KU LEUVEN

Digital Signal Processing 2

Les 2: Lineaire predictie

Prof. dr. ir. Toon van Waterschoot

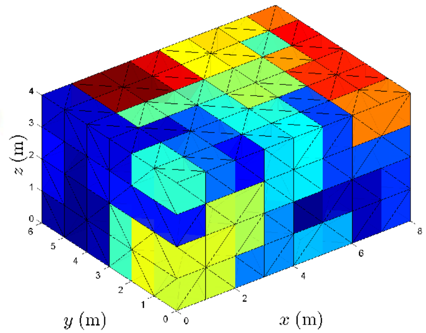
Faculteit Industriële Ingenieurswetenschappen
ESAT – Departement Elektrotechniek
KU Leuven, Belgium

Onderzoeksafdeling

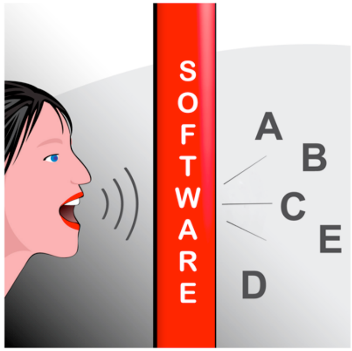
- **STADIUS** Centrum voor Dynamische Systemen, Signaalverwerking en Data-Analyse:
  - Dynamische Systemen: identificatie, optimalisatie, regeltechniek, systeemtheorie
  - Signaalverwerking: spraak- & audioverwerking, digitale communicatie, biomedische signaalverwerking
  - Data-Analyse: machine learning, bio-informatica

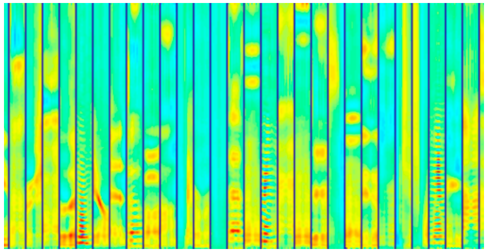
- **AdvISe** – Advanced Integrated Sensing Lab:
  - Biomedisch: biomedische technologie, ambient assisted living
  - Audio: akoestische modellering, audio-analyse, akoestische signaalverbetering
  - Chip-ontwerp: stralingsharde elektronica

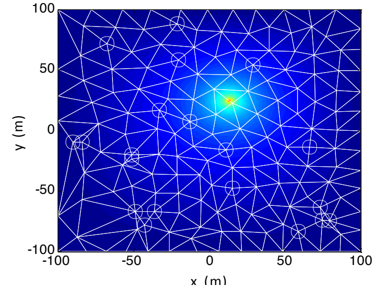
Onderzoekstopics



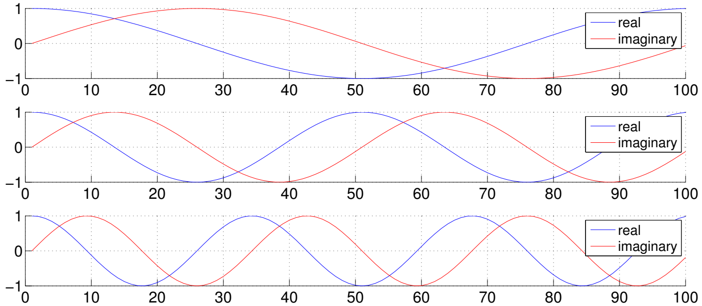
Audio signal analysis
- speech recognition
- event detection
- source localization
- audio classification







Acoustic modeling
- ear modeling
- room modeling
- loudspeaker modeling
- signal modeling



Acoustic signal enhancement
- noise reduction
- echo/feedback control
- room equalization

KU LEUVEN

Contactgegevens

Toon van Waterschoot

- Mail: toon.vanwaterschoot@esat.kuleuven.be

- Kantoor (enkel op Campus Leuven):
  Departement Elektrotechniek (ESAT-STADIUS)
  Kasteelpark Arenberg 10, 3001 Leuven
  telefoon: +32 16 321788

