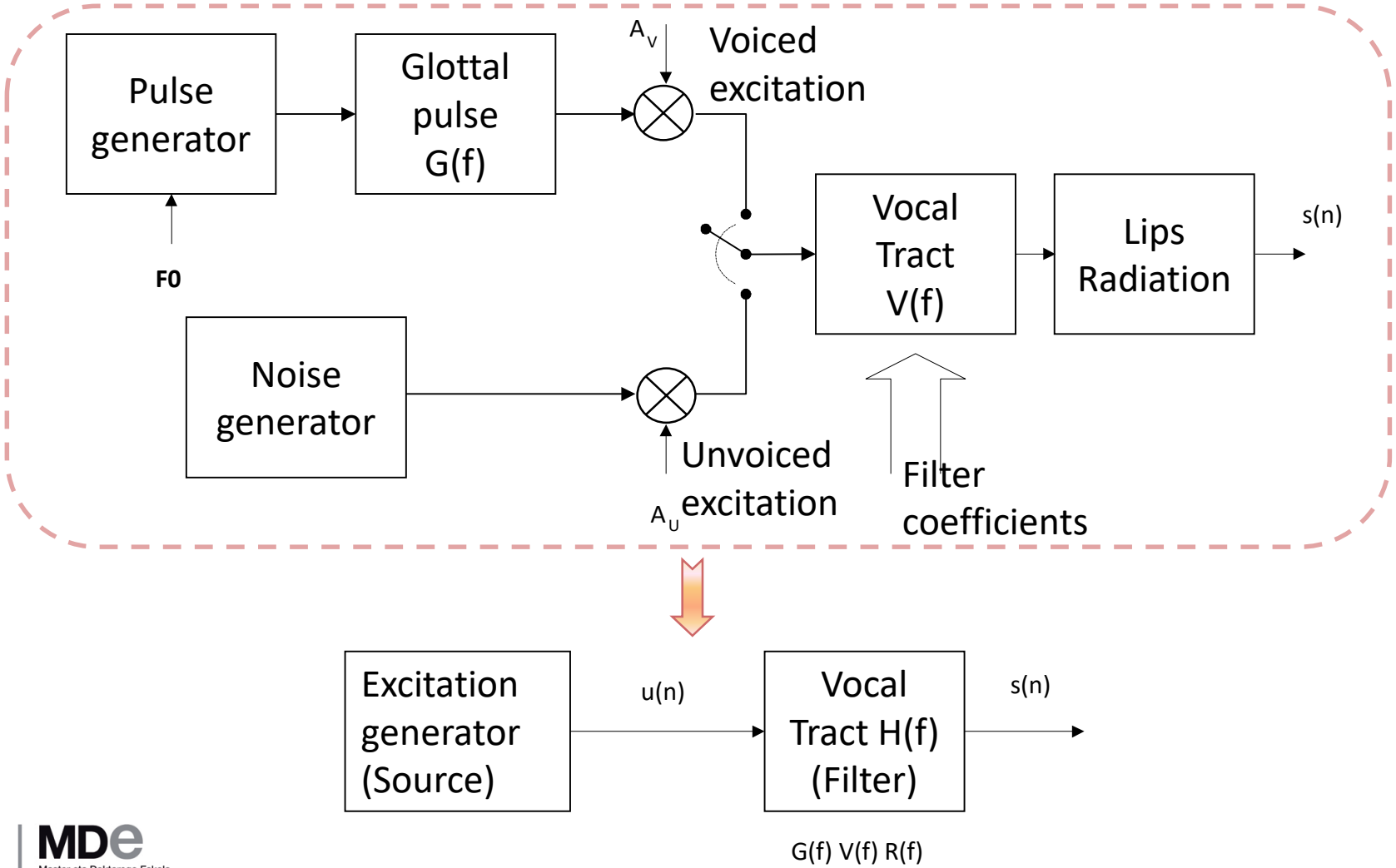
Basic concepts about signals & systems (Part III)

Outline

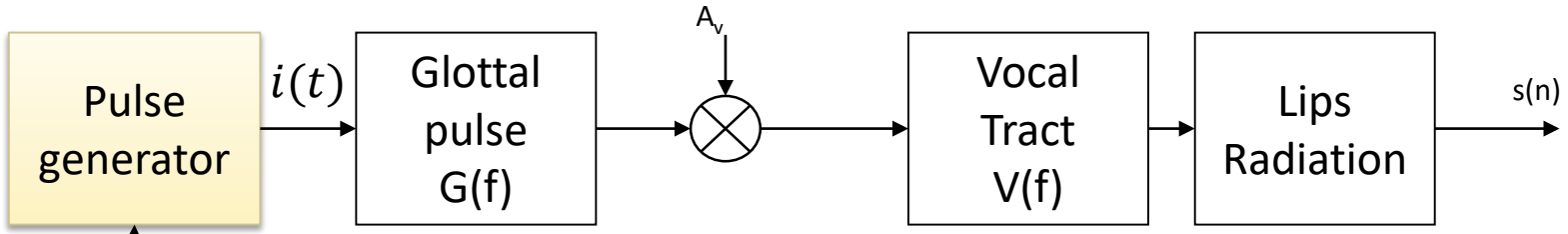
1. Introduction
2. Basic signals and operations
3. Linear Time Invariant Systems
4. The Fourier Transform
5. Filters and resonators
6. The source-filter model

