

DSP Basics

Let's recapitulate:

- Spectrum, DFT, DTFT and FFT
- the connection between periodical signal and harmonic spectrum
- time-frequency resolution

1 Spectrum, DFT, DTFT and time-frequency resolution

Let's catch the signal

$$s(n) = [1 \quad 1 \quad 2 \quad -2 \quad -1]$$

and it's DFT

$$S(k) = \sum_{n=0}^4 s[n] \exp(-jnk * 2\pi/5),$$

which can be found in figure 1. For instance values $S(5)$ and $S(17)$ can be obtained by the same formula. As a result we get that DFT is periodical, in this case the length of the period is 5 which implies $S(0) = S(5) = S(10)$, $S(1) = S(6) = S(11)$, and so forth.

DFT $S(k)$ contains the same information as the original signal $s(n)$, even without redundancy, which is good for instance in signal compression. However in signal analysis it is often useful to have more redundant frequency representation. Let's examine the value

$$S[2] = \sum_{n=0}^4 s[n] \exp(-jn2 * 2\pi/5),$$

in DFT, it is the innerproduct of signals $s(n)$ and $\exp(-jn2 * 2\pi/5)$, in other words it tells roughly, how much signal $\exp(-jn2 * 2\pi/5)$ is contained in $s(n)$ (enthusiastic mathematicians can ponder this more carefully by considering the innerproduct of vectors in \mathbf{C}^n). The signal

$$\exp(-jn2 * 2\pi/5)$$

is simply complex exponent signal, with period of 5 samples, which therefore can be written as well in form

$$\exp(-jn\omega),$$

where $\omega = 2 * 2\pi/5$. Having get this far, we can set ω other real values such as 0, $2\pi/5$, $2 * 2\pi/5$, $3 * 2\pi/5$, $4 * 2\pi/5$. For example if $\omega = 0.1 * 2\pi$, then the sum

$$\sum_{n=0}^4 s[n] \exp(-jn\omega)$$

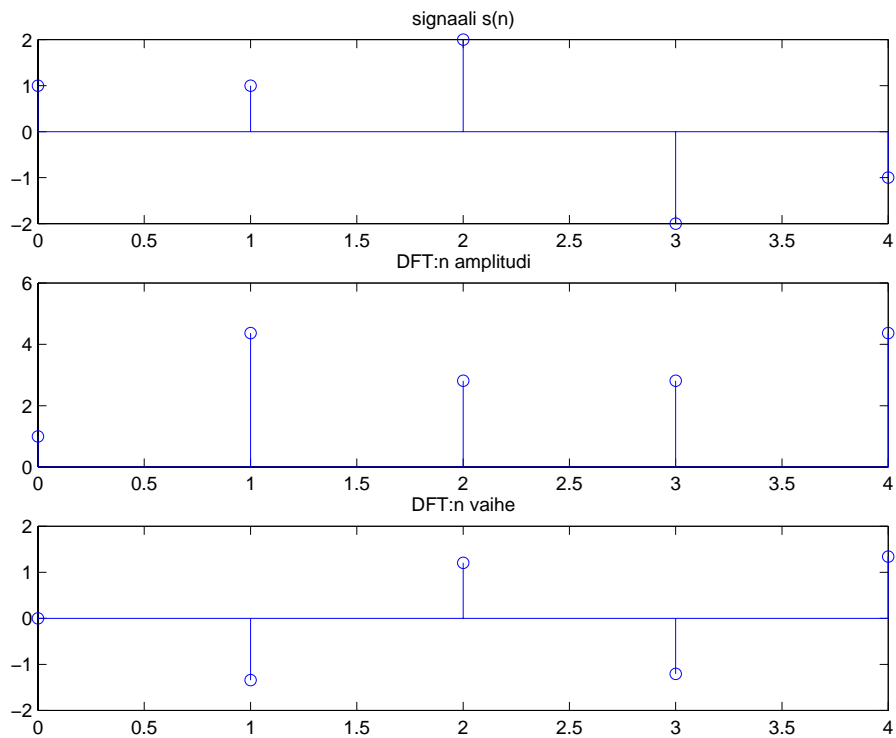


Figure 1: Waveform and its DFT (amplitude and phase separately).

tells roughly how much the signal

$$\exp(-jn\omega)$$

(compl. exp. signal, period 10) is contained in $s(n)$.

