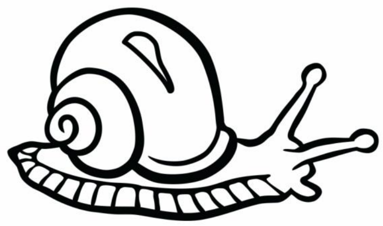
Обработка и интерпретация речевого сигнала. Обработка сигнала в частотной области

П. А. Холявин

p.kholyavin@spbu.ru

19.09.2024



Дискретное преобразование Фурье

$$X_N[k] = \sum_{n=0}^{N-1} x_N[n] e^{-j2\pi nk/N} \qquad 0 \le k < N$$

$$x_N[n] = \frac{1}{N} \sum_{k=0}^{N-1} X_N[k] e^{j2\pi nk/N} \qquad 0 \le n < N$$

Дискретное преобразование Фурье

$$\tilde{x}_N[n] = \frac{1}{N}\sum_{k=-18}^{18} X_N[k]e^{j2\pi nk/N} = \frac{X_N[0]}{N} + \frac{2}{N}\sum_{k=1}^{18} X_N[k]\cos(2\pi nk/N)$$

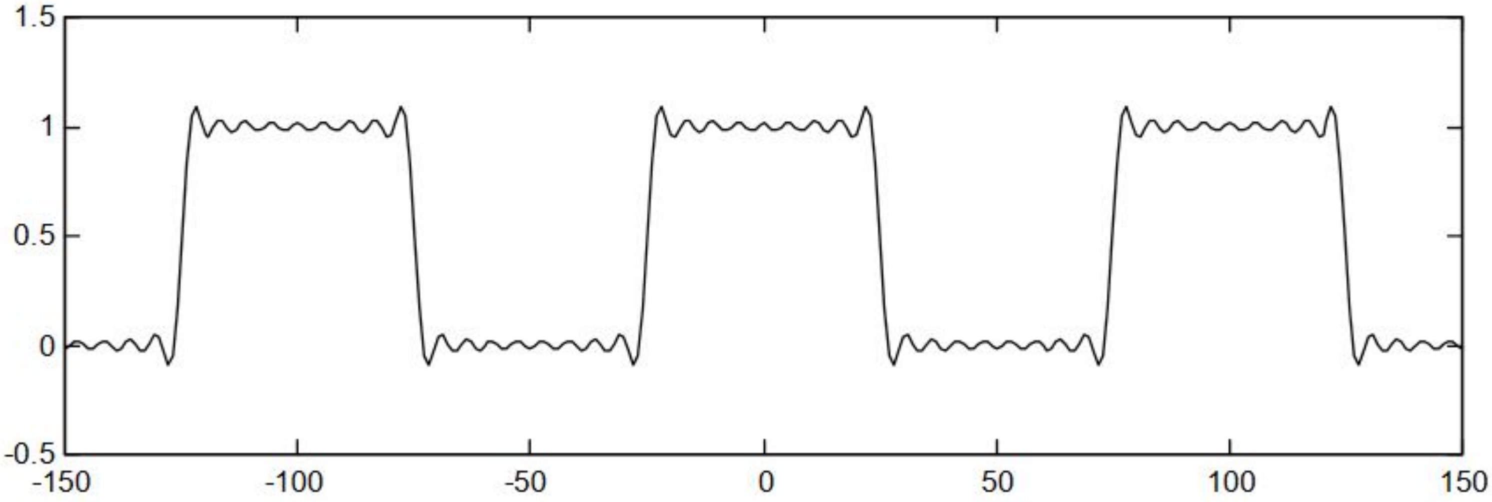


Figure 5.11 Decomposition of a periodic square signal with period 100 samples as a sum of 19 harmonic sinusoids with frequencies $\omega_k = 2\pi k/100$.

Дискретное преобразование Фурье

Если дан фрейм сигнала в N отсчётов с частотой дискретизации F, то как определить, какой частоте соответствует n-ный отсчёт спектра?

