Signals and systems class, HSE, Spring 2015, A. Ossadtchi, Ph.D.

Lecture 7

Frequency domain portraits of random processes

Introduction

We have explored in quite a bit of details the intuition behind transforms. We started from a very simple case of discrete finite length sequences and illustrated the main principles that remain true for any orthogonal transform. Then, we considered the Discrete Fourier Transform, a technique very commonly used in practice to elucidate hidden periodic structures present in the data. We then considered Discrete Time Fourier Transform, a technique that we obtain by replacing the discreet index reflecting oscillation frequency with a continuous variable ω. This allowed to increase the frequency domain resolution of our analysis and explain the values of the DFT coefficients we get when working with sequences of finite length. We have then looked at how to describe LTI system in the frequency domain and studied the concept of system's transfer function $H(\omega)$ describing system's reaction to harmonic excitations of each particular frequency. Having in hand the DTFT based decomposition of the input signal as a superposition of harmonic sequences and using system's transfer function we formulated a simple way to compute the DTFT coefficients $Y(\omega)$ of LTI system output signal $y[n]$ in response to the input signal $x[n]$ with DTFT coefficients $X(\omega)$. This is $Y(\omega) = H(\omega)X(\omega)$, since all components of this equation are in general complex we can split the amplitude and phase components as

$$|Y(\omega)|e^{j\varphi_y} = |H(\omega)|e^{j\varphi_H}|X(\omega)|e^{j\varphi_X} = |H(\omega)||X(\omega)|e^{j(\varphi_X+\varphi_H)}$$

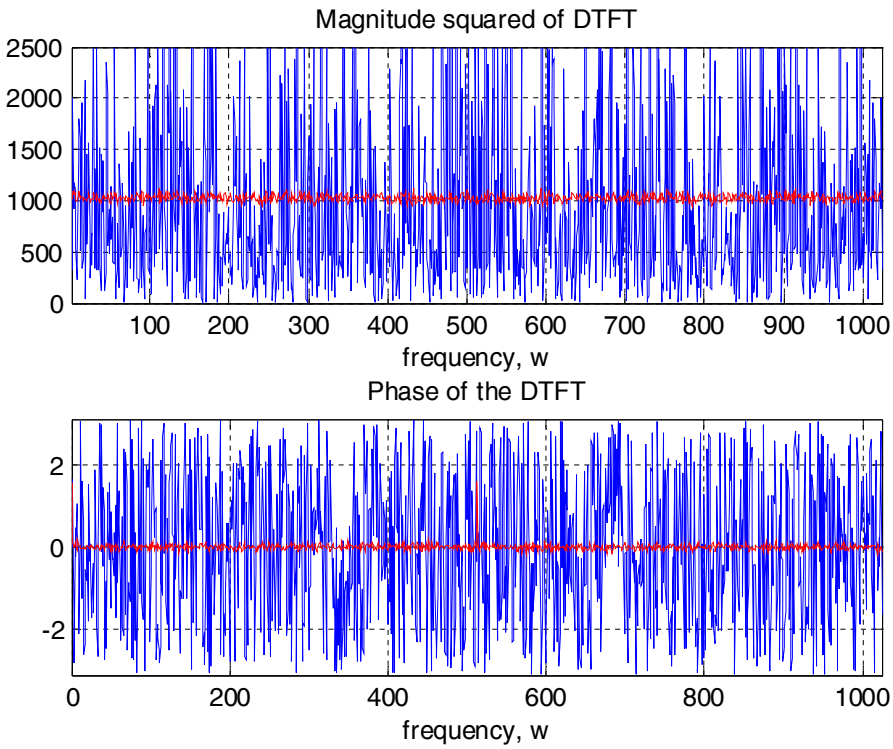
Clearly, the analysis of LTI systems in the Fourier domain boils down to a mere multiplication of amplitude components and summation of the phase components describing the output signal and the system.