KU LEUVEN

Digital Signal Processing 2: Vakinhoud

- Les 1: Eindige woordlengte
- Les 2: Lineaire predictie
- Les 3: Optimale filtering
- Les 4: Adaptieve filtering
- Les 5: Detectieproblemen
- Les 6: Spectrale signaalanalyse
- Les 7: Schattingsproblemen 1
- Les 8: Schattingsproblemen 2
- Les 9: Sigma-Deltamodulatie
- Les 10: Transformatiecodering

Digital Signal Processing 2: Tijdschema

- **Hoorcolleges: donderdag 8:25 – 10:25**
  - 25/9: Les 1 (P. Karsmakers)
  - 01/10: geen les (practicum)
  - 08/10: geen les (practicum)
  - 15/10: Les 2
  - 23/10: Les 3
  - 30/10: Les 4
  - 06/11: Les 5
  - 13/11: geen les (practicum)
  - 20/11: Les 6
  - 27/11: Les 7
  - 04/12: Les 8
  - 11/12: Les 9
  - 18/12: Les 10

Digital Signal Processing 2: Lesmateriaal

- **Slides**
  - slides = basis lesmateriaal = leerstof voor examen
  - beschikbaar op Toledo

- **Cursustekst**
  - geen vaste cursustekst
  - voor de meeste lessen wordt er een hoofdstuk uit een (Engelstalig) boek of een artikel op Toledo geplaatst

- **Software**
  - tijdens enkele lessen worden Matlab-oefeningen gemaakt of opdrachten gegeven, waarvan de oplossing op Toledo komt

Digital Signal Processing 2: Labo

- **Doel:** Implementatieaspecten van DSP + implementatieproject op TMS320C5515 DSP
- **Docent:** Peter Karsmakers (peter.karsmakers@kuleuven.be)
- **Uurrooster:** 13 x 2u (wekelijks op donderdagvoormiddag na hoorcollege DSP-2)

KU LEUVEN

Digital Signal Processing 2: Examen

- **Examenvorm theorie:**
  - mondeling met schriftelijke voorbereiding
  - gesloten boek (enkel rekentoestel en formularium toegelaten)
  - theorievragen en oefeningen
- **Puntenverdeling:**
  - eindcijfer = gewogen gemiddelde van alle onderwijs- en leeractiviteiten (OLAs)
  - gewichtsfactor = verhouding studiepunten OLA/OPO
    - DSP-1: 34%
    - DSP-1 practicum: 12%
    - DSP-2: 34%
    - DSP-2: practicum: 20%
- **Voorbeeldexamen/Formularium:** zie Toledo

KU LEUVEN

Digital Signal Processing 2: Vakinhoud

- Les 1: Eindige woordlengte
- Les 2: Lineaire predictie
- Les 3: Optimale filtering
- Les 4: Adaptieve filtering
- Les 5: Detectieproblemen
- Les 6: Spectrale signaalanalyse
- Les 7: Schattingsproblemen 1
- Les 8: Schattingsproblemen 2
- Les 9: Sigma-Deltamodulatie
- Les 10: Transformatiecodering

Les 2: Lineaire predictie

- **Parametric signal models**
  non-parametric vs. parametric, AR, ARMA, …

- **Linear prediction**
  prediction error, autocorrelation method, covariance method, …

- **Linear predictive modeling/coding of speech**
  speech production, LP speech model, LP speech coding, …

- **Exercise/homework**

Les 2: Literatuur

- **Parametric signal models**

  B. Porat, *A Course in Digital Signal Processing*

  - Ch. 13, “Analysis and Modeling of Random Signals”
    - Section 13.3, “Rational Parametric Models of Random Signals”

- **Linear prediction**
- **Linear predictive modeling/coding of speech**
- **Exercise/homework**

  T. Dutoit, *Applied Signal Processing*

  - Ch. 1, “How is speech processed in a cell phone conversation?”