Быстрое преобразование Фурье

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi nk/N} = \sum_{n=0}^{N-1} x[n] W_N^{nk} \qquad 0 \le k < N \qquad W_N = e^{-j2\pi/N}$$

$$X[k] = \sum_{n=0}^{N/2-1} f[n] W_{N/2}^{nk} + W_N^k \sum_{n=0}^{N/2-1} g[n] W_{N/2}^{nk} = F[k] + W_N^k G[k]$$

$$F[k + N/2] = F[k]$$

$$G[k + N/2] = G[k]$$

Спектр и спектральная плотность

The *complex spectrum* of a sound $x(t)$ in the time range (t_1, t_2) is

$$X(f) \equiv \int_{t1}^{t2} x(t)\, e^{-2\pi ift}\, dt$$

for any frequency f in the two-sided frequency domain $(-F, +F)$. If $x(t)$ is expressed in units of Pascal, $X(f)$ is expressed in units of Pa/Hz. In Praat, this complex spectrum is the quantity stored in a Spectrum.

From the complex spectrum we can compute the *one-sided power spectral density* in Pa^2/Hz as

$$PSD(f) \equiv 2|X(f)|^2 / (t_2 - t_1)$$

where the factor 2 is due to adding the contributions from positive and negative frequencies. In Praat, this power spectral density is the quantity stored in a Spectrogram.

Спектр и спектральная плотность

It is often useful to express the power spectral density in dB relative to $P_{ref} = 2 \cdot 10^{-5}$ Pa:

$$PSD_{dB}(f) = 10 \log_{10} \{ PSD(f) / P_{ref}^2 \}$$

Since the argument of the logarithm is in units of Hz^{-1}, this spectral measure can loosely be said to be in units of “dB/Hz”. In Praat, this logarithmic power spectral density is the quantity stored in an Ltas; it is also the quantity shown in pictures of a Spectrum and a Spectrogram.

Фильтрация сигнала

low-pass filter

high-pass filter

band-pass filter

band-stop filter

Спектральные моменты

If the complex spectrum is given by $S(f)$, where f is the frequency, the *centre of gravity* is given by

$$\int_0^\infty f\,|S(f)|^p\,df$$

divided by the "energy"

$$\int_0^\infty |S(f)|^p\,df$$

Thus, the centre of gravity is the average of f over the entire frequency domain, weighted by $|S(f)|^p$. For $p = 2$, the weighting is done by the power spectrum, and for $p = 1$, the weighting is done by the absolute spectrum. A value of $p = 2/3$ has been seen as well.

Спектральные моменты

If the complex spectrum is given by $S(f)$, the nth central spectral moment is given by

$$\int_0^\infty (f - f_c)^n \, |S(f)|^p \, df$$

divided by the "energy"

$$\int_0^\infty |S(f)|^p \, df$$

In this formula, f_c is the spectral centre of gravity (see Spectrum: Get centre of gravity...). Thus, the nth central moment is the average of $(f - f_c)^n$ over the entire frequency domain, weighted by $|S(f)|^p$. For $p = 2$, the weighting is done by the power spectrum, and for $p = 1$, the weighting is done by the absolute spectrum. A value of $p = 2/3$ has been seen as well.

Спектральные моменты

For $n = 1$, the central moment should be zero, since the centre of gravity f_c is computed with the same p. For $n = 2$, you get the variance of the frequencies in the spectrum; the standard deviation of the frequency is the square root of this. For $n = 3$, you get the non-normalized spectral skewness; to normalize it, you can divide by the 1.5 power of the second moment. For $n = 4$, you get the non-normalized spectral kurtosis; to normalize it, you can divide by the square of the second moment and subtract 3. Praat can directly give you the quantities mentioned here:

Спектральные признаки

- **Alpha Ratio**, ratio of the summed energy from 50–1000 Hz and 1–5 kHz
- **Hammarberg Index**, ratio of the strongest energy peak in the 0–2 kHz region to the strongest peak in the 2–5 kHz region.
- **Spectral Slope 0–500 Hz and 500–1500 Hz**, linear regression slope of the logarithmic power spectrum within the two given bands.

Спасибо за внимание!