Speech production: The source-filter model

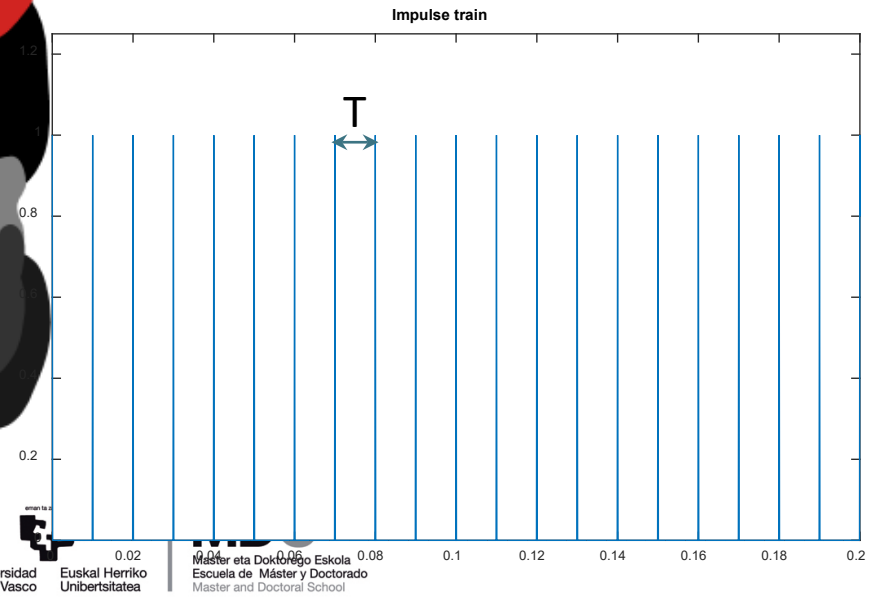


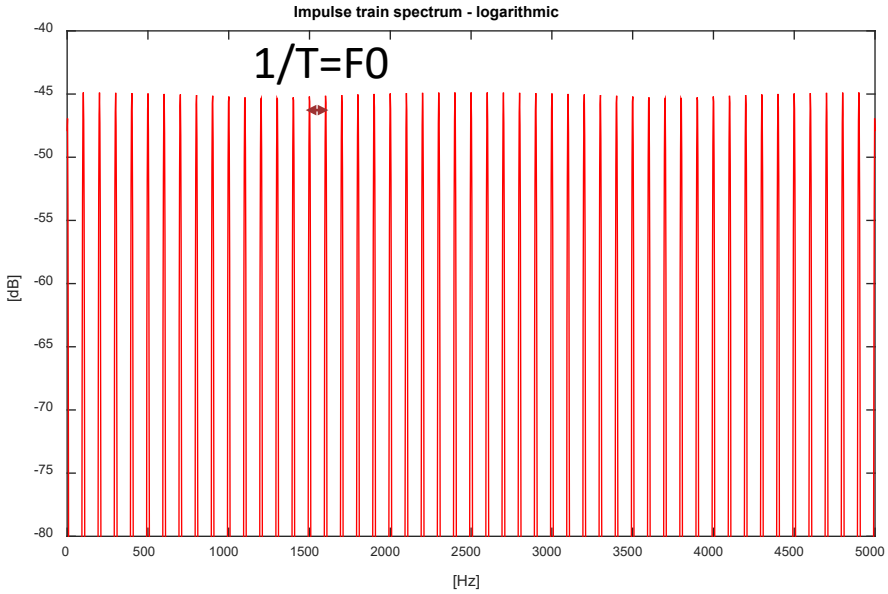
Speech production: The source-filter model

Voiced sounds



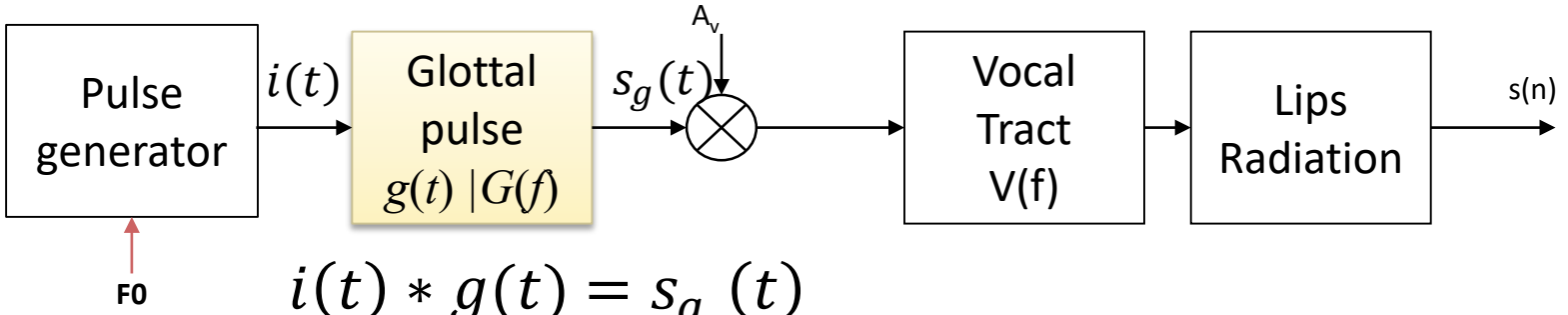
$$i(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT) \quad \xrightarrow{F} \quad I(f) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T}\right)$$





