Learning objectives and goals:

In this problem set, we will analyze data recorded in the scalp electroencephalogram (EEG) of a rat that is a genetic model for human absence epilepsy. The EEG provides a measure of brain voltage activity with high temporal resolution (typically on the order of milliseconds) but poor spatial resolution (on the order of 10 cm^2 of cortex). Here we consider EEG activity recorded from a single scalp electrode located in the frontal cortex. We will analyze these data to determine what (if any) rhythmic/periodic activity is present. Typical rhythms in EEG recordings span the range of 2-80 Hz. It is important to mention, that EEG recordings are typically contaminated with 60 Hz noise coming from the electrical outlet. This exercise will provide us an opportunity to apply the concepts of filtering, time series analysis and spectral analysis, to this data set.

These topics were covered in lectures 10-13 and recitation 8. The accompanying MATLAB code to the lecture notes will be a very valuable resource to complete this PSET.

Our goal is to analyze a single recording of EEG data by characterizing the observed activity. There are many ways to do so, and we focus on developing techniques to characterize the observed rhythms. By the end of this PSET, you should be familiar with:

- time series analysis techniques,
- the principles behind Fast Fourier transform (FFT)
- computing the periodogram and plot power on a linear and decibel (dB) scale.
- filtering signals in the frequency domain and take the inverse Fourier transform.
- computing and visualizing multi-tapered spectra.
- producing spectrograms to analyze signals whose spectrum changes over time.

MATLAB functions you will need and advice for multi-taper analysis:

The following functions will be helpful:

`fft`	Compute Fast Fourier Transform
`ifft`	Compute inverse Fast Fourier Transform
`xcorr`	Compute autocorrelation
`abs`	Compute magnitude of a complex number
`surf`	to visualize spectrograms
`log10`	Compute logarithm in base 10
`semilogx`	Plot y versus x on an log-scale x-axis

We have provided two functions, `WSpec` and `WSpecgram` to compute spectra and spectrograms respectively, using a multi-taper approach. To become familiar with the syntax of these functions type ‘`help WSpec`’ (or `WSpecgram`) on the MATLAB command window, or check the code accompanying lecture notes. In this PSET, we will guide you through the process of choosing parameters for using these functions. In real life, when analyzing your own data, you will need to explore different parameter combinations to obtain a set of visualizations that help you interpret your data in a meaningful way.

Here are some general guidelines so that your computations do not take forever when exploring parameter combinations.

Both functions are computationally intensive routines, but you can mitigate this if you understand how the parameters affect its operation.

Remember that `WSpecgram` is doing the same operation as `WSpec` for many overlapping time windows, specified by `Twin` and `Tstep`. The smaller your `Tstep`, the longer your program will run. A simple way to make sure your code basically works is to use a large `Tstep` (up to `Twin`) at first, and then reduce it once you know it is working to improve the resolution.
Another thing to note is that your computer/graphics card may have issues plotting the spectrogram if the resulting matrix is big. The number of columns in this matrix is controlled mostly by the step size.
The other parameters are `P`, the time-bandwidth product, and `Zpad` the zero-padding factor.
The larger `P` is, the more tapers, and therefore the more data generated per time window, which will slow computations (more FFT's to compute). Start with a smaller `P` (1 or 2) and you can increase later to improve the quality of your spectrogram if necessary.
The same logic applies for the zero-padding factor (i.e. start without zero-padding).

PART 1: Time series analysis and periodogram.

The EEG signals in file `EEGepilepsy.mat` were obtained from electrodes placed on the right frontal cortex of a male adult WAG/Rij rat[1], which is a model for human absence epilepsy. The file contains two data arrays: `epileptic`, the EEG scalp recording in units of microvolts during an epileptic event characterized by spike waves; and `normal`, the EEG scalp recording in units of microvolts in the absence of epileptic seizures. Both signals wEEgepilepsyhere referenced to an electrode placed in the cerebellum and sampled at 1 kHz.

We will start by analyzing the epileptic recording. Load the data to MATLAB and complete the following tasks:

1. **What is the sampling interval for this recording? What is the total duration of the recording?**

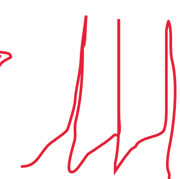
   Sampling interval : 1 ms
   Total Duration: 5 seconds = 5000 ms

2. **Assume the recording starts at t = 0. Make a time vector in MATLAB and plot the epileptic EEG data as a function of time and visually inspect it. In one or two sentences briefly describe if you observe any rhythmic activity in this recording.**

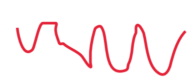
   time = 1:5000;
   plot(time, epileptic)

   There seems to be increasing signal amplitude then a spike then the whole thing repeats again

   Top →

   Bottom

3. **Plot the raw autocorrelation of the EEG signal as a function of the lags. Remember that the lags are in units of the sampling interval so scale your x-axis accordingly. Briefly describe what you see. What is the main period and frequency of the signal, if any?**

   plot(xcorr(epileptic));
   xticks([0:1000:10000]);
   xticklabels([-5000:1000:5000]);
   xlabel("Lag in ms");

   Maximum autocorrelation when lag is 0 as expected and it seems to oscillate between high autocorrelation and low autocorrelation every 8(to 8.5) milliseconds and the absolute value of autocorrelation decreases linearly as lag is increased (or decreased). The period of the signal is about 16(to 16.5 ms) (but the amplitude changes) and frequency of the signal is about 60 Hz

[1] From the labs of Giles van Luijtelaar and Joyce Welting. Phys. Rev. E, 2002; 65: 041903.