So far, in the context of Fourier domain analysis we have restricted ourselves to the deterministic signals. However, in real life we are usually dealing with realizations of random signals. I think we are now well equipped to explore the ways we can represent signals in the Fourier domain. Besides a very useful tool, such representation furnishes a fruitful framework for your thinking about data analysis problems you may encounter in your daily life as a researcher.

An experiment with a random process

One of the most frequently used random process is so called Gaussian white noise. The defining property of this process is the absence of statistical dependence of its samples. You have already played with this process during your first homework. From the frequency domain point of view this process contains contributions of equal extent from the harmonics of all the frequencies. Let's try to see it. The naïve approach would be to apply the DTFT to a realization of the process and to look at the magnitude and phase spectra of the result.

Let's do it. Consider the figure below



The blue trace represents the magnitude squared and the phase of the DTFT coefficients computed on a single realization of the white noise random process with N = 1024. The squared magnitude of the DTFT coefficients normalized by the length of the sample is called periodogram. This is really very different from what we expected to see that would show equal contribution of power on all the frequencies. More interesting is if we increase the length of the data the power of deviations around the true values of "something"(see the next section) showing us the distribution of random process power over frequencies will not decrease. Let's try to generate 1000 realizations of this random process, calculate magnitude and phase and average the magnitude and phase. The results of this procedure is shown on the same axis in red. We can see that average of the magnitude converges to some non-zero value that remains constant over all frequencies. This is exactly what the white noise process is about! Pay attention to the phase, for all (except $\omega = 0$ and its mirror) values of frequency the averaged phase approaches zero values and therefore is the phase is not informative as a descriptor of this random process. In general, for Fourier domain description of the wide sense stationary random processes the phase information is not used. For some specific random processes, however, the phase may carry the important information (e.g. phase-locked Evoked Response Potentials).

Periodogram vs Power Spectral Density

So, in the previous section we first calculated the periodogram

$$\tilde{P}(\omega) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x[n]e^{-j\omega n}\right|^2$$

and realized that it gives a very noisy version of what we are expecting to see. Another experiment (not shown here) would demonstrate that the increase of N does not reduce the noise. As another experiment shown, averaging $\tilde{P}(\omega)$ calculated over several realizations allows to reduce the noise. In fact the

variance of the estimate decreases as $1/N_{avg}$. With the infinitely large number of averages we calculate the mean value (or expected value) of $\tilde{P}(\omega)$. We will call this quantity Power Spectral Density as instead of energy concentration function $P(\omega)$ defined during the previous lecture we are using the quantity normalized by the length of the sample and therefore proportional to the energy per time unit which is power.