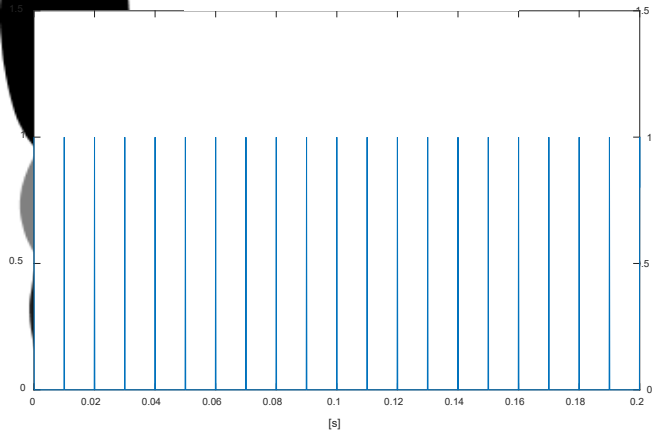
Speech production: The source-filter model

Voiced sounds

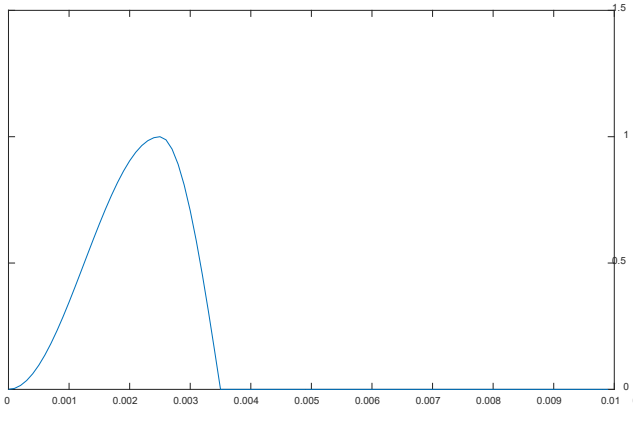


$$i(t) * g(t) = s_g\ (t)$$

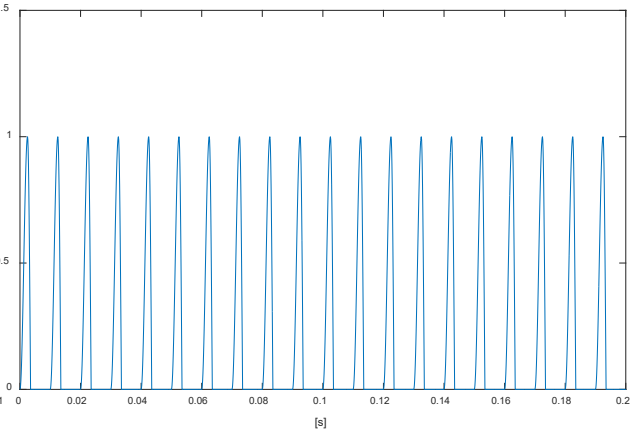
Impulse train



Glottal pulse (impulse response)

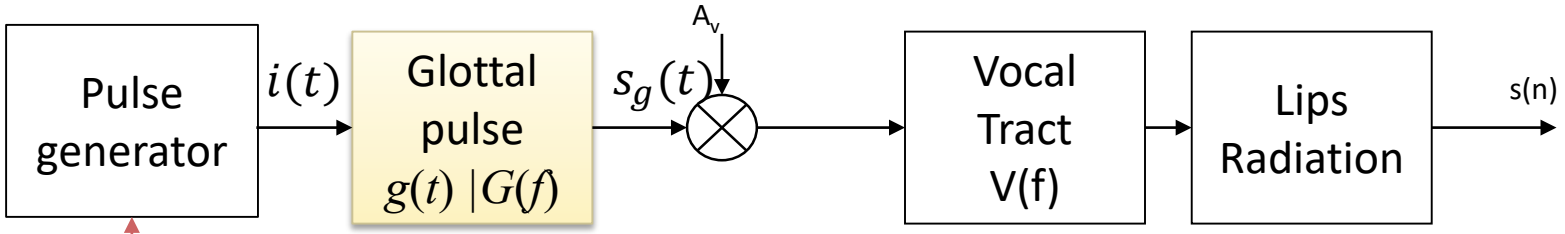


Glotal signal (pulse train)



