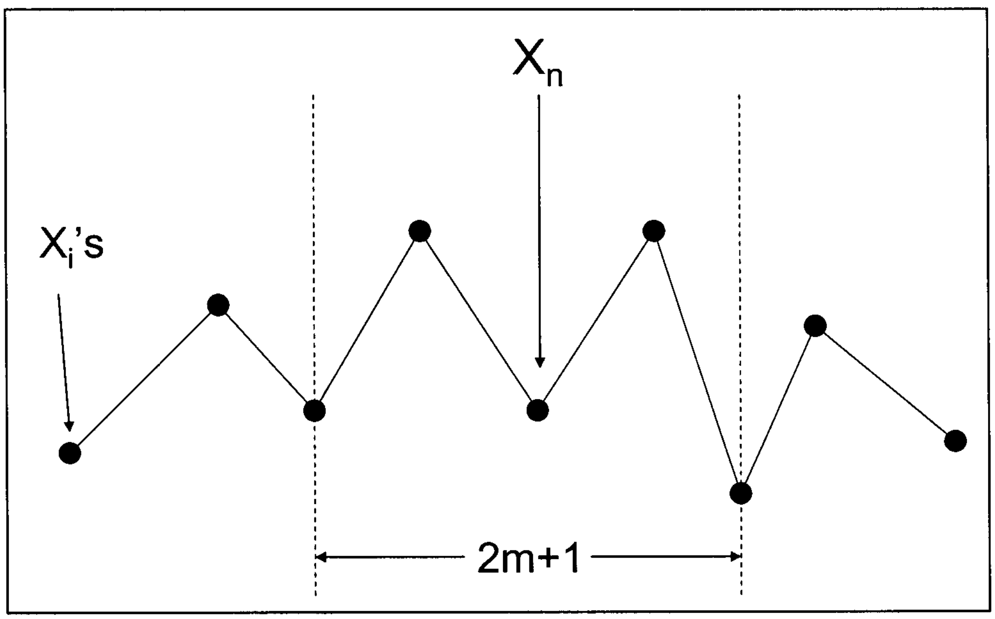


Figure 5

A. Moving Average Convolution Filters

In a moving average smoothing filter, all the filtering coefficients F_i are equal to one. The filter simply consists of replacing the data point to be smoothed by the average of the data points in the selected interval around it. A plot of the filtering coefficients is a rectangle centered around the data point to be filtered as shown in Figure 6.

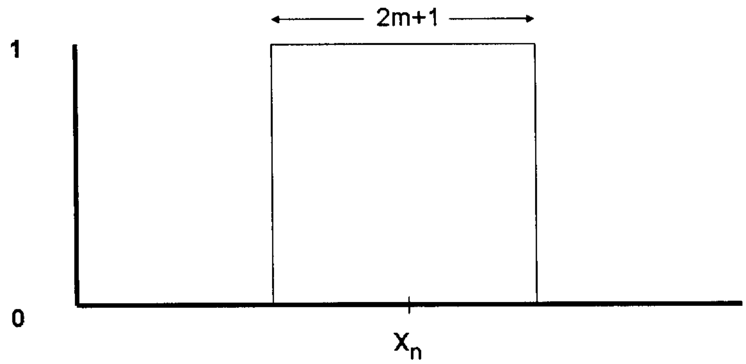


Figure 6

Exercise #4 – Implementing Moving Average Convolution Filters in Microsoft Excel

1. Start Microsoft Excel. Enter the worksheet file **filters.xls** from the directory **Ch471**. In column B is a signal of a noisy Gaussian peak as might be encountered in a chromatographic separation.

2. Move the cell pointer to cell c6 and enter the formula **=average(b4:b8)**.

3. Copy the formula from cell c7 to c101.

4. Using the chart wizard, plot the data in columns A-C. You should see the original noisy Gaussian with the moving average smooth superimposed on top as shown in Figure 7(a).