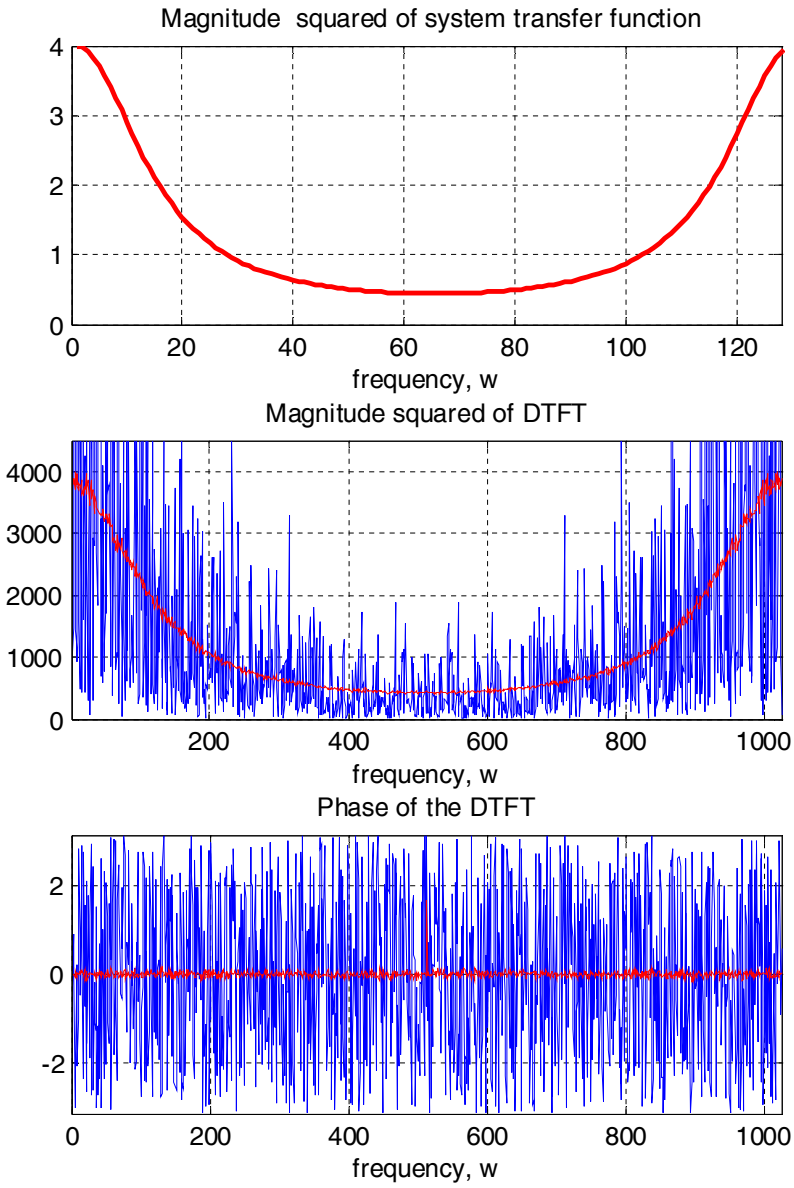
Passing a random process through an LTI system (Shaping filter)

Many processes can be obtained from the white noise process by means of so called shaping filters. Shaping filter is basically an LTI system with some predefined pulse response. The pulse response is chosen to ensure that the response of the LTI system to the white noise input will have desired PSD. Since the PSD of random process *y* as defined above is simply $S_{yy}(\omega) = \lim_{N\to\infty} \frac{1}{N} E\{Y_N(\omega)Y_N^*(\omega)\}$ we can see what will be the PSD of the output of the LTI system. We know that the DTFT of the signal output by an LTI system with transfer function $H(\omega)$ is $Y_N(\omega) = H(\omega)X_N(\omega)$, then the PSD of this output signal is simply $S_{yy}(\omega) = \lim_{N\to\infty} \frac{1}{N} E\{Y_N(\omega)Y_N^*(\omega)\} = \lim_{N\to\infty} \frac{1}{N} E\{H(\omega)X_N(\omega)H^*(\omega)X_N^*(\omega)\} =$
$= H(\omega)H^*(\omega) \lim_{N\to\infty} \frac{1}{N} E\{X_N(\omega)X_N^*(\omega)\} = |H(\omega)|^2 S_{xx}(\omega)$.
This relation stating that $S_{yy}(\omega) = |H(\omega)|^2 S_{xx}(\omega)$ allows to find the transfer function that would guarantee the specific shape of the PSD of the output signal. This becomes especially simple if the input signal is the white noise process with all frequency components present to equal extent with PSD $S_{NN}(\omega) = \sigma^2 1(\omega)$. Then, the PSD of the output signal will be simply equal to the squared norm of the system's transfer function.

Let's do the experiment. Remember during the last lecture I showed to you that the DTFT of the sequence $h[n] = 0.5^n u[n]$ that happened to be $H(\omega) = \frac{1}{1-0.5e^{-j\omega}}$. If we take sequence $h[n] = 0.5^n u[n]$ to be the pulse of our system then in order to implement it (at the current level of your knowledge) we would need to perform a convolution with an infinitely long sequence. To avoid this, let's truncate this pulse response at *N*=128 points. The squared magnitude of the DTFT of this truncated pulse response is shown in the top panel of the figure below.