By doing the same for a big bunch of different values of ω we get a function

$$S(\omega) = \sum_{n=0}^4 s[n] \exp(-jn\omega),$$

which is DTFT (discrete time Fourier transform) shown in figure 2. We observe that also DTFT is periodic with period 2π (devious thing since it is not dependent on the length of the signal), moreover its values at $0, 2\pi/5, 2 * 2\pi/5, 3 * 2\pi/5, 4 * 2\pi/5$ give precisely the DFT.

The squared amplitude of DTFT $|S(\omega)|^2$ is called spectrum (in fact it can be defined in many ways but this suffices for us). What's the use of this spectrum then? E.g. the following: take 10 samples of sinusoidal signal

$$s(n) = \sin(n * 2\pi/3),$$

with period 3. The signal and the amplitude of the DFT are shown in figure 3. Since $s(n)$ is fully periodic signal, we could expect that there is only one component (in addition one component corresponding to the negative frequency but this is not essential consideration here), but what is shown to us, is a big bunch of different frequencies in DFT.

The explanation is that there is no frequency $2\pi/3$ in DFT, of length 10. The closest frequencies are $3 * 2\pi/10$ and $4 * 2\pi/10$.

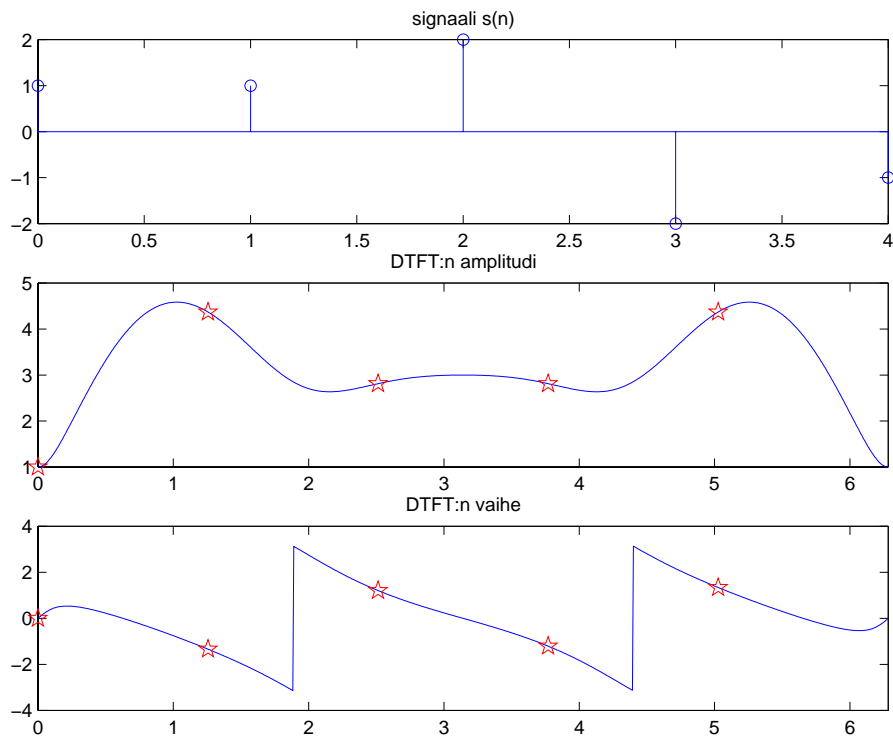


Figure 2: *DTFT*.

If the length of the signal happens to be a multiple of its period, then there is only one nonzero value in the DFT.

However the spectrum of the signal is shown in figure 4. It shows that the peaks have spread to entire frequency band in such a way that the highest peak is in the frequency of the signal.

Interpolated DFT gives more 'fair' picture of the signal, because its values do not depend so much on how the length of the signal is related to its period.

By considering frequency estimation of a sinusoidal signal we arrive at more general concept called Heisenberg's uncertainty principle of time-frequency-resolution: If time resolution of the signal is good, then the frequency resolution is bad and vice versa.