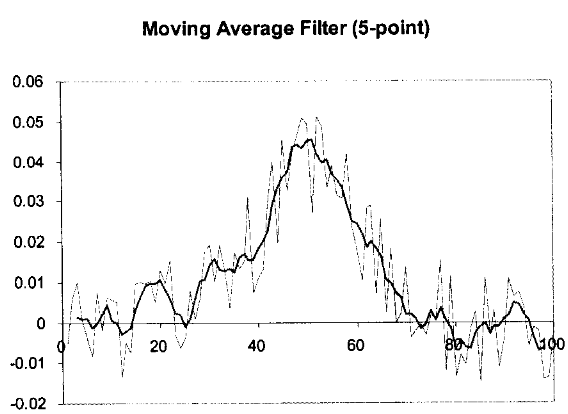


Figure 7(a)

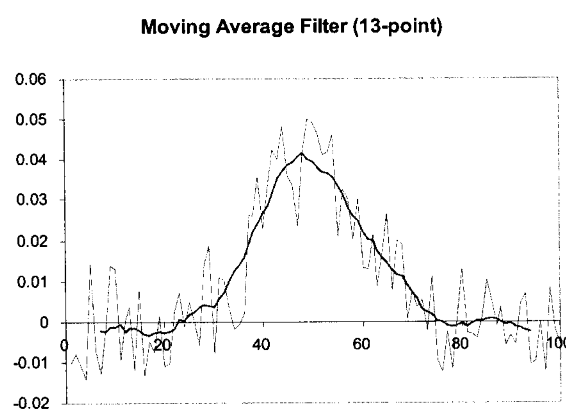


Figure 7(b)

5. Move to cell d10 and enter the formula **=average(b4:b16)**. Copy the formula from cells d11 to d97. Plot the data and smooth using the chart wizard. You should see a plot like Figure 6(b). By extending the smoothing range, the data has an even greater signal to noise ratio.

6. Print out your graph and hand in with the exercise.

7. Now you will use the convolution theorem to perform the same Finite Impulse Response (FIR) filtering operation. We will take advantage of the fact that the convolution of a noisy signal with a rectangular moving average filter is equal to the product of their Fourier transforms.

B. Savitzky-Golay Convolution Filters

Exercise #5 – Implementing Savitzky-Golay Convolution Filters in Microsoft Excel

8. Return to Excel and the spreadsheet filters.xls. Move to cell e6 and enter the formula -

   = -.0857*e4+.3429*e5+.4857*e6+.3429*e7-.0857*e8

   Watch those negative signs! Copy the formula down to cell e101. Plot the data and smooth. A polynomial smooth such as this one follows the finer details in the data instead of smoothing them out. That's the value of a Savitsky-Golay filter - it preserves fine spectral detail. If you want more smoothing, then increase the size of the filter.

9. Print out your graph and hand in with the exercise.

Exercise #6 –Convolution Filters via Digital Filtering and the Convolution Theorem

10. Start the program **LabVIEW Student Edition 7.0** by double-clicking the desktop icon. Open the "virtual instrument" (also called a "vi") FIR.vi. You will learn how to construct vi's in the next lab on programming A/D converters.

11. To start the program, press the arrow icon at the top of the window. A dialog box opens; change the directory to **c:\ch471** and enter the text file **"MA5pt.txt"**. Then when the second dialog box opens, select "gauss.txt". The file MA5pt.txt contains the moving average convolution filter with a width of 5 points, so it's a 5-point moving average filter.

12. Scroll down to the bottom of the window to observe the results of the convolution process.

13. Repeat step 11 for the remaining moving average filters given in Table 1 on the next page.

Table 1: Moving Average Convolution Filters

Moving Average Filter	Estimate S/N	Sketch FFT of Filter*	Change in Width?
5-point			
13-point			
33-point			
65-point	increasing	narrower with less high frequency response	
*Make sure that each sketch illustrates relative widths of each filter			

14. Now repeat the process for the first Savitzky-Golay filter file named **"SG5_2_3.txt"**, and then complete Table 2.

Table 2: Savitzky-Golay Convolution Filters

Moving Average Filter	Estimate S/N	Sketch FFT of Filter*
5-point, $2^{nd}/3^{rd}$ order		
13-point, $2^{nd}/3^{rd}$ order		
21-point, $2^{nd}/3^{rd}$ order	increasing	narrower with less high frequency response
13-point, $4^{th}/5^{th}$ order	actually worse	
*Make sure that each sketch illustrates relative widths of each filter		

A REVIEW: Using a convolution filter consists of three steps:

1. Separately Fourier transform the data and convoluting filter
2. Multiply the transforms together to obtain the product.
3. Inverse Fourier transform the product back.