Speech production: The source-filter model

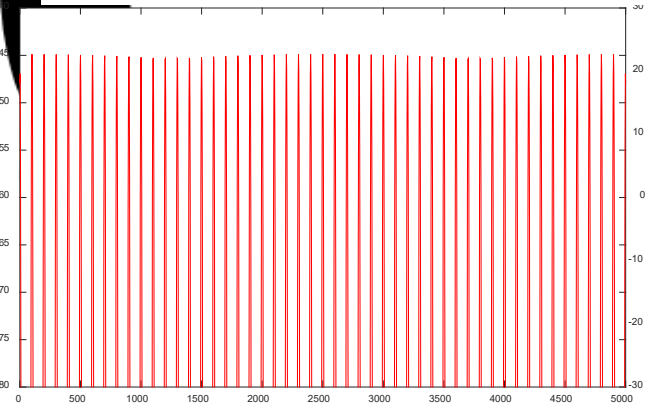
Voiced sounds



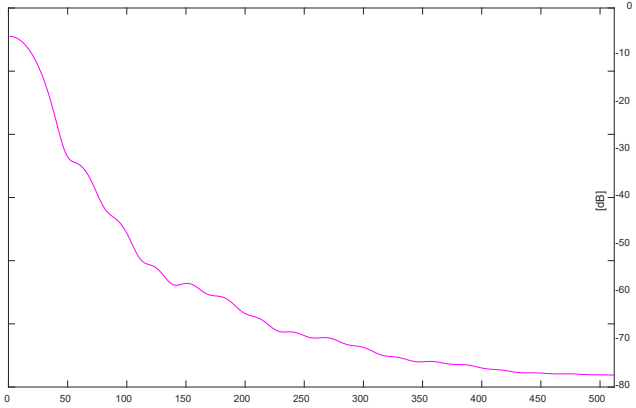
$$I(f) \cdot G(f) = S_g(f)$$

$$20 \log I(f) + 20 \log G(f) = 20 \log S_g(f)$$

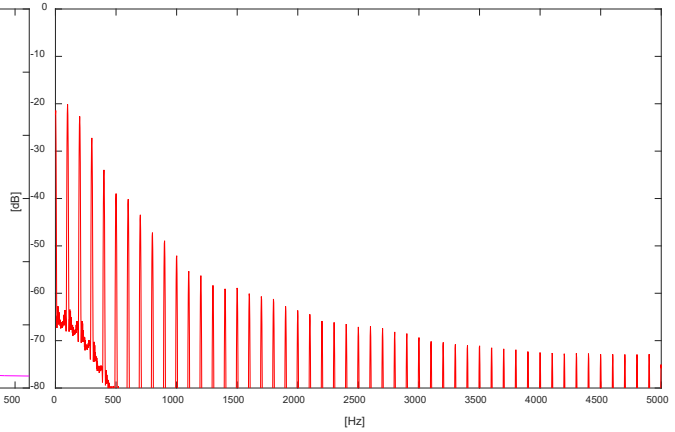
Glottal impulse train log-spectrum



Glottal pulse log- spectrum

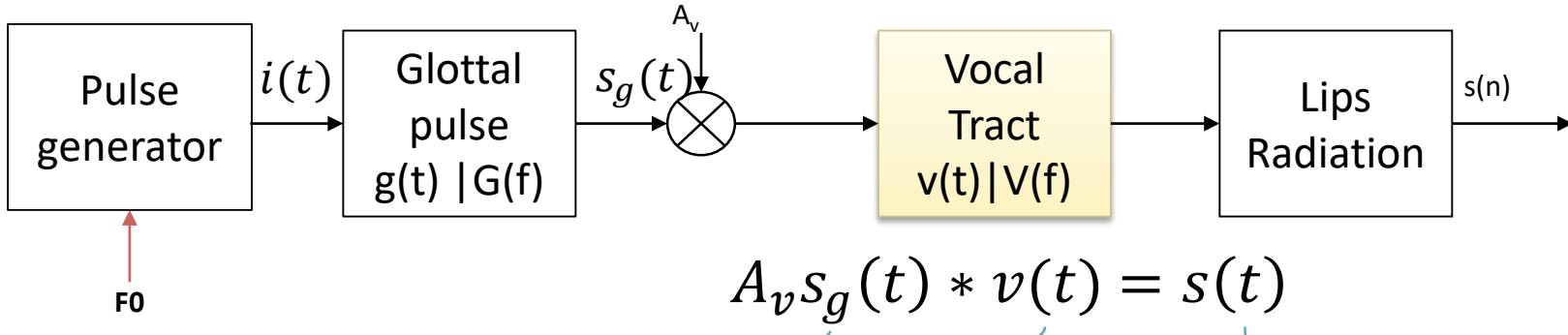


Glottal pulse train log-spectrum



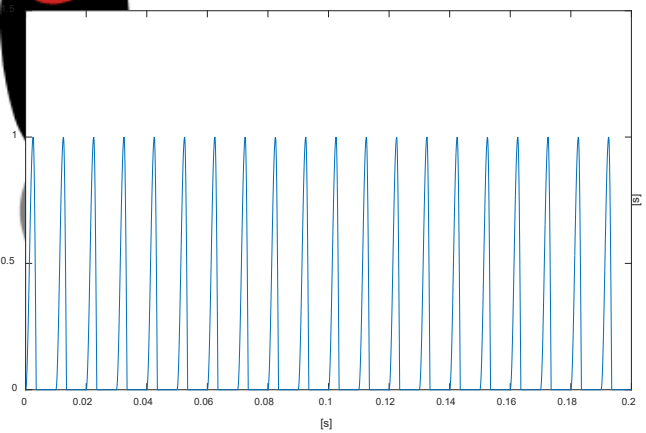
Speech production: The source-filter model

