Cepstrum

G. Anushiya Rachel
Assistant Professor, CSE
Shiv Nadar University Chennai

Homomorphic Systems

- Homomorphic systems for convolution obey a generalized principle of superposition.
- Superposition principle for conventional linear systems:

$$L[x(n)] = L[x_1(n) + x_2(n)]$$
$$= L[x_1(n)] + L[x_2(n)]$$
$$= y_1(n) + y_2(n) = y(n)$$

and $L[ax(n)] = aL[x(n)]$

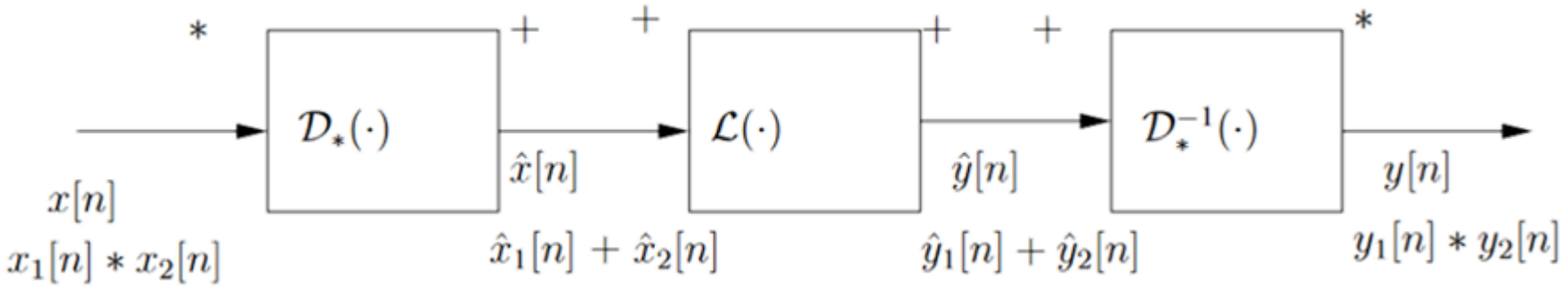
  - If the input is a linear combination of elementary signals then the output is a linear combination of corresponding outputs.
  - + → additive combination
- We can define a class of systems where addition is replaced by convolution in the superposition principle → Homomorphic systems

Homomorphic Systems

- Homomorphic systems will have the following property:

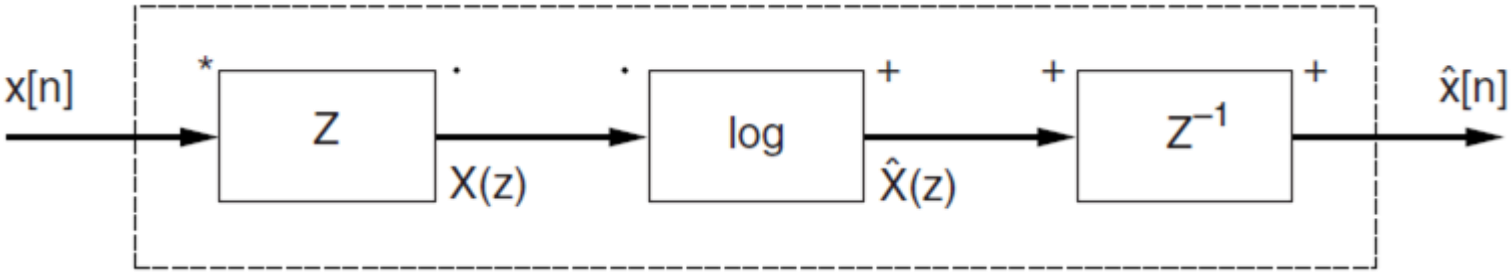
$$\begin{aligned} H[x(n)] &= H[x_1(n) * x_2(n)] \\ &= H[x_1(n)] * H[x_2(n)] \\ &= y_1(n) * y_2(n) = y(n) \end{aligned}$$

- Any homomorphic system can be represented as a cascade of three systems.
  - Characteristic system ($D_*(\ )$) – transforms convolution to addition
  - Linear system that obeys the superposition principle
  - Inverse characteristic system ($D_*^{-1}(\ )$) – transforms addition to convolution