KU LEUVEN

Les 2: Lineaire predictie

- **Parametric signal models**
  non-parametric vs. parametric, AR, ARMA, …

- **Linear prediction**
  prediction error, autocorrelation method, covariance method, …

- **Linear predictive modeling/coding of speech**
  speech production, LP speech model, LP speech coding, …

- **Exercise/homework**

Parametric signal models

- Non-parametric vs. parametric signal models
- Linear parametric signal models

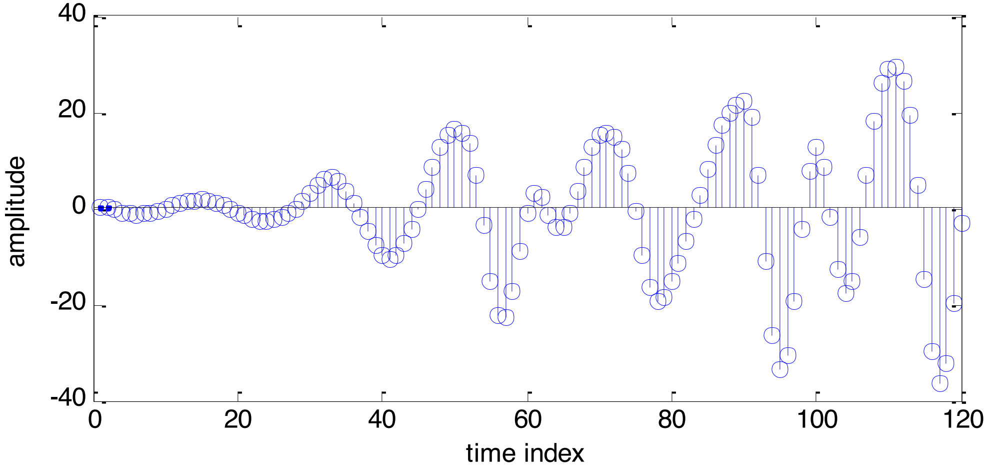
Non-parametric / parametric signal models (1)

- What is a **non-parametric** signal model?
  - non-parametric models are used to represent signals directly by their magnitude values

example 1: time-domain waveform

model parameters = time-domain samples

$$\mathbf{x} = [x(t_1) \; \ldots \; x(t_N)]^T$$



KU LEUVEN

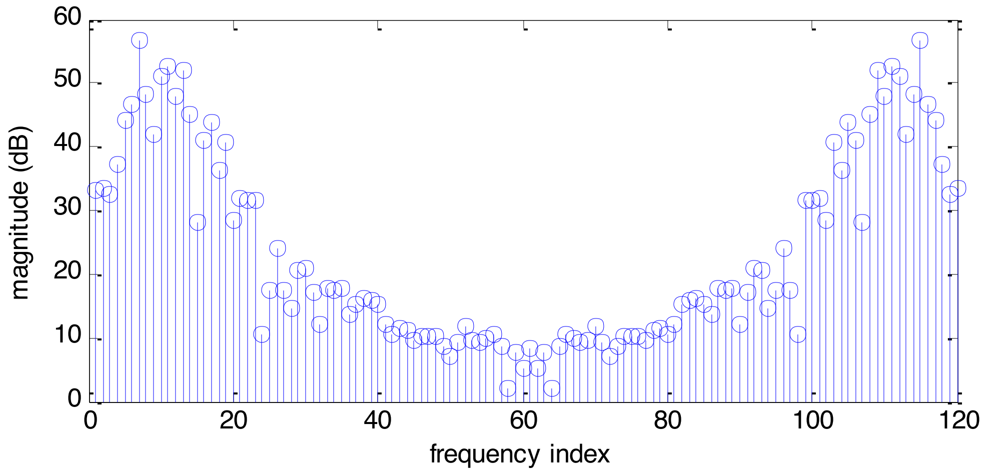
Non-parametric / parametric signal models (2)

