# Short-time Fourier Transform (STFT)

$$X_l[k] = \sum_{n=-N/2}^{N/2-1} w[n]\, x[n+lH]\, e^{-j2\pi kn/N} \qquad l = 0, 1, \ldots,$$

w: analysis window
l: frame number
H: Hop size

---

The output is a sequence of spectra. each one have same size of magnitudes and phase. but each one differently because input will be a different fragment of the sound.

---

Transform of a windowed sinewave

$$x[n] = A_0 \cos(2\pi k_0 n/N) = \frac{A_0}{2} e^{j2\pi k_0 n/N} + \frac{A_0}{2} e^{-2\pi k_0 n/N}$$

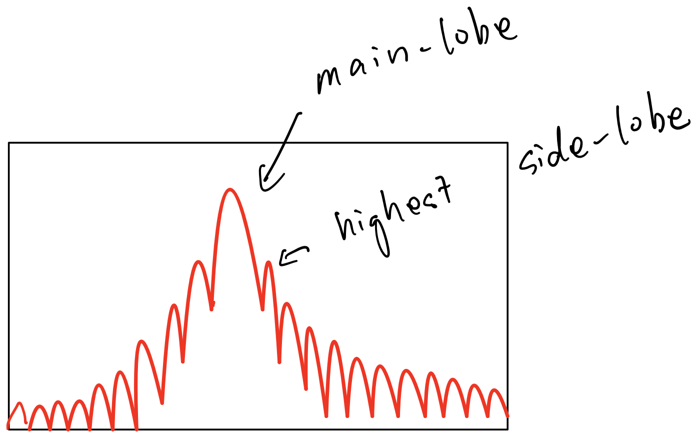
$$x[k] = \sum_{n=-N/2}^{N/2-1} w[n]\, x[n]\, e^{-j2\pi kn/N}$$

$$= \sum_{n=N/2}^{N/2-1} W[n] \left( \frac{A_0}{2} e^{j2\pi k_0 n/N} + \frac{A_0}{2} e^{-j2\pi k_0 n/N} \right) e^{-j2\pi kn/N}$$

$$= \sum_{n=-N/2}^{N/2-1} W[n] \frac{A_0}{2} e^{j2\pi k_0 n/N} e^{-j2\pi kn/N} + \sum_{n=-N/2}^{N/2-1} W[n] \frac{A_0}{2} e^{-j2\pi k_0 n/N} e^{-j2\pi kn/N}$$

$$= \frac{A_0}{2} \sum_{n=-N/2}^{N/2-1} W[n]\, e^{-j2\pi(k-k_0)n/N} + \frac{A_0}{2} \sum_{n=-N/2}^{N/2-1} W[n]\, e^{-j2\pi(k+k_0)n/N}$$

$$= \frac{A_0}{2} W[k - k_0] + \frac{A_0}{2} W[k + k_0]$$




① We wish to have narrow main-lobe and low side-lobe.

② magnitude spectrum as smooth as possible.

---

## Window size.

larger window size will have more information capture.

---

## Even-odd window size

won't affect magnitude spectrum, but affect phase spectrum

odd size of samples can be perfectly symmetric around zero and get the zero phase value

## FFT size

larger FFT size will have much smoother magnitude spectrum.

---

## Hop size

window size · $\frac{1}{4}$ is good

---

## Time frequency compromise

window size ↓

time resolution ↑ frequency resolution ↓

window size ↑

time resolution ↓ frequency resolution ↑

Inverse STFT

shift progressively

DFT

$$y[n] = \sum_{l=0}^{L-1} Shift_{lH,n} \left[ \frac{1}{N} \sum_{k=-N/2}^{N/2-1} X_l[k] \, e^{j2\pi kn/N} \right]$$

each output frame is :

$$yw_l[n] = x(n + lH) \, w[n]$$

and the output sound is :

$$y[n] = \sum_{l=0}^{L-1} yw_1[n] = x[n] \sum_{l=0}^{L-1} w[n - lH]$$

STFT system