Now, let's generate a sample $x[n]$ of the white noise random process and pass it through the system. To do so, we convolve this sample with $h[n]$, i.e. $y[n] = h[n] * x[n]$.

The periodogram of a single realization and the estimate of the PSD



Wiener-Khintchine theorem

We have stated that the best way to describe a random process is to define joint the probability density function of its samples. In practice, however, dealing with this way of defining a random process is not practical. Let us limit ourselves to considering only stationary random processes and recall the definition of the autocorrelation function $(n_1, n_2) = \iint_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2, n_1, n_2) dx_1 dx_2$. For a stationary process this function depends only on the distance $m = n_2 - n_1$ and therefore we get $(m) = \iint_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2, m) dx_1 dx_2$.

The autocorrelation function is the second order statistics of a random process. Autocorrelation function of a zero-mean Gaussian random process completely describes this process. We will come

back to this fact later in the lecture when we will consider state space representation of timeseries.

Let's do an exercise. In the first section we have introduced the notion of periodogram as the normalized square magnitude of the DTFT coefficients computed on a sample of a random process $\tilde{P}(\omega) = \frac{1}{N}|X(\omega)|^2 = \frac{1}{N}X(\omega)X^*(\omega)$. Let's now assume that we received a realization $x[n]$ of a random process *x* and calculated its periodogram as

$$\tilde{P}(\omega) = \frac{1}{N}\left(\sum_{n=0}^{N-1} x[n]e^{-j\omega n}\right)\left(\sum_{m=0}^{N-1} x[m]e^{-j\omega m}\right)^* = \frac{1}{N}\left(\sum_{n=0}^{N-1} x[n]e^{-j\omega n}\right)\left(\sum_{m=0}^{N-1} x^*[m]e^{j\omega m}\right) =$$
$$= \frac{1}{N}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} x[n]x[m]^*e^{-j\omega(n-m)}$$

Note, that in order to carry out the double summation we need a new index of time *m*. If we kept the same index serving both summations we would have summed only the diagonal terms and completely omitted the cross-terms.
We are interested in some sort of statistics that would reflect statistical properties of our random process. Since *x*[*n*] is basically a realization of random process we can take the expectation operation and obtain

$$E\{\tilde{P}(\omega)\} = \frac{1}{N}E\left\{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} x[n]x[m]^*e^{-j\omega(n-m)}\right\} = \frac{1}{N}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} E\{x[n]x[m]^*\}e^{-j\omega(n-m)}$$

We can recognize that the term $E\{x[n]x[m]\}$ is autocorrelation function $R(n,m)$ of the random process. In case the random process is stationary $R(n,m) = R(n-m) = R(k)$ and then we can rewrite the above expectation as

$$E\{P(\omega)\} = \frac{1}{N}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} R(n-m)e^{-j\omega(n-m)}$$

Now things will get a bit more complicated but only technically. We need to convert this double summation into a single summation of the autocorrelation sequence lags weighted by the number of times each of the lags appears in this double sum. Consider this table of *n-m* values for *N* = 5

n, m	**1**	**2**	**3**	**4**	**5**
1	0	1	2	3	4
2	-1	0	1	2	3
3	-1	-1	0	1	2
4	-3	-2	-1	0	1
5	-4	-3	-2	-1	0

Therefore, we can conclude that we will get the following count of lags in this double summation

n-m	0	1	2	3	4	-1	-2	-3	-4
Count	5	4	3	2	1	4	3	2	1

Therefore, the summation for $\sum_{n=0}^{4}\sum_{m=0}^{4} R(n-m)e^{-j\omega(n-m)}$ can be written as $5R(0)e^{j\omega 0} + 4R(1)e^{-j\omega 1} + 3R(2)e^{-j\omega 2} + 2R(3)e^{-j\omega 3} + 1R(4)e^{-j\omega 4} + 4R(-1)e^{j\omega 1} + 3R(-2)e^{j\omega 2} + 2R(-3)e^{j\omega 3} + 1R(-4)e^{j\omega 4}$