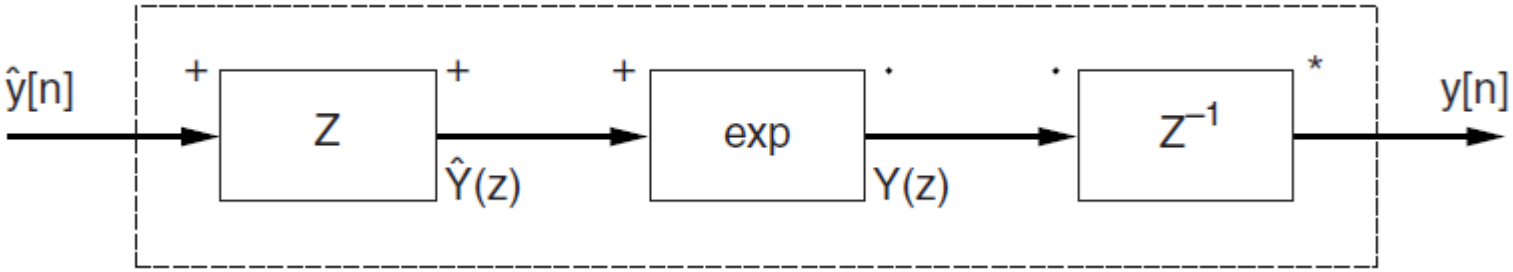


Homomorphic Systems

- Characteristic system:
  - To convert convolution to addition, convert from time domain to frequency and domain and take the logarithm.



- Inverse characteristic system:



Cepstrum

- In the characteristic system, we compute the logarithm of $X(z)$ or $X(e^{j\omega})$, which is complex and is given by

$$\hat{X}(e^{j\omega}) = \log|X(e^{j\omega})| + j\arg[X(e^{j\omega})]$$

- Inverse FT of the above is the output of the characteristic system and is given by

$$\hat{x}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\hat{X}(e^{j\omega})e^{j\omega n}d\omega$$

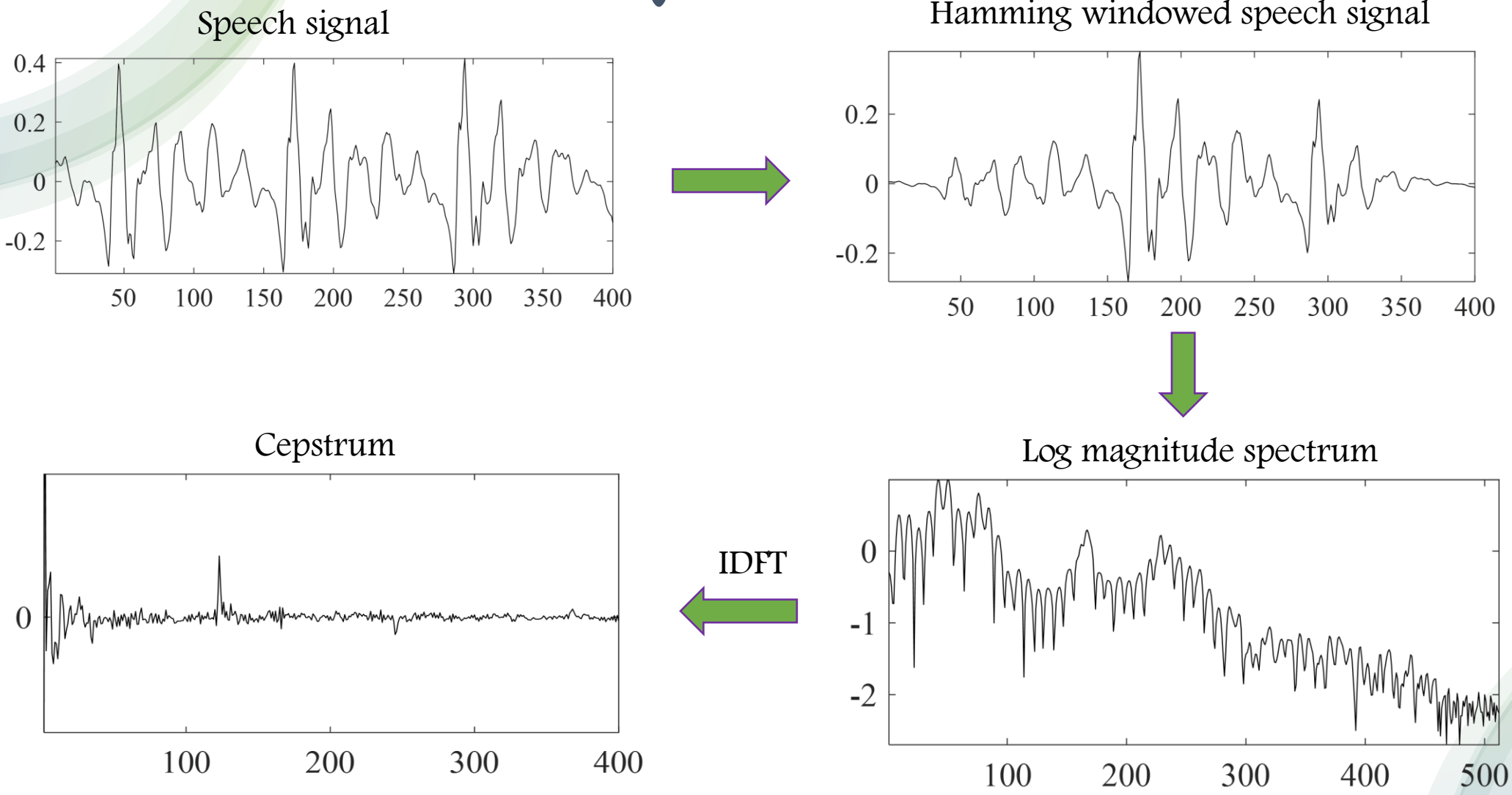
This is called the complex cepstrum.