- What is a **non-parametric** signal model?
  - non-parametric models are used to represent signals directly by their magnitude values

example 2: frequency magnitude spectrum

model parameters = discrete Fourier transform (DFT) samples

$$\mathbf{x} = [X(\omega_1) \ \ldots \ X(\omega_N)]^T$$



KU LEUVEN

Non-parametric / parametric signal models (3)

- What is a **parametric** signal model?
  - parametric models are used to *approximately* represent signals by a small number of parameters

$$\boldsymbol{\theta} = [\theta_1 \ \ldots \ \theta_P]^T$$

$$\boxed{\dim(\boldsymbol{\theta}) = P \ll N = \dim(\mathbf{x})}$$

Non-parametric / parametric signal models (4)

- Why is a parametric signal model useful?
  - **coding:** represent, store, and transmit signals using relatively small number of parameters

    (e.g., speech, audio, and video compression and streaming)
  - **analysis:** summarize characteristic signal behavior in low-dimensional parameter space

    (e.g., pitch + spectral envelope estimation of speech and audio)
  - **synthesis:** generate synthetic signals from limited number of parameters

    (e.g., music synthesizers, automated speech messages)
  - **whitening:** invertible parametric signal models can be useful for signal whitening

    (e.g., speech and audio signal decorrelation in adaptive filtering)

Parametric signal models

- Non-parametric vs. parametric signal models
- Linear parametric signal models

Linear parametric signal models (1)

- Autoregressive (AR) model:
  - **linear prediction interpretation:** prediction of current signal sample based on past signal samples and excitation signal

$$\boxed{\text{AR: } x(t) = -a_1 x(t-1) - a_2 x(t-2) - \ldots - a_P x(t-P) + e(t)}$$

  - **source-filter interpretation:** modeling signal as output of linear all-pole filter driven by excitation (source) signal

$$\boxed{\text{AR:} \quad x(t) = \frac{1}{A(z)} e(t) \qquad A(z) = 1 + a_1 z^{-1} + \ldots + a_P z^{-P}}$$

KU LEUVEN

Linear parametric signal models (2)

- Autoregressive moving average (ARMA) model:
  - **linear prediction interpretation:** prediction of current signal sample based on past signal samples and moving average of excitation signal

ARMA: $$x(t) = -a_1 x(t-1) - a_2 x(t-2) - \ldots - a_P x(t-P) \\ +b_0 e(t) + b_1 e(t-1) + \ldots b_Q e(t-Q)$$

  - **source-filter interpretation:** modeling signal as output of linear pole-zero filter driven by excitation (source) signal

ARMA: $$x(t) = \frac{B(z)}{A(z)} e(t)$$

$$A(z) = 1 + a_1 z^{-1} + \ldots + a_P z^{-P}$$

$$B(z) = b_0 + b_1 z^{-1} + \ldots + b_Q z^{-Q}$$

KU LEUVEN

Les 2: Lineaire predictie

- **Parametric signal models**
  non-parametric vs. parametric, AR, ARMA, …
- **Linear prediction**
  prediction error, autocorrelation method, covariance method, …
- **Linear predictive modeling/coding of speech**
  speech production, LP speech model, LP speech coding, …
- **Exercise/homework**

Linear prediction

- Linear prediction signal model
- Autocorrelation method
- Covariance method

Linear prediction signal model (1)

- Observed signal:

$$x(t) = -a_1 x(t-1) - a_2 x(t-2) - \ldots - a_P x(t-P) + e(t)$$

- Linear prediction (LP) signal model = AR signal model:

$$\hat{x}(t, \mathbf{a}) = -a_1 x(t-1) - a_2 x(t-2) - \ldots - a_P x(t-P)$$

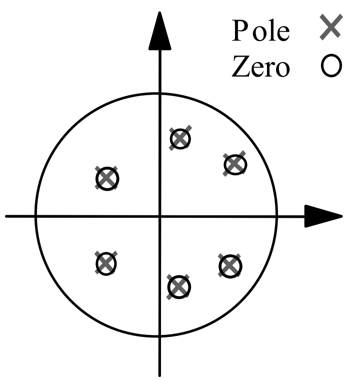
- Goal: estimate model parameters such that predicted model output matches with observed signal in the “best possible way”