Voiced sounds

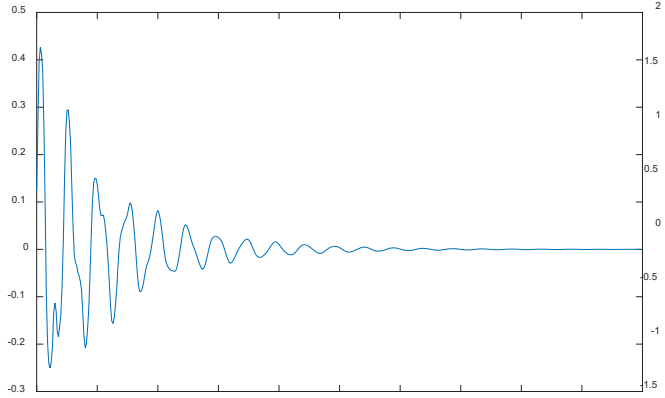


$$A_v s_g(t) * v(t) = s(t)$$

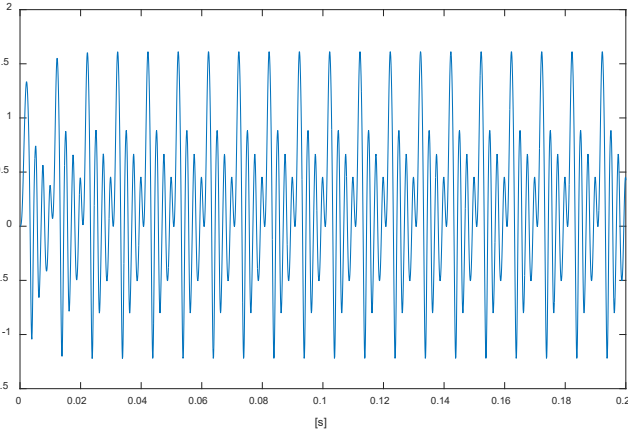
Glottal signal (pulse train)



Vocal tract impulse response

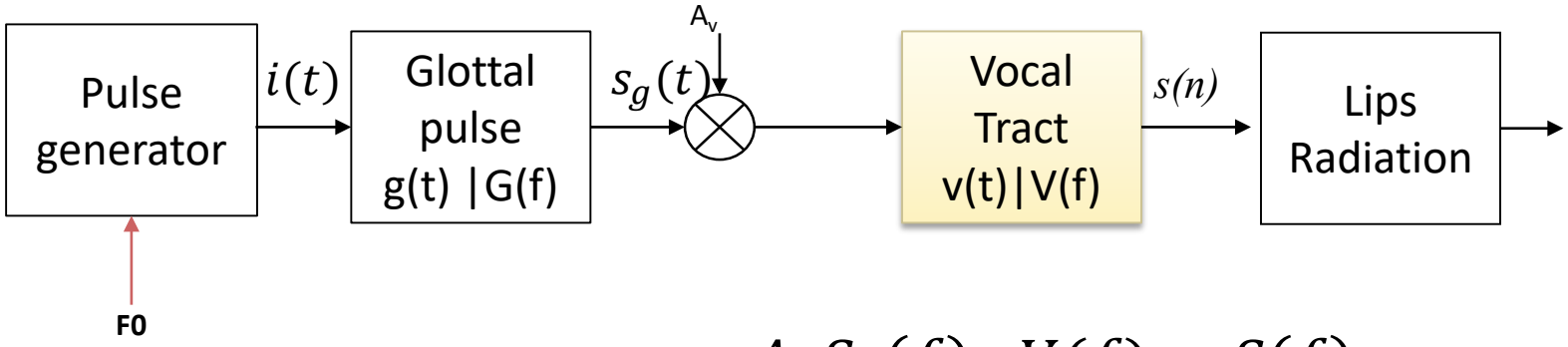


Speech signal (before lips radiation)