- If we consider only the real part of $\hat{X}(e^{j\omega})$,

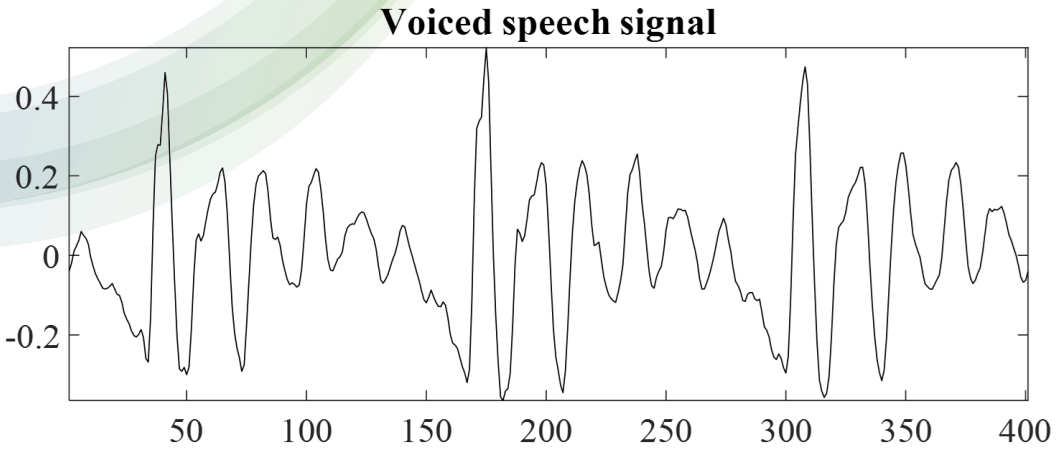
$$c(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\log|X(e^{j\omega})|\,e^{j\omega n}d\omega$$

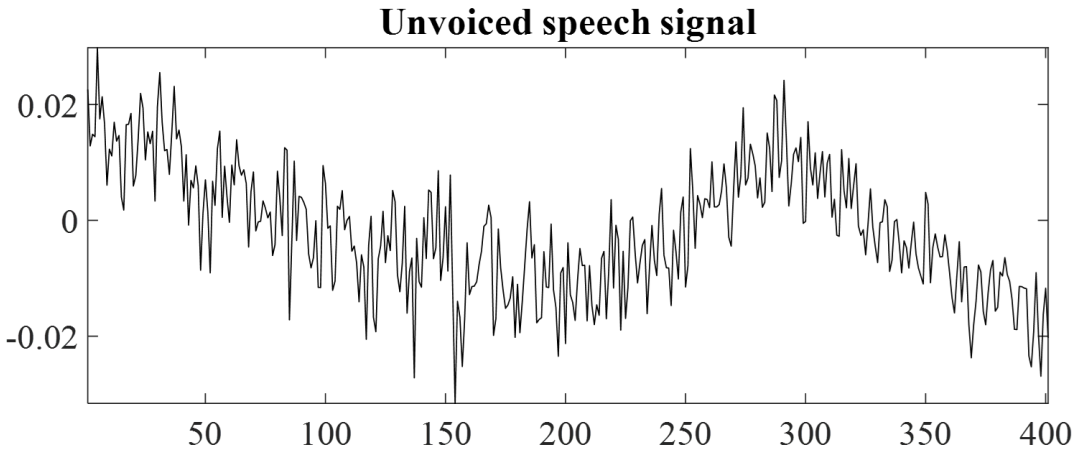
This is called the cepstrum.