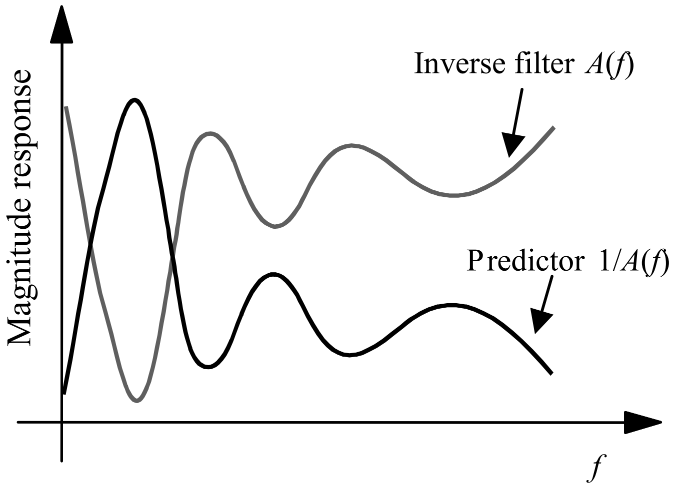
Linear prediction signal model (2)

- Prediction error:

$$
\begin{aligned}
\varepsilon(t,\mathbf{a}) &= x(t) - \hat{x}(t,\mathbf{a}) \\
&= x(t) + a_1 x(t-1) + a_2 x(t-2) + \ldots + a_P x(t-P) \\
&= A(z)x(t)
\end{aligned}
$$

- Prediction error filter (PEF): $A(z) = 1 + a_1 z^{-1} + \ldots + a_P z^{-P}$





KU LEUVEN

Linear prediction

- Linear prediction signal model
- Autocorrelation method
- Covariance method

Autocorrelation method (1)

- Autocorrelation method
  - observed signal:

$$x(t) = -a_1 x(t-1) - \ldots - a_P x(t-P) + e(t)$$

  - left multiply with *x*(*t*)

$$x(t)x(t) = -a_1 x(t)x(t-1) - \ldots - a_P x(t)x(t-P) + x(t)e(t)$$

  - take expectation

$$E\{x(t)x(t)\} = -a_1 E\{x(t)x(t-1)\} - \ldots - a_P E\{x(t)x(t-P)\} + E\{x(t)e(t)\}$$

  - substitute autocorrelation function $r_x(p) \triangleq E\{x(t)x(t-p)\}$

$$r_x(0) = -a_1 r_x(1) - \ldots - a_P r_x(P) + E\{x(t)e(t)\}$$

KU LEUVEN

Autocorrelation method (2)

- Yule-Walker equations:
  - repeat procedure, left multiplying with $x(t-1), \ldots, x(t-P)$
  - force prediction error to be independent: $E\{x(t-p)e(t)\} = 0$

$$\underbrace{\begin{bmatrix} r_x(0) & r_x(1) & \ldots & r_x(P-1) \\ r_x(1) & r_x(0) & \ldots & r_x(P-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(P-1) & r_x(P-2) & \ldots & r_x(0) \end{bmatrix}}_{\mathbf{R_x}} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{bmatrix} = -\underbrace{\begin{bmatrix} r_x(1) \\ r_x(2) \\ \vdots \\ r_x(P) \end{bmatrix}}_{\mathbf{r_x}}$$

- Prediction error variance: $\sigma^2 = \sum_{p=0}^{P} a_p r_x(p) \quad (a_0 \triangleq 1)$

Autocorrelation method (3)

- Naïve solution (matrix inversion): complexity $O(P^3)$

$$\mathbf{a} = \mathbf{R_x}^{-1}\mathbf{r_x}$$

- Efficient solution:
  - exploit properties of $\mathbf{R_x}$ (symmetric and Toeplitz)
  - Levinson-Durbin algorithm = order-recursive algorithm
    - estimate LP model of order *P* = 1
    - for *p* = 2:*P*

      calculate LP model of order P from LP model of order P-1
    - end
  - overall complexity: $O(P^2)$