4. **Use the FFT in MATLAB to compute the periodogram of the signal (power spectrum). Plot power as a function of frequency on a linear scale for power. Restrict to your x-axis to the range 0-80 Hz.**
5. **What is the Nyquist frequency and the frequency resolution on the periodogram?** 500 Hz 0.2 Hz
6. **Replot the power spectrum on a dB scale for power. Restrict to your x-axis to the range 0-80 Hz. What do you see? Are there any other components in this signal?** There are components of other frequencies which were being lost in the plot when linear scale of power was used
7. **Finally replot the power spectrum on a dB-scale and a logarithmic axis for frequency (x-axis, again in the range 0-80 Hz). What do you see? Are there any other components in this signal?** There is a lot of positive power in a lot of frequency components <60 Hz with the frequency near 10 Hz having the most amplitude as compare to >60 Hz which mostly have negative power

PART 2: Removing electric outlet noise.

In the previous section you identified that the epileptic EEG signal is heavily contaminated with 60 Hz noise. The objective of this section is to apply filtering in the frequency domain.

1. **Your job is to write code to obtain a filtered signal. The steps are:**
   a. **Take the FFT of the signal.**
   b. **Plot the magnitude of the Fourier coefficients for both positive and negative frequencies.**
   c. **Based on this plot, choose a reasonable threshold to annihilate the offending frequency (electrical outlet noise). Use this threshold to set the Fourier coefficients whose magnitude is greater than the threshold to zero.**
   d. **Take the inverse Fast Fourier Transform to send the filtered signal in the frequency domain back to the time domain. This is your filtered data**
2. **Plot the filtered data as a function of time. Visually check that there are no 60 Hz components.**
3. **Use FFT to plot the periodogram of the filtered signal on a linear scale for frequency and power. Restrict to your x-axis to the range 0-80 Hz. Use dots for your plot rather than a continuous line. Again, corroborate that the 60 Hz noise is gone but all other frequencies in the original signal remain present.**
4. **The periodogram is very jagged. Replot the periodogram by applying a zero-pad factor of 5 (i.e. the padded signal to FFT should be 5 times the length of the unpadded signal). Use dots for your plot rather than a continuous line. Compare this plot to the previous one. What tis the effect of applying zero-padding?** More dots, smoother plot, more details preserved

PART 3: Multi-taper spectra and spectrograms on the filtered data

In this section we will compute and visualize spectra and spectrograms using a multi-tapered approach.

1. **Compute the spectrum for the filtered signal using the multi-tapered method. Do so using time-half bandwidth (TW) products of 1.5, 2, and 5. Set the zero-pad factor to 0 and the time window to 1 seconds. Use subplot to visualize the three spectra on a single figure panel. Use a linear scale for power and a linear scale for frequency (restrict the x-axis to the range 0-80 Hz)**

2. **Repeat the same analysis but setting the zero-pad factor to 5. Make a new figure with 3 subplots with same consideration for axes scaling/range.**
3. **In a few sentences interpret the spectra and describe which rhythms are present in this signal. What combination of multi-taper parameters do you prefer for this signal? Defend your choice. How does the multi-taper results compare to the periodogram? Which one is more informative?**
4. **Use surf to plot the spectrogram as a function of frequency and time. Se the power on a linear scale and restrict he y-axis to the range 0-30 Hz. Use a time-bandwidth product of 2, a time window of 1 seconds and a time-step of 0.05 seconds. Apply zero-padding with a factor of 5. At which time do the rhythmic activities turn on and off?**

The main rhythm in decreasing power is around 8 Hz (possibly Theta ?) , around 16 Hz (possible Beta?) and around 25 Hz (again Beta?). I prefer TW = 1.5 and zero_pad = 5. A bigger TW loses resolution and zero padding of 5 is giving more resolution in the power for similar frequencies. Multi Taper is giving a variety of results at different resolution so its better to see the kind of possibilities and to know if the best parameters have been chosen. It offers more resolution. Multi Taper is more informative

activity: 0.5s to 1s , 2.5 s to 3s

PART 4: Comparison to normal data

The normal data is not contaminated with 60 Hz noise. We will carry some analyses to compare the normal and epileptic recordings.

1. **Plot the normal EEG data as a function of time and visually inspect it. In one or two sentences briefly describe if you observe any rhythmic activity in this recording. How does this differ from the filtered epileptic data?**

epileptic had definite spiking after intervals. normal doesn't have such activity just general activity in the PFC without such major patterns

2. **Plot the raw autocorrelation of the EEG signal as a function of the lags. Remember that the lags are in units of the sampling interval so scale your x-axis accordingly. Briefly describe what you see. What is the main period and frequency of the signal, if any?**

The signal has low to none underlying periodic activity. Every 150 ms there is some slight increase in amplitude which may just be normal PFC activity.

3. **Plot the spectrogram with identical parameters and scaling as for the epileptic case. Briefly describe what you see. What are the main differences with epileptic recording?**

No turn on turn off in epileptic period defined above. There is 1 Hz turn on turn off in 0.5s to 1.5s. Brain does not register additional activity besides this. The epileptic may be the seizure happening so there is spiking but this is just a guess.