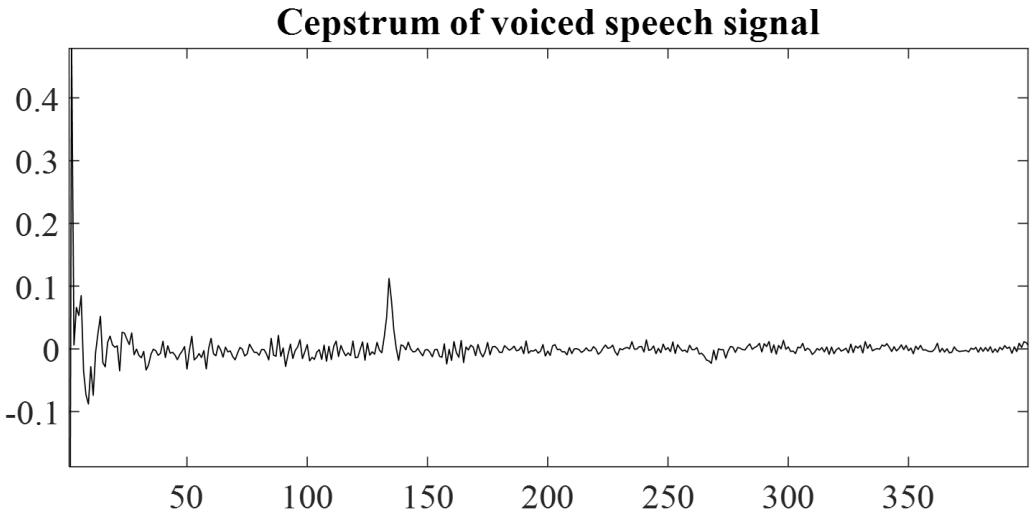
Cepstrum

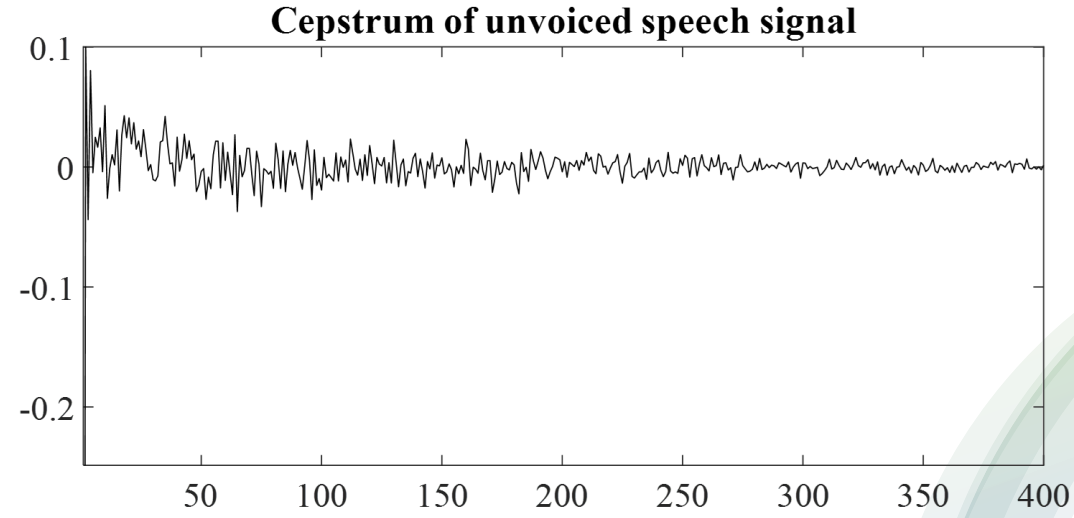

Speech signal
0.4
0.2
0
-0.2
50 100 150 200 250 300 350 400
Hamming windowed speech signal
0.2
0
-0.2
50 100 150 200 250 300 350 400
Log magnitude spectrum
0
-1
-2
100 200 300 400 500
IDFT
Cepstrum
0
100 200 300 400