Time resolution is roughly the window length of the signal and frequency resolution tells how accurate is the interpolated DFT.

If we take a window of 10 samples of the signal, we know pretty well the position of this window in time. On the other hand it is hard to make any precise frequency analysis based on this short window: it could be sinusoid of some frequency and little bit noise or some other sinusoid with entirely different frequency and more noise.

If we take a window of 1024 samples, we can tell quite well which one of the frequencies considered above shows in the signal. But now the moment of time is much less accurately spreaded over 1024 samples.

This principle can be formulated a little bit more accurately but for this course it is wise to keep things simple and in mind that longer window implies better frequency resolution (but worse time resolution).

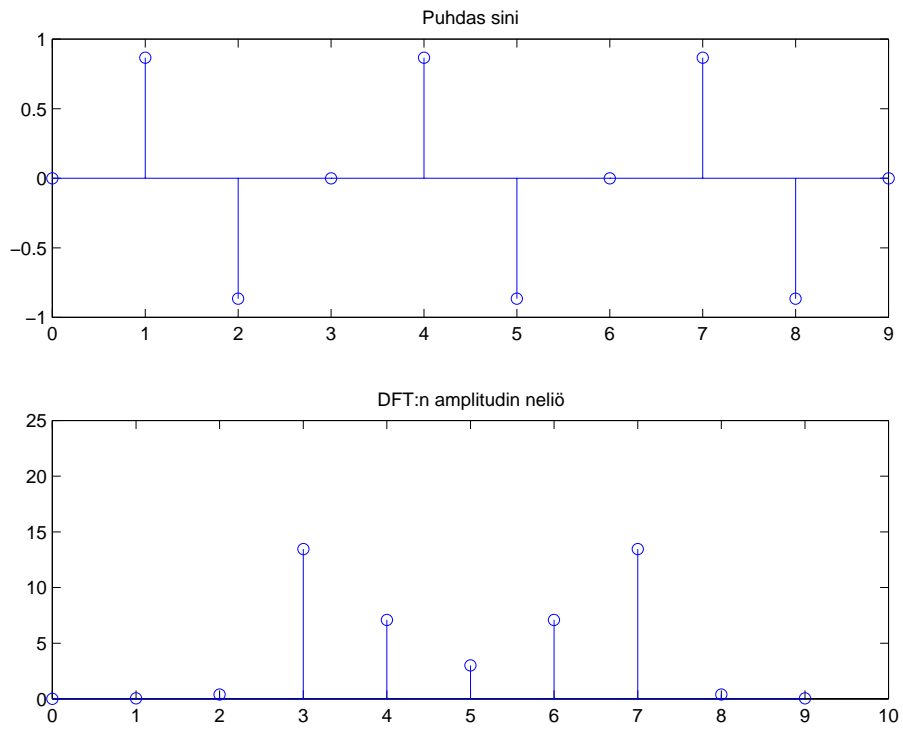


Figure 3: 10 samples of sinusoidal signal having period 3 and the squared amplitude of DFT.