Speech production: The source-filter model

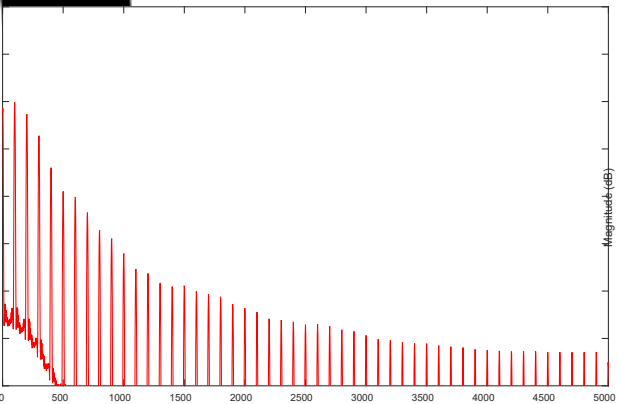
Voiced sounds



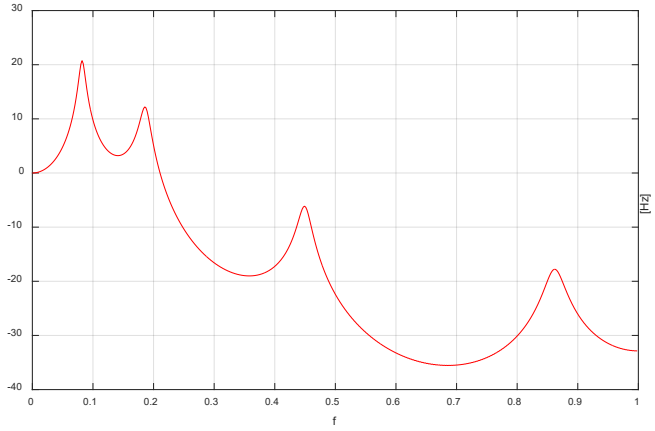
$$A_v S_g(f) \cdot V(f) = S(f)$$

$$20\log(A_v S_g(f)) + 20\log V(f) = 20\log S(f)$$

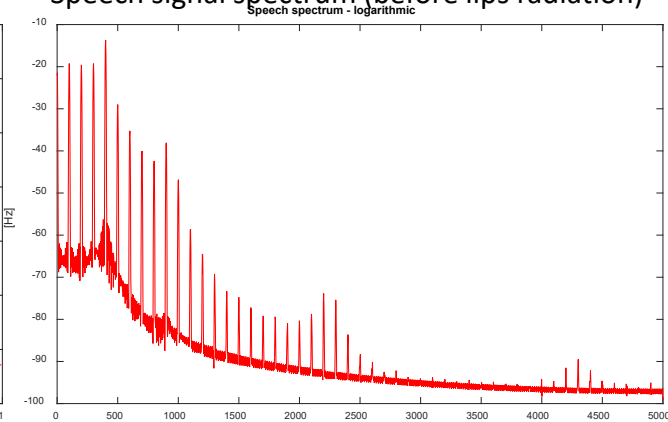
Glottal signal log spectrum



Vocal tract frequency response

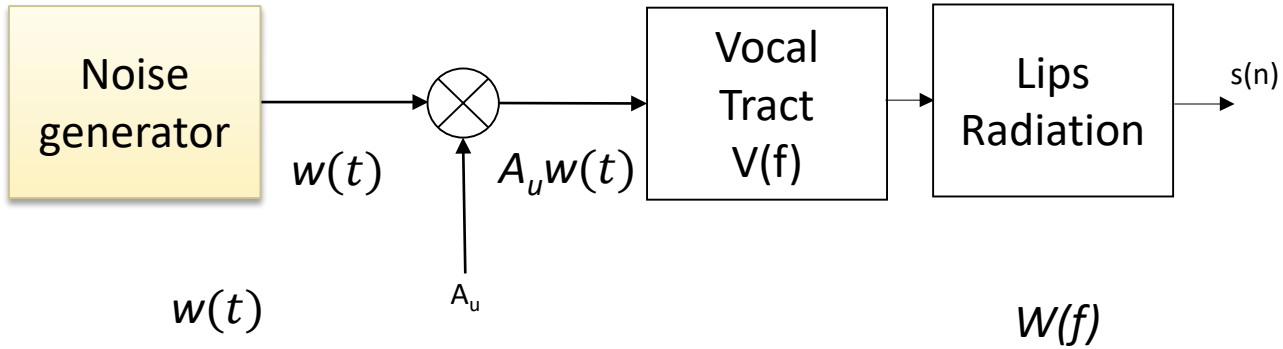


Speech signal spectrum (before lips radiation)



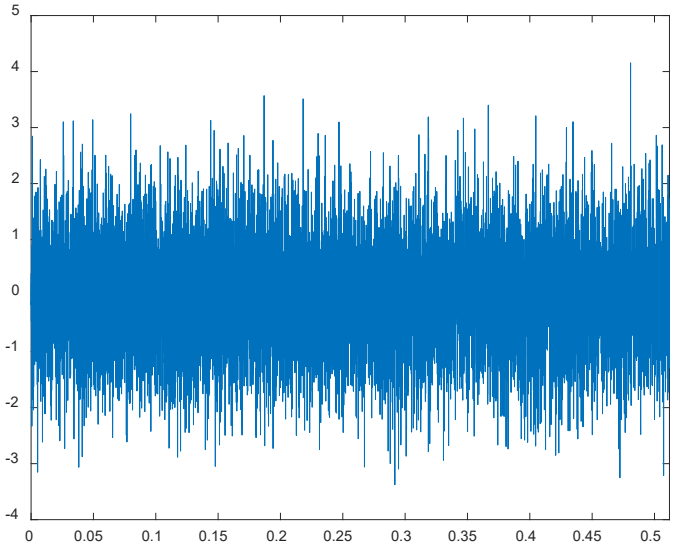
Speech production: The source-filter model