Cepstrum – Voiced vs Unvoiced







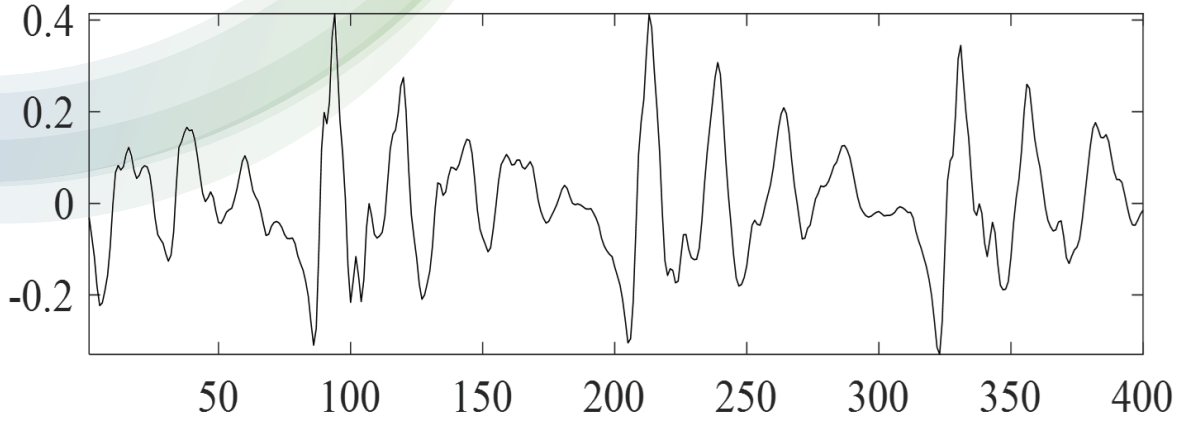


Cepstrum-Based Pitch Estimation

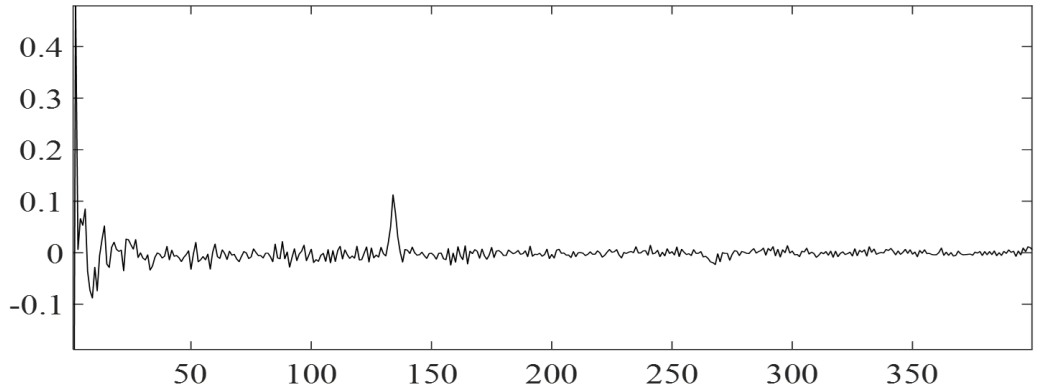
- The cepstrum of a voiced segment of speech contains a strong peak in the range of 3 to 20 ms.
- The location of this peak gives an estimate of the pitch period.
- For an unvoiced sound, there is no dominant peak in the cepstrum.
- Drawback:
  - Certain voiced segments of speech may not show a strong peak in the cepstrum.
  - The existence of the peak depends on a number of factors:
    - Type and length of the window applied
- With appropriate window length, the location and amplitude of the cepstrum peak provide a reliable estimate of the pitch and voicing.
- In cases where the cepstrum fails, other information such as ZCR and STE may be additionally used and the pitch and voicing may be forced to vary smoothly.