PART V: Infinite Impulse Response (IIR) Filters Using LabVIEW

Infinite Impulse Response filters differ from FIR filters in that the frequency response of the filter never becomes zero. This means that we cannot apply the summation formula seen previously. The algorithm that is used to practically apply an IIR filter is *recursive* in nature in that the present response depends on past iterations. Without going into the details, it is sufficient to say that our convolution ideas are still valid here. A IIR filter is still Fourier transformed, multiplied times the transformed data, and inverse Fourier transformed. You are going to use a program called LabVIEW to do it for you.

Two important IIR filters are shown in Figure 8. The first type of filter is known as a **Butterworth** filter. The Butterworth filter [Figure 15(a)] is defined as:

$$\frac{V_{in}}{V_{out}} = \frac{1}{[1 + (f/f_c)^{2n}]^{1/2}}$$

where n = the order of the filter (number of "poles"), and f_c is the critical cut-off frequency. Note that the sharpness of the cut-off region increases with the order of the filter.

The second kind of filter is the **Chebyshev** filter. Chebyshev filters trade off an increased sharpness of the cut-off frequency with *passband ripple.* In other words, we tolerate some signal amplitude distortion for less noise. Chebyshev filters are defined by -

$$\frac{V_{in}}{V_{out}} = \frac{1}{[1 + \varepsilon^2 C_n^2 (f/f_c)]^{1/2}}$$

where the magnitude of the ripple is given by ε, and C_n is a Chebyshev polynomial of degree n defined as $C_n(x) = \cos(n \arccos x)$

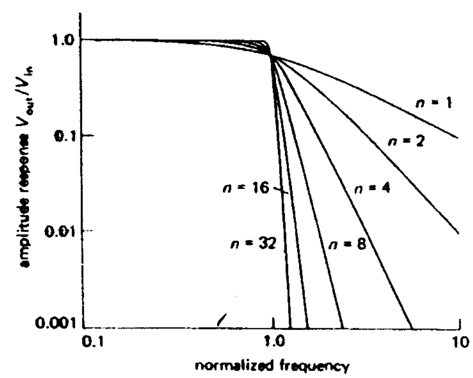


Figure 8(a)

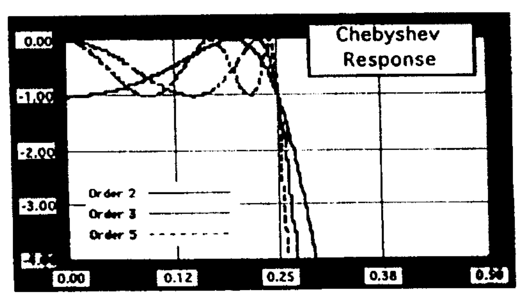


Figure 8(b)

Exercise #6 – Experimenting With Butterworth and Chebyshev IIR Filters

15. From Windows, start the program LVSE31 (**LabVIEW Student Edition 3.1**, an older version of LabVIEW with some nice instructional vi's). From the file menu, select open followed by the directories **examples** and **analysis**. Open the "vi" **IIR Filter Design**. The screen in Figure 9 will be displayed.