Unvoiced sounds

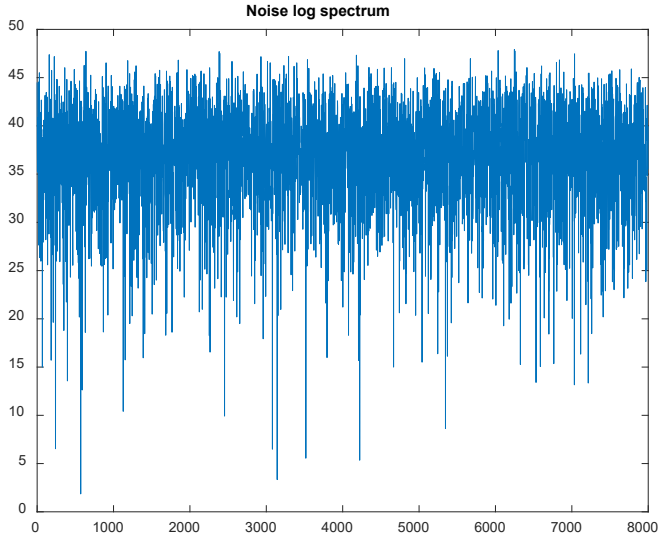


$w(t)$

Noise excitation

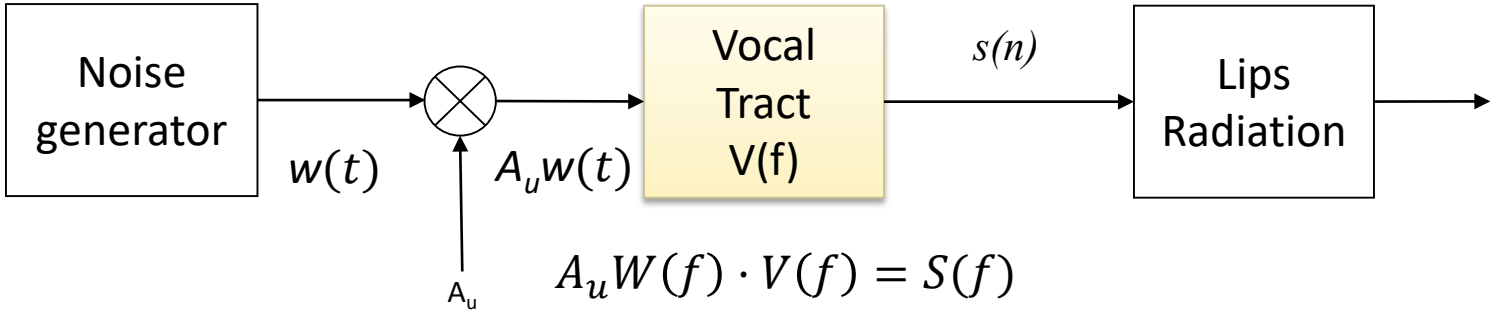


$W(f)$



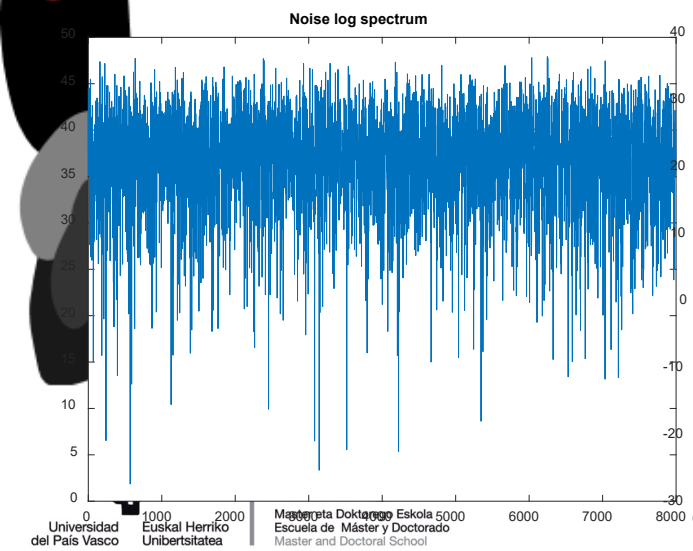
Speech production: The source-filter model