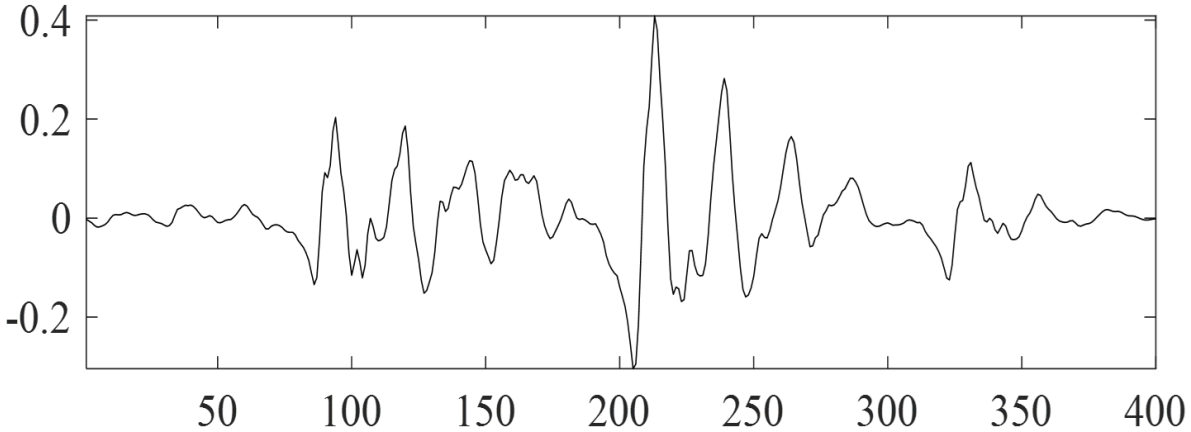
Cepstrum-Based Pitch Estimation

Rectangular-windowed speech signal

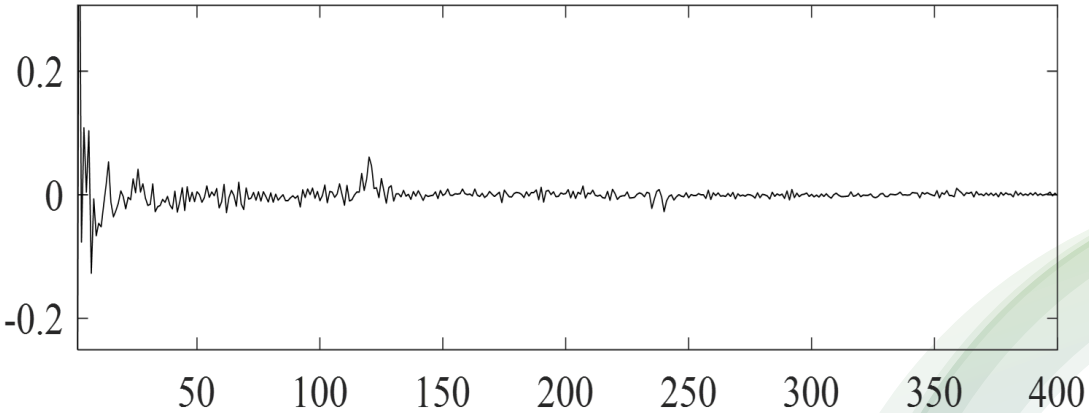


Cepstrum of rectangular-windowed speech





Hamming-windowed speech signal with only 1 prominent excitation

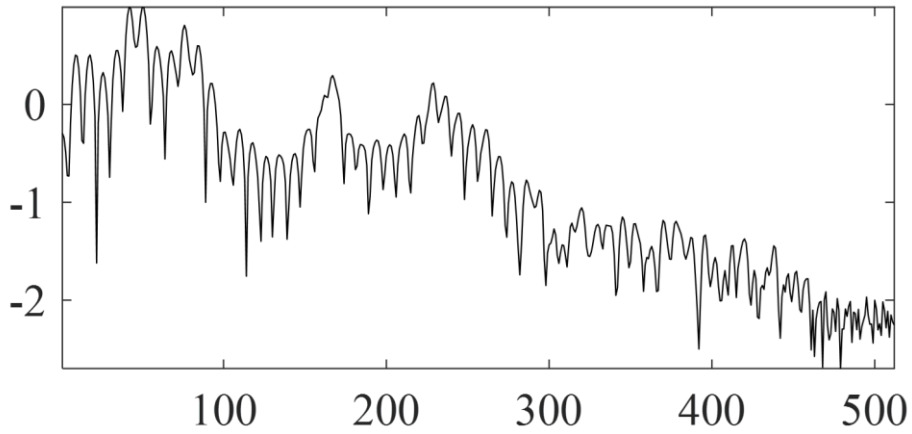


Cepstrum when the signal has only one excitation

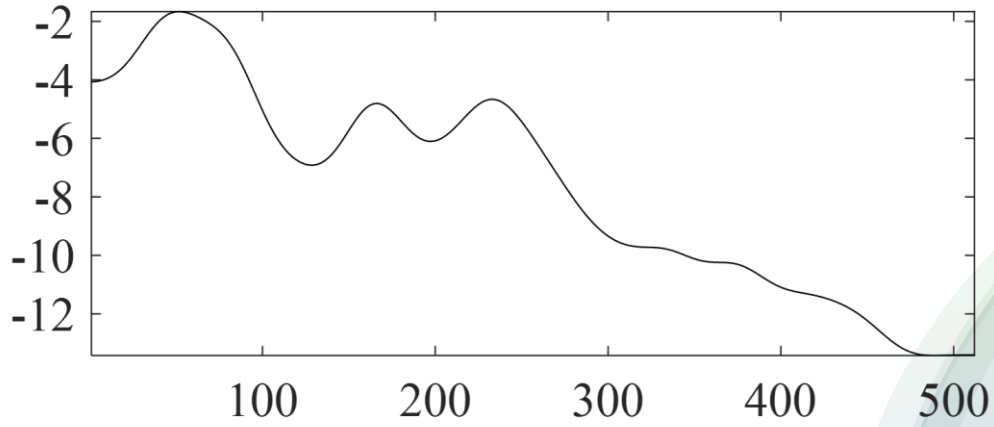
Cepstrum-Based Formant Estimation

- Lower half of the cepstrum contains the system information.
- Apply a one-sided hamming window to extract the first few samples (around 20 samples, less than one pitch period) from the cepstrum.
- Derive the magnitude spectrum from the above signal.
- The magnitude spectrum will contain peaks at the formant frequencies.
- Therefore, the location of the peaks will yield an estimate of the formant frequencies.