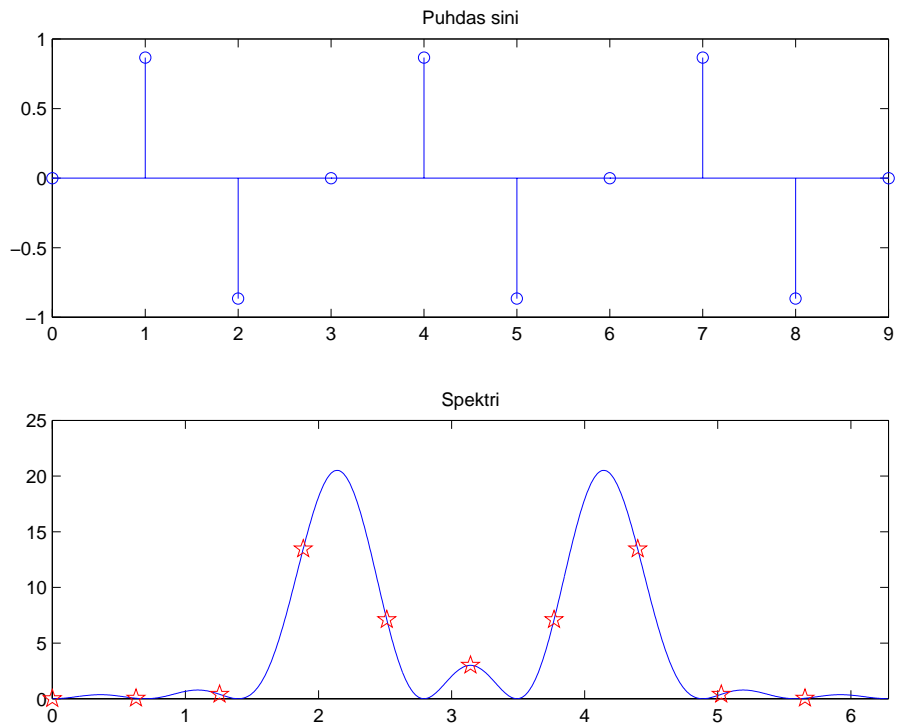


Figure 4: 10 samples of sinusoidal signal with period 3 and its spectrum.

2 Periodic signal → harmonic spectrum

Zero padded signal → spectrum

The waveform of voiced speech (e.g. vowels) is often nearly periodic. The periodicity in turn is shown in the Fourier-transform s.t. its DFT is harmonic, which is most of its energy is on the fundamental frequency f_0 and its multiples $2f_0, 3f_0, 4f_0, \dots$. But why?

Explanation 1 (hard). Let's calculate cold bloodedly

$$S_1(\omega) = \sum_{n=0}^{N-1} s_1(n) \exp(-jn\omega),$$

where $s_1(n)$ is non-periodic signal of length N . Let

$$s_2(n) = [s_1(n) \ s_1(n)]$$

be two periods of the signal $s_1(n)$, then

$$\begin{aligned} S_2(\omega) &= \sum_{n=0}^{2N-1} s_2(n) \exp(-jn\omega) \\ &= \sum_{n=0}^{N-1} s_1(n) \exp(-jn\omega) + \sum_{n=0}^{N-1} s_1(n) \exp(-j(n+N)\omega) \\ &= S_1(\omega) + \exp(-jN\omega) S_1(\omega) \\ &= S_1(\omega)(1 + \exp(-jN\omega)). \end{aligned}$$

And in general we can see that

$$S_K(\omega) = S_1(\omega) [1 + \exp(-jN\omega) + \dots + \exp(-j(K-1)N\omega)].$$

Now the question is how

$$E_K(\omega) = 1 + \exp(-jN\omega) + \dots + \exp(-j(K-1)N\omega)$$

behaves as K increases. Since $E_K(\omega)$ is a geometric serie, we obtain (by remembering high school math)

$$E_K(\omega) = \frac{1 - \exp(-jNK\omega)}{1 - \exp(-jN\omega)}.$$

The denominator of this function is 0 when $\omega = 0, 2\pi/N, 2 * 2\pi/N, \dots, (N-1) * 2\pi/N$ when also the numerator is zero. To calculate the value of this we need l'Hopitals rule

$$\begin{aligned} \lim_{\omega \rightarrow \phi} E_K(\omega) &= \lim_{\omega \rightarrow \phi} \frac{1 - \exp(-jNK\omega)}{1 - \exp(-jN\omega)} \\ &= \frac{\left(\frac{d}{d\omega}\right) (1 - \exp(-jNK\omega))|_{\omega=\phi}}{\left(\frac{d}{d\omega}\right) (1 - \exp(-jN\omega))|_{\omega=\phi}} \\ &= \frac{-\exp(-jNK\phi)(-jNK)}{-\exp(-jN\phi)(-jN)} \\ &= \frac{jNK}{jN} = K. \end{aligned}$$

In addition $E_K(\omega) = 0$ when ω is multiple of $2\pi/(NK)$ if it is not the multiple of $2\pi/N$.