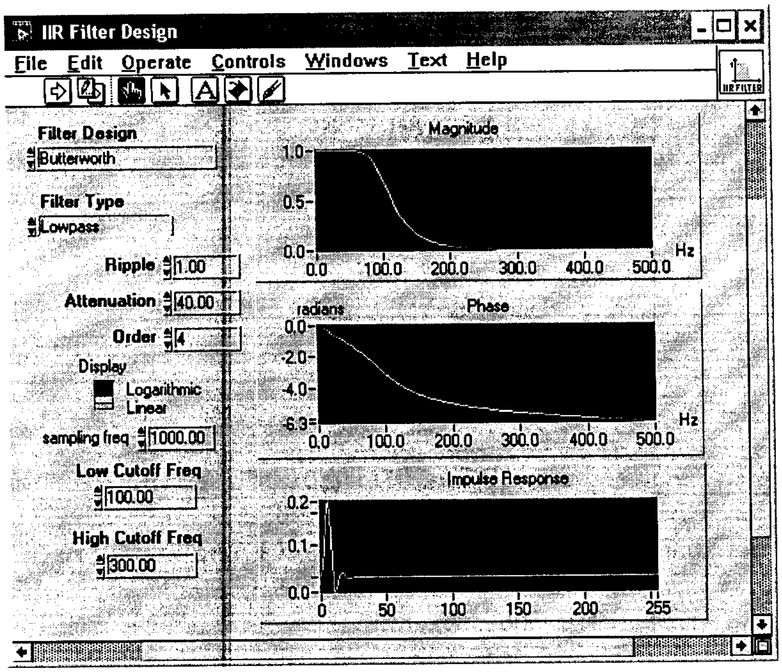


Figure 9

16. Make sure that the "hand cursor" is displayed; different cursors can be selected from the toolbar at the top of the screen. Start the program by clicking on the right-hand arrow.

17. A **Butterworth** filter is calculated. Change the filter type from lowpass to highpass. Adjust the ripple, attenuation, and order. Finally, adjust the low and high cutoff frequencies. Write a short description and explanation of the changes you observe and what is causing them in question ~~#13~~ at the end of the lab handout.
#15

18. Now enter the file **IIR.VI** from the directory **c:\ch471**. Enter the following parameters in the input boxes on the right side of the "instrument panel" -

Filter design = 0	Sampling frequency = 1000
Filter type = 0	Low cutoff freq = 100
Ripple = 1.00	High cutoff freq = 300
Attenuation = 40.00	
Order = 5	

19. To start the program, click on the right arrow in the toolbar. Then flip the toggle switch up to the start position. Press the right arrow again. Now turn the control dial to increase the noise level and look at the graph windows. The top left graph contains the noise free sine wave with the noisy signal superimposed on top. The upper right window contains the power spectrum. The bottom right graph is a display of the filtered sine wave.

20. A filter design of 0 corresponds to a Butterworth filter, and a filter type = 0 corresponds to low pass. Change the filter design to 1 for a Chebyshev filter. Play with the ripple also. Describe and explain your observations in question #14.

PART VI: Questions

1. Explain how can you decrease (a) thermal noise, and (b) shot noise.

(a) decrease T, resistor value, and signal bandwidth

(b) decrease average current out of detector and signal bandwidth

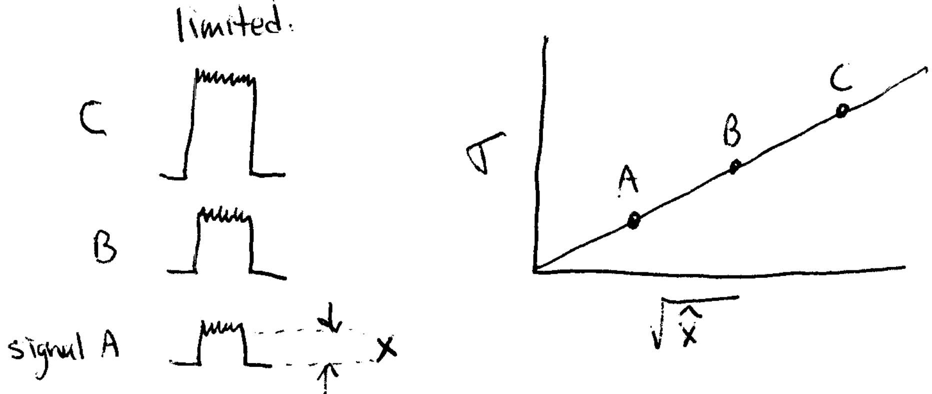
2. What does it mean to be "shot noise limited"?

All of the constant noise has been suppressed (using techniques in #1)

3. How can you determine if the noise in your signal is only shot noise? Draw a graph to explain your answer.

The noise distribution is Poisson, not Gaussian. So $\sigma = \sqrt{\hat{x}}$

If a plot of σ vs. $\sqrt{\hat{x}}$ is linear, then you are shot-noise limited.