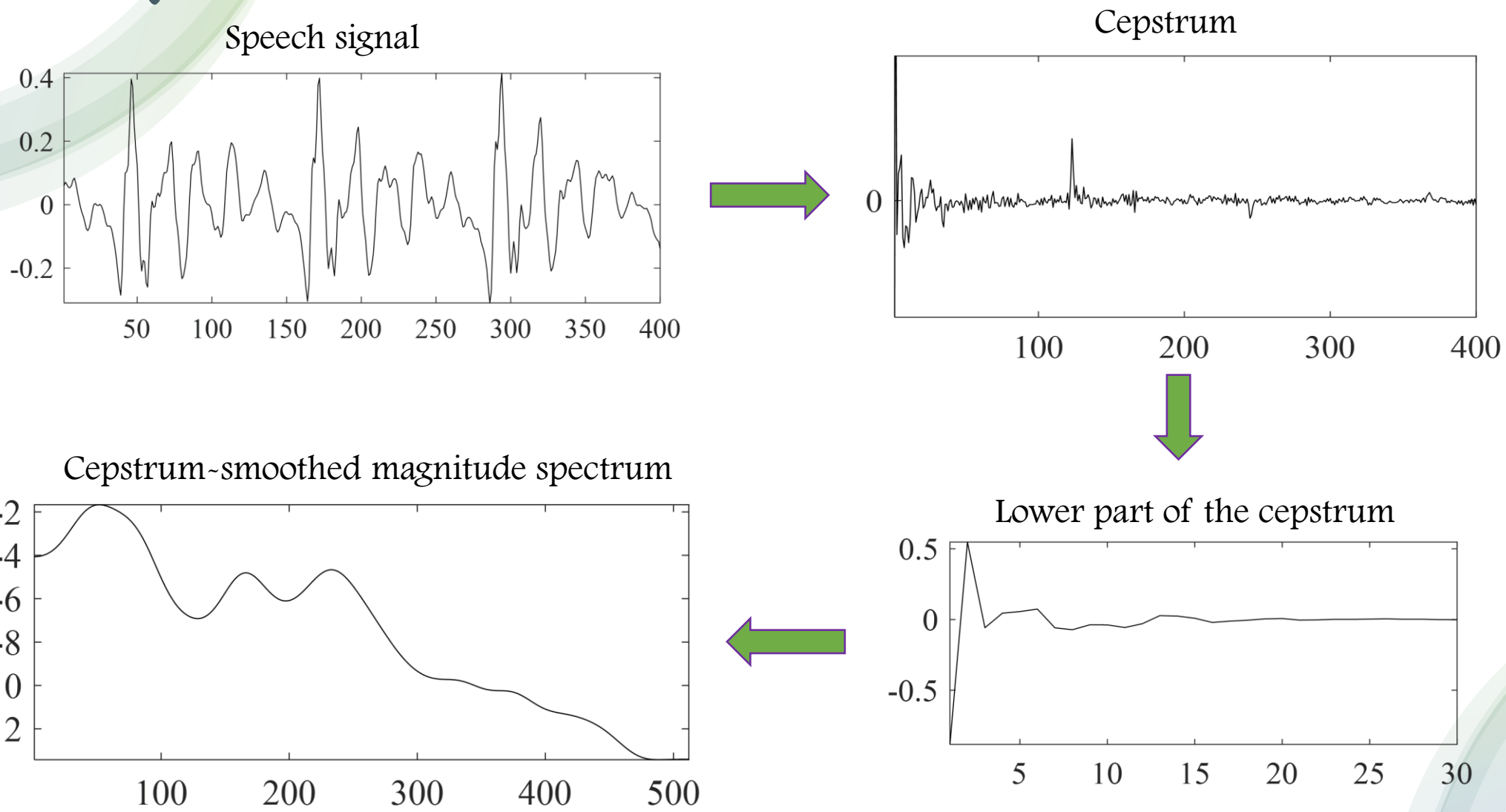
Original magnitude spectrum



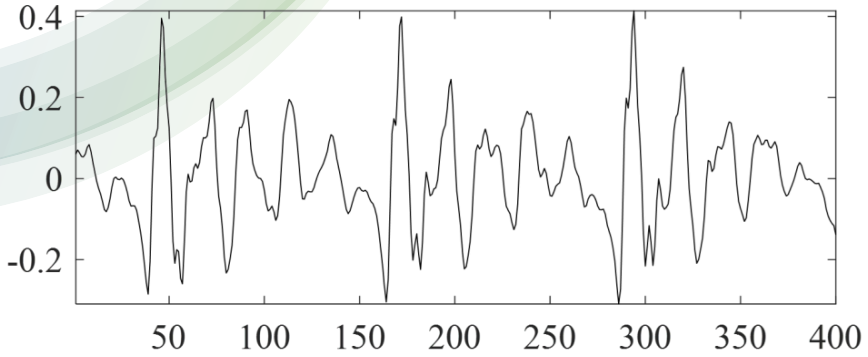
Cepstrum-smoothed magnitude spectrum

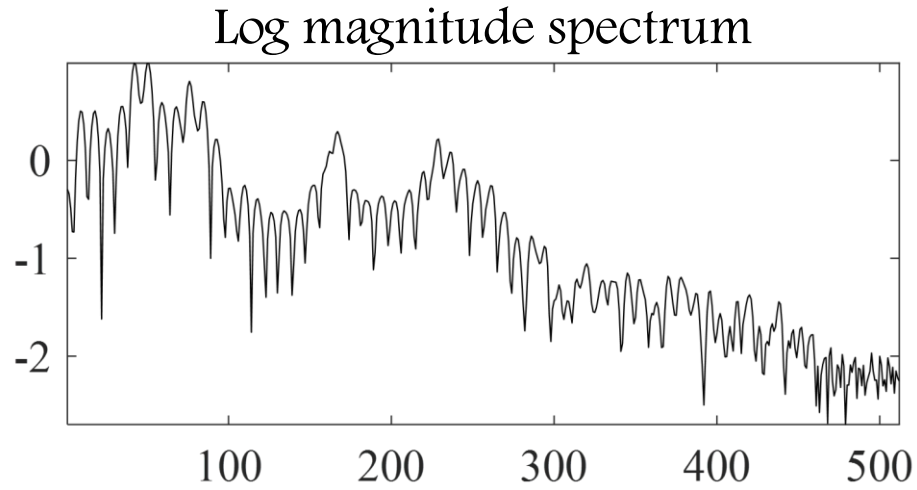


Cepstrum-Based Formant Estimation



Mel Frequency Cepstral Coefficients





- Mel Frequency Cepstral Coefficients
  - Convert linear frequency scale of the magnitude spectrum to Mel scale and then obtain the cepstrum.