Spectral analysis: laboratory

Your report must be uploaded on `hippocampus` *server at most* *10 days after the last session. You can work by pairs and send only one report. It should include your answers, appropriate figures, comments on obtained results and Matlab codes* *as a unique PDF file.*

Think in giving titles to your figures: `title('My figure')` **and correctly label axes:** `xlabel('Time')` / `xlabel('Frequency'), ...`

Recall that, for an N-point signal sampled at frequency F_S, its Fourier transform is computed by `Fx = fft(x,Nf);` where `Nf` should be an integer power of 2 (see Matlab function `nextpow2`), significantly bigger than N. Then, the corresponding frequency axis corresponds to `Nf` regularly spaced frequencies, from 0 to $\frac{Nf-1}{Nf}Fs$. In Matlab: `freq = 0:Fs/Nf:(Nf-1)/Nf*Fs`. The periodogram is then displayed by `plot(freq,abs(Fx).^2);`...

1 Detection of oscillations with periodograms

The data file `signals.mat` available on the web server contains fours different signals `x1`, `x2`, `x3` and `x4`, composed of one or several sine waves, some of them with noise. Here, we consider normalized frequencies, that is, the sampling frequency is $Fs = 1$.

By using appropriate periodogram analysis (basic periodogram, windowed periodogram or Welch periodogram with appropriate settings), what can you conclude about the frequencies present in each data set?

For each case, draw the time-domain signal and its spectral analysis, with what you consider to be the most relevant choice of periodogram(s), and indicate how you chose the parameters.

Tips:

- Basic and windowed periodograms will be computed using the `fft` function. Most windowing functions (Hamming, Hanning, Blackman) are already implemented in Matlab. For example, `h=hamming(N);` generates a Hamming window with N points.
- The Welch periodogram can be computed with the Matlab function:

  ```
  periodogram_welch = pwelch(x,window,noverlap,Nf,Fs,'twosided');
  ```

  where `window` is the size of each segment, `noverlap` is the number of overlapping points between two segments, `Nf` is the number of frequencies and `Fs` is the sampling frequency (here, `Fs` $= 1$).

2 High-resolution methods

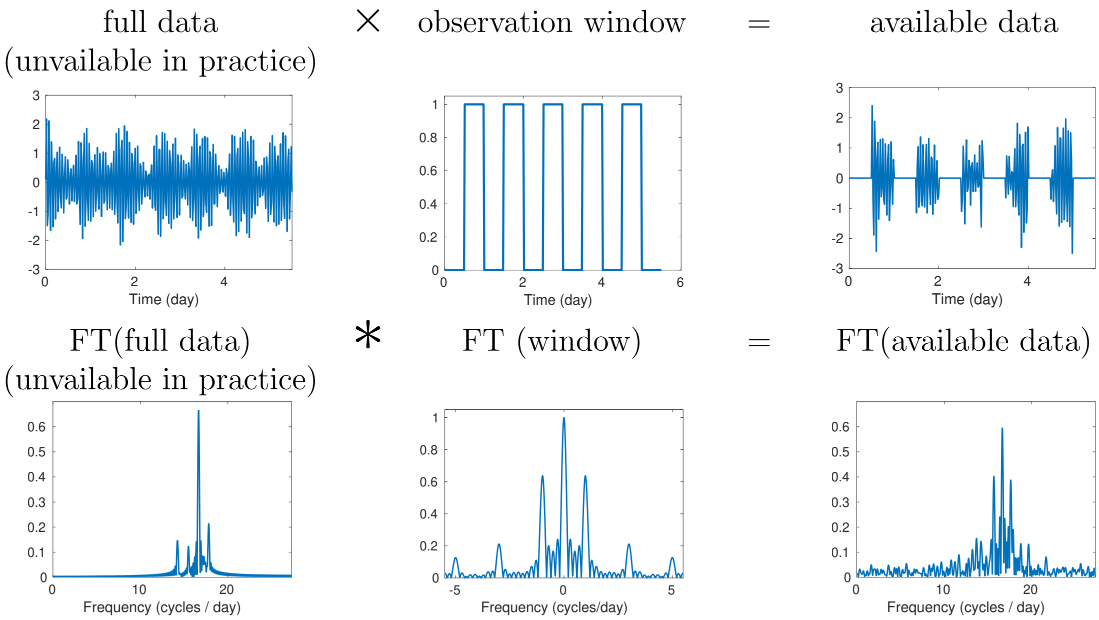
Compare the obtained results to those obtained by the three following high-resolution methods:

1. The *autoregressive* (AR) spectral analysis method can be computed as follows:
   (a) First, define the order P of the model (tip: one complex exponential corresponds to one pole in the transfer function...)
   (b) Then, estimate the AR coefficients. Among the several possible methods, you may use the covariance one: `[a,sigma2] = arcov(x,P);`
   (c) Finally, the spectrum is computed by `ARspectrum = sigma2./abs((fft(a,Nf))).^2;`. Explain why.
2. The *MUSIC* method will be directly implemented from the Matlab function:
   `S_MUSIC = pmusic(x,P,freq,Fs);`
   where `P` is the number of *complex* exponentials that are searched and `freq` is the corresponding frequency axis on which the MUSIC indicator value is computed.
3. The *sparsisty-regularized Fourier spectrum* method will be implemented from a specific function, avilable on hippocampus:
   `u = minl1_Fourier(x,freq,mu);`
   computes the minimizer of the least-squares problem penalized with an ℓ_1-norm term (see course slide 44). `freq` is the corresponding frequency axis and μ is the weight of the ℓ_1-norm term, to be tuned for each case.

3 Exoplanet detection from the spectral analysis of time series

In Astrophysics, a classical method for extrasolar planet detection consists in **searching periodic components** in an observable quantity such as the apparent velocity of the star (that can be obtained through spectroscopic measurements and preprocessing steps). When orbiting a star, each planet creates slight periodic variations of the mass center of the system, which can be approximately modeled as a sinusoidal component under some simplifying assumptions (circular orbit, line of sight in the same plane as the star-planet system, ...)

Ground-based observations suffer from **irregularity**: the star is only visible at night, therefore long-time observations contain **gaps**. Equivalently, such **missing data** can be described by means of an *observation window*, equal to 1 when data are available (night) and 0 otherwise (day). Subsequently, the Fourier transform of the available data can be written as the *convolution* of the true spectrum by the Fourier transform of such observation window:



The file `stardata.mat` contains typical data (variables `t` and `x`), corresponding to five observation nights.

1. Variable `win` contains the observation window, and `t` contains the corresponding time axis. Compute and graph its Fourier transform in magnitude (use 4096 frequencies between 0 and F_s). Comment its shape.

2. Compute and graph the periodogram of the data (variable `x`), and try to interpret it.

3. *In order to reduce artifacts, a classical method*[1] *consists in the following iterative procedure: the main frequency is estimated by the Fourier transform, then its contribution to the data is removed. The residual signal is then processed again and a new frequency is searched in its Fourier transform, until no significant component can be found in the residual.*

   Implement this method where, at each iteration:

   i) the main frequency f_i is estimated at the maximum of the periodogram of the residual signal;

   ii) the corresponding cosine and sine components are obtained by least-squares estimation. This is a trivial operation, which is performed by the given function:

   ```
   [alpha,beta] = estim_cos_sin(residual,t,fi);
   ```

   The contribution of the corresponding sine wave is then $\alpha\cos(2\pi \mathtt{f_i}\, t) + \beta\sin(2\pi \mathtt{f_i}\, t)$.

   iii) The analysis continues or terminates, depending on whether relevant features are found in the periodogram.

4. Implement this algorithm and display, at each iteration, the residual data as a function of time, its Fourier transform and the detected components.

5. Comment on the results. How many frequencies (and therefore, exoplanets) can you detect?

[1] see D.H. Roberts *et al.*, *Time Series Analysis with Clean - Part I - Derivation of a Spectrum.* Astronomical Journal 93, 1987.