Generalizing we get

$$E\{\tilde{P}(\omega)\} = \frac{1}{N}\sum_{k=-N+1}^{N-1}(N-|k|)R(k)e^{-j\omega k} = \sum_{k=-N+1}^{N-1}\left(1-\frac{|k|}{N}\right)R(k)e^{-j\omega k}$$

Let's consider the limit of the above as *N* goes to infinity:

$$S_{XX}(\omega) = \lim_{N\to\infty} E\{P(\omega)\} = \lim_{N\to\infty}\sum_{k=-N+1}^{N-1}\left(1-\frac{|k|}{N}\right)R(k)e^{-j\omega k}$$

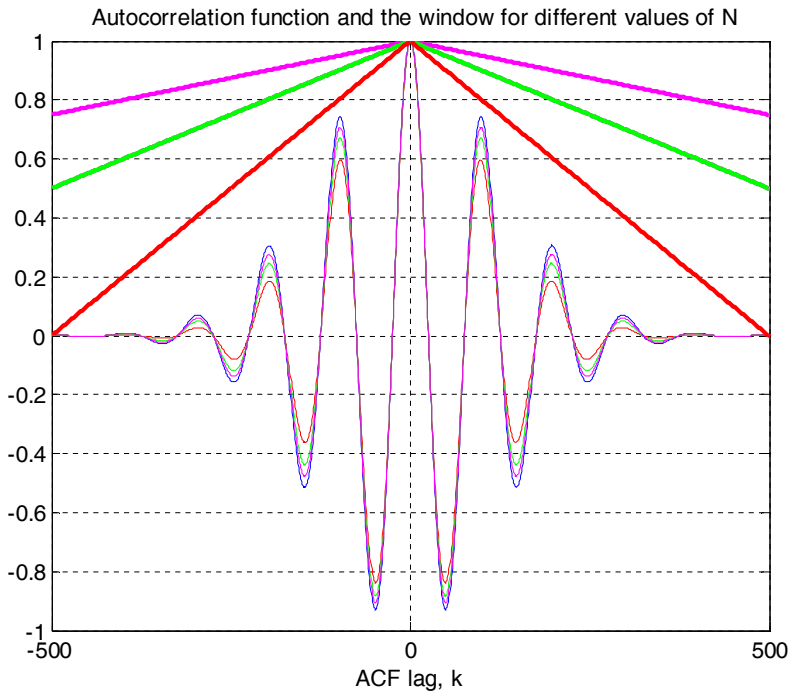


Figure 1: *Autocorrelation function (blue, thin) , three different windows corresponding to N = 500,1000,2000 (thick lines) and the windowed ACF for these three different windows (colored curves). As we can see with N growing large the windowed ACF (red->green->magenta) approaches the original ACF (blue thin line)*

The theorem we are trying to derive here is valid only for the processes with finite energy, in other words for the processes whose ACF is square summable. In case of practical usage, the finite length of sequences guarantees this. The autocorrelation function at lag *k* shows the amount of linear statistical dependence of process samples situated *k* time points apart. Clearly, in the real-life situations this dependence diminishes with increasing *m* and therefore the magnitude of the ACF decreases as *k* grows. At the same time the slope of the triangular window function $1 - \frac{|k|}{N}$ decreases with growing *N* and thus its product with the original ACF tends to this original ACF as $N \to \infty$. These two facts together make the limit in the above expression evaluates as

$$S_{XX}(\omega) = \lim_{N\to\infty} \sum_{k=-N+1}^{N-1} \left(1 - \frac{|k|}{N}\right) R(k) e^{-j\omega k} = \sum_{k=-N+1}^{N-1} R(k) e^{-j\omega k}$$

This expression is called Wiener-Khintchine theorem and states that the power spectral density $S_{XX}(\omega)$ of a random process is related to its autocorrelation sequence via Fourier transform, the DTFT in case of discrete sequences.