4. What type of instruments are inherently less noisy, those with shot-noise limited detectors or thermal noise? Give examples of several types of instruments that fall into each category.

shot-noise limited less noisy = PMT or photodiode detectors e.g. fluorometer

thermal noise more noisy = FTIR

5. How could the bandwidth of a signal be changed electronically (think back to the analog electronics lab) ? Determine the values of the electronics components needed to produce a bandwidth of 20 MHz.

Build an RC Filter where $f_c = 20\,MHz$

$$f_c = \frac{1}{2\pi RC}$$

pick $R = 1K\Omega$ and solve for C (units must be in Ω's and Farads)

$$20\times10^6\,\Omega = \frac{1}{2\pi(1000\,\Omega)C}$$

$$C = 8\,pF$$

6. Mathematically state how the signal-to-noise ratio varies with the number of averages.

$$S/N \propto \sqrt{N}$$

e.g. 100 averages then $\sqrt{100} = 10X$ improvement

7. Mathematically state the Convolution Theorem.

$$f_1(t) * f_2(t) = \mathcal{F}^{-1}\left[\mathcal{F}_1(w) \cdot \mathcal{F}_2(w)\right]$$

Answer questions 8-13 based on Tables 1 and 2.

8. Define what a Finite Impulse Response Filter (FIR) is.

The frequency response drops off to zero.

9. Why was the 5-point moving average filter better than the 5-point Savitzky-Golay filter?

The SG filter had more amplitude at higher frequencies, allowing more high-frequency noise to remain in the signal.

10. Why did the moving average convolution filter distort the width of the original data?

SKIP

11. Which moving average convolution filter had the best S/N and why? ~~Was there a trade-off between S/N and width distortion?~~ ← SKIP

The greater the number of points (width) the better the S/N. Again, the S/N improves for the same reason as in #9

12. Which 2nd/3rd order Savitzky-Golay filter was best and why?

The one with the greatest width for the same reason as #9

13. Compare and contrast the two 13-point Savitzky-Golay filters.

Surprisingly the lower order was better

Questions about Infinite Impulse Response (IIR) Filters.

14. Define what an Infinite Impulse Response Filter (IIR) is.

The frequency response never drops off to zero.

15. Give your observations and explanations from step 17.

$$\frac{V_{in}}{V_{out}} = \frac{1}{[1+(f/f_c)^{2n}]^{1/2}}$$

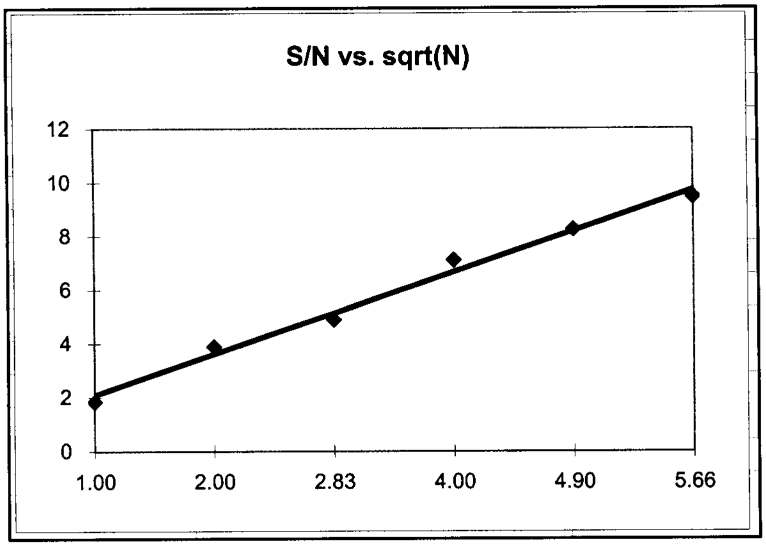
ripple and attenuation have no effect

Increasing order (n) steepens cutoff and improves S/N by more efficiently removing high frequency noise

16. Give your observations and explanations from step 20.

$$\frac{V_{in}}{V_{out}} = \frac{1}{[1+\varepsilon^2 C_n^2 (f/f_c)]^{1/2}}$$

increasing order (n) steepens cutoff frequency at the cost of more rippling - and therefore distortion of the input signal. This can be mitigated by increasing the attenuation (decreasing ε) of the filter


S/N vs. sqrt(N)
12
10
8
6
4
2
0
1.00
2.00
2.83
4.00
4.90
5.66