So: by concatenating the signal by K copies, we get that the DTFT is $K S_1(\omega)$ when ω is multiple of $2\pi/N$ and it is 0 when ω is multiple of $\frac{2\pi}{NK}$ (except in multiples of $2\pi/N$). Figure 5 aims to illustrate this.

Perhaps this explanation is not the most illustrative one. So let's see yet another one...

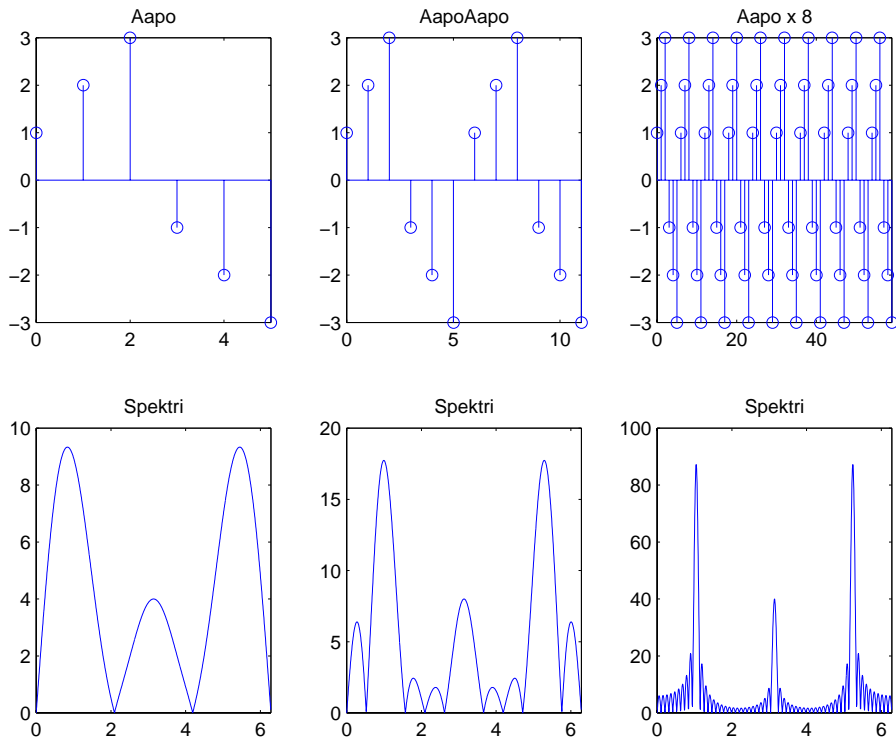


Figure 5: Catenated signals and their spectrums.

Explanation 2 (easier, less accurate). Let's take again K copies of the signal $s_1(n)$ of length N in $s_K(n)$. Let's calculate its DFT $S_K(k)$ (not interpolated), whereupon the inverse DFT, which is the signal $s_K(n)$, is given by summing the signals

$$S_K(k) \exp\left(jnk \frac{2\pi}{NK}\right)$$

over all indices $k = 0, 1, \dots, NK - 1$.

Now the exponent functions corresponding to indices $0, K, 2K, \dots, (N - 1)K$ are periodic with period N .

Is it so? Yes because the period of the signal

$$\exp\left(jnk \frac{2\pi}{NK}\right)$$

is NK/k , and N is the multiple of this when k is multiple of K (to be honest it can be so in other cases too but let's not mourn for it in this inaccurate explanation).

Thus: we have seen that the values of DFT of $s_K(n)$ at $0, K, \dots, (N - 1)K$ correspond to signals with periods whose multiple N is.

This is good, because we know that the signal $s_K(n)$ itself is periodic with period N , so any values in $S_K(k)$ with $k = 0, K, 2K, (N - 1)K$ give an inverse signal (after IDFT) with period N (can be $N/2$ or $N/3$ and so forth but at most N).

Vice versa: other frequencies that every K th, must be zero to ensure the time domain signal to be of period N .

Well one can't say this explanation was particularly clear.

Explanation 3 (ability to explain is almost unactual, but goes as a rule of thumb).

When time domain signal is copied K times, in its spectrum will appear K zeros in between of every original frequencies.