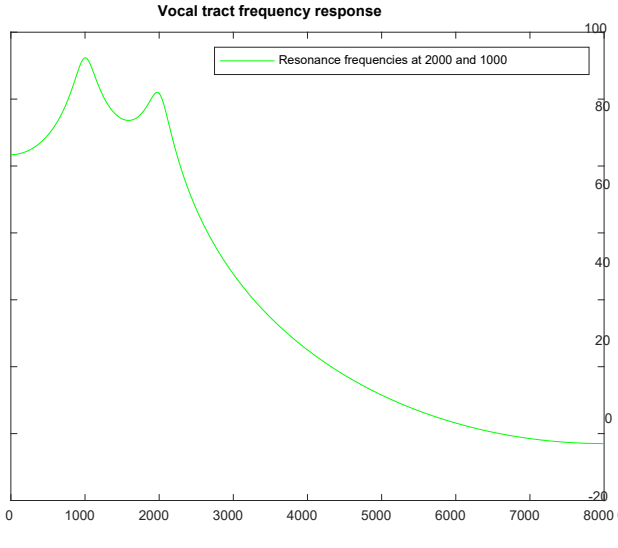
Unvoiced sounds

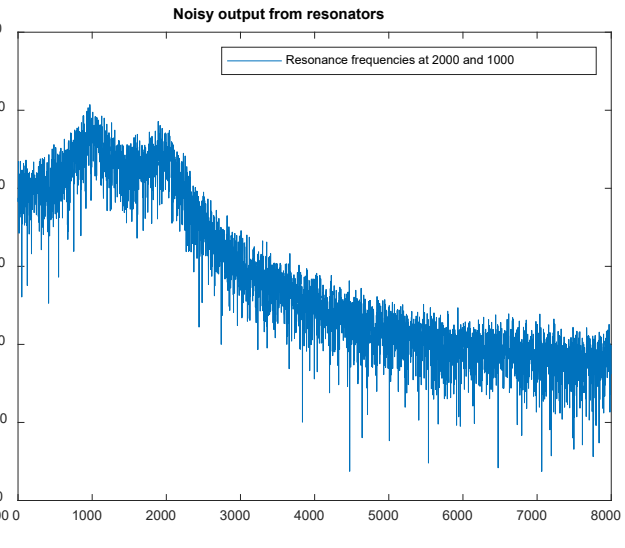


$$A_u W(f) \cdot V(f) = S(f)$$

$$20\log(A_u W(f)) + 20\log V(f) = 20\log S(f)$$







Speech production: The source-filter model

Klatt formant synthesizer

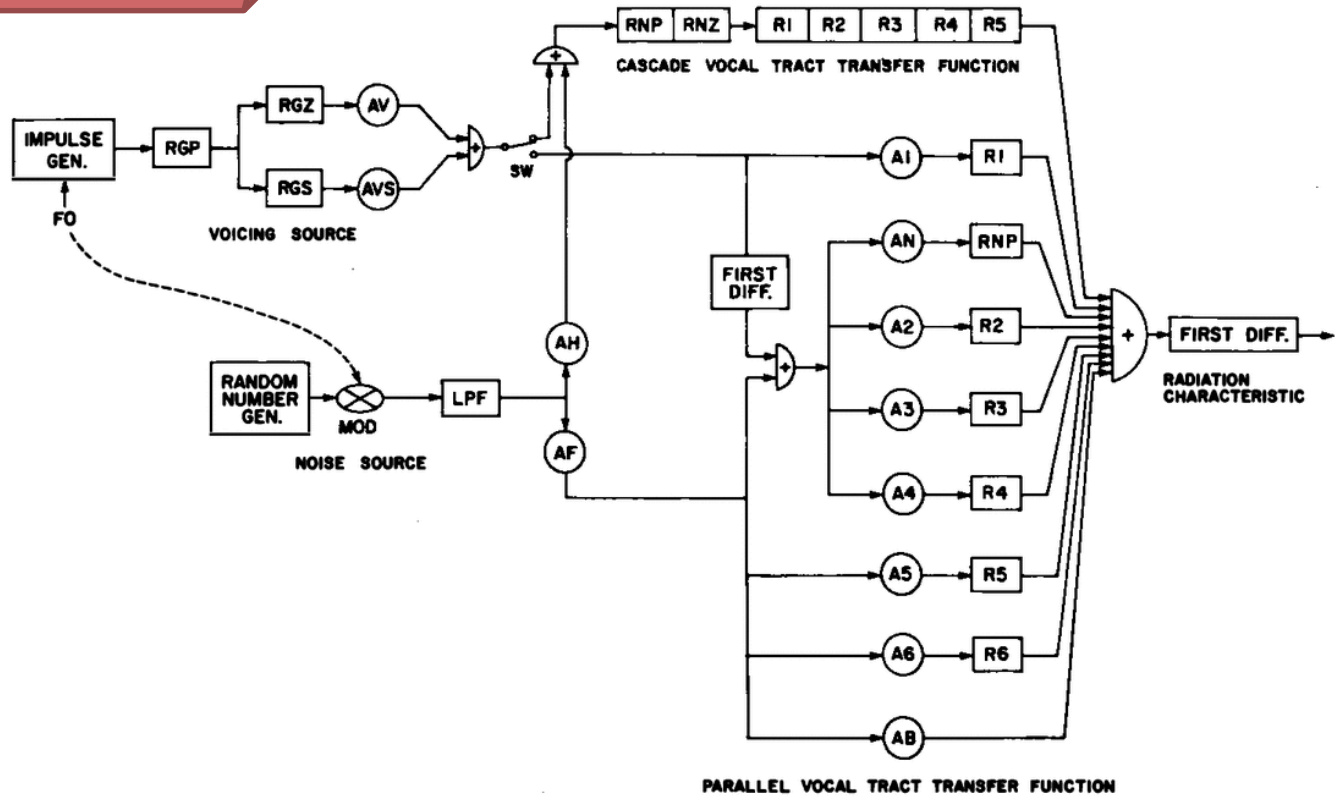


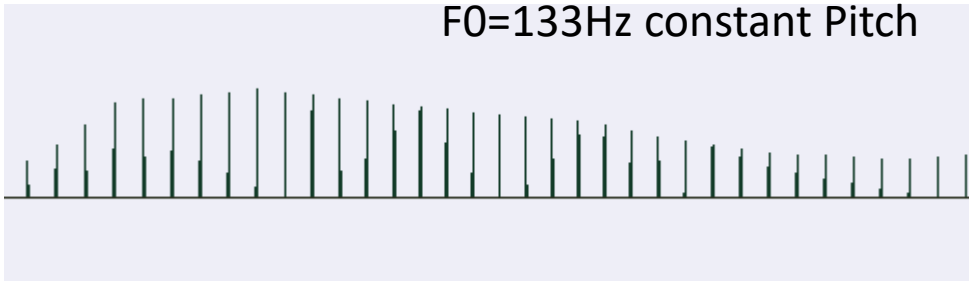
Figure 2. Block diagram of a formant synthesizer (Klatt, 1980).

Phonetic Commands Converted to Motions of the Articulators

Universidad del País Vasco
Euskal Herriko Unibertsitatea
Master eta D
Escuela de
Master and

Original

The source-filter model



Excitation generator (Source)

u(n)

Vocal Tract H(f) (Filter)

s(n)

All voiced excitation

All unvoiced excitation

The source-filter model

- Separates the characteristics of the source and the vocal tract (some models include interaction)
- It is defined for stationary sounds: works better when the parameters change slowly
- In many systems, V(f) is an *all-resonances* filter: it cannot model the *antiresonances* needed for nasal sounds.
- Typically, the source must be either voiced or unvoiced, which is not always realistic (voiced fricatives). Improves with mixed excitation
- Sudden change of source: non-realistic. Improves with smoothing.

Outline

1. Introduction
2. Basic signals and operations
3. Linear Time Invariant Systems
4. The Fourier Transform
5. Filters and resonators
6. The source-filter model

Recommended Lectures

- Lawrence Rabiner and Ronald W. Schafer “Digital Processing of Speech Signals” 1979 (Ed.Prentice Hall) (Chapter 3)

- Klatt, D.; *Software for a cascade/parallel formant synthesizer*, Journal of the Acoustical Society of America, 67(3) 1980

- Klatt, D.; Klatt, L.; *Analysis, synthesis and perception of voice quality variations among male and female talkers*. The Journal of the Acoustical Society of America, 1990, DOI:10.1121/1.398894