Thus copying the signal in time-domain brings zeros in spectrum. Another operation which is perhaps more useful is padding zeros in the end of the signal before calculating the DFT.

Suppose we have a signal with 256 samples $s_1(n)$ in which we pad zeros so that the length of the signal is 1024. Let's denote this by $s_2(n)$. ($s_1(n)$ is the beginning of $s_2(n)$ and the rest is zeros.)

The figure 6 shows the DFT of the original signal and the DFT of the zero-padded signal. We notice that DFT will be interpolated. Why?

This can be explained by calculating the spectrum. That is to say the samples of spectrum $S_1(\omega)$ at frequencies $\omega = 0, 2\pi/256, 2 * 2\pi/256, \dots, 255 * 2\pi/256$ are given by

$$S_1(\omega) = \sum_{n=0}^{255} s_1[n] \exp(-jn\omega).$$

On the other hand in the spectrum $S_2(\omega)$ can be found frequencies $\omega = 0, 2\pi/1024, \dots, 1023 * 2\pi/1024$

$$S_2(\omega) = \sum_{n=0}^{1023} s_2[n] \exp(-jn\omega).$$

But wait a minute! Because the values of first 256 samples in the $s_2(n)$ are same as in the $s_1(n)$ and the rest of samples are zeros, we obtain

$$S_2(\omega) = \sum_{n=0}^{255} s_1[n] \exp(-jn\omega),$$

which is the same as $S_1(\omega)$.

Thus the spectrum of the zero-padded signal is precisely same as the original, except the frequency sampling is more dense.

Good thing about calculation of DFT of a zero-padded signal is its low computational complexity due to implementation with FFT. (DFT is the transform which is calculated by the FFT algorithm).

In Matlab the DFT of the zero-padded signal is calculated with command `fft(x, n)` where n is the desired length.

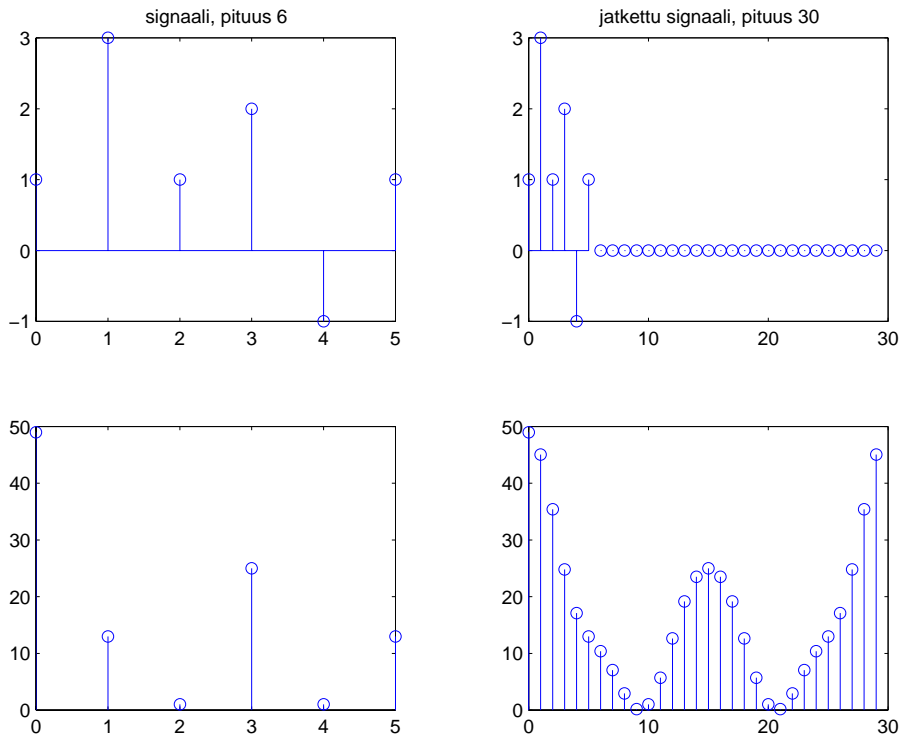


Figure 6: Signal, zero-padded signal and DFTs of both.