KU LEUVEN

Linear prediction

- Linear prediction signal model
- Autocorrelation method
- Covariance method

Covariance method (1)

- Covariance method
  - define cost function to minimize prediction error

$$\min_{a_1,\ldots,a_P} \frac{1}{2}\sum_{t=1}^{N} e^2(t) \quad \text{with} \quad e(t) = x(t) + a_1 x(t-1) + \ldots + a_P x(t-P)$$

or $$\min_{\mathbf{a}} \frac{1}{2}\mathbf{e}^T\mathbf{e} \quad \text{with} \quad \mathbf{e} = \mathbf{x} + \mathbf{X}\mathbf{a}$$

$$\underbrace{\begin{bmatrix} e(1) \\ e(2) \\ \vdots \\ e(N) \end{bmatrix}}_{\mathbf{e}} = \underbrace{\begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(N) \end{bmatrix}}_{\mathbf{x}} + \underbrace{\begin{bmatrix} x(0) & x(-1) & \ldots & x(-P+1) \\ x(1) & x(0) & \ldots & x(-P+2) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-1) & x(N-2) & \ldots & x(N-P) \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{bmatrix}}_{\mathbf{a}}$$

Covariance method (2)

- Covariance method
  - minimize function = set derivative to zero (for each a_p!)

$$
\begin{aligned}
\frac{\partial}{\partial \mathbf{a}}\left[\frac{1}{2}\mathbf{e}^T\mathbf{e}\right] = 0 \quad &\Leftrightarrow \quad \frac{\partial}{\partial \mathbf{a}}\left[\frac{1}{2}(\mathbf{x}+\mathbf{X}\mathbf{a})^T(\mathbf{x}+\mathbf{X}\mathbf{a})\right] = 0 \\
&\Leftrightarrow \quad \mathbf{X}^T(\mathbf{x}+\mathbf{X}\mathbf{a}) = 0 \\
&\Leftrightarrow \quad \mathbf{X}^T\mathbf{X}\mathbf{a} = \mathbf{X}^T\mathbf{x} \\
&\Leftrightarrow \quad \boxed{\mathbf{R}_x\mathbf{a} = \mathbf{r}_x}
\end{aligned}
$$

= normal equations

  - $\mathbf{R_x}$ symmetric but not Toeplitz
  - algorithms based on symmetric matrix decomposition

    = *square-root or Cholesky algorithm:* $\frac{1}{6}P^3 + O(P^2)$

Covariance method (3)

- When are both methods **equivalent**?
  - signal = zero outside modeling interval $[1, N]$:

$$x(0) = x(-1) = \ldots = x(-P) = 0$$

  - signal = stationary and ergodic

$$E\{x(t-p)x(t)\} = \sum_t x(t-p)x(t)$$

Covariance method (4)

- Which method to choose?
  - **autocorrelation method:**
    - *Levinson-Durbin algorithm* requires $P^2 + O(P)$ multiplications
    - resulting all-pole filter guaranteed to be stable
    - signal periodicity destroyed by zero padding
  - **covariance method:**
    - *square-root or Cholesky algorithm* (requires $\frac{1}{6}P^3 + O(P^2)$ mult.)
    - resulting all-pole filter not guaranteed to be stable
    - signal periodicity is retained

Les 2: Lineaire predictie

- **Parametric signal models**
  non-parametric vs. parametric, AR, ARMA, …

- **Linear prediction**
  prediction error, autocorrelation method, covariance method, …

- **Linear predictive modeling/coding of speech**
  speech production, LP speech model, LP speech coding, …

- **Exercise/homework**

LP modeling/coding of speech

- Speech production
- LP modeling of speech
- LP coding (LPC) of speech

Speech production (1)

Lungs

Lungs produce air flow (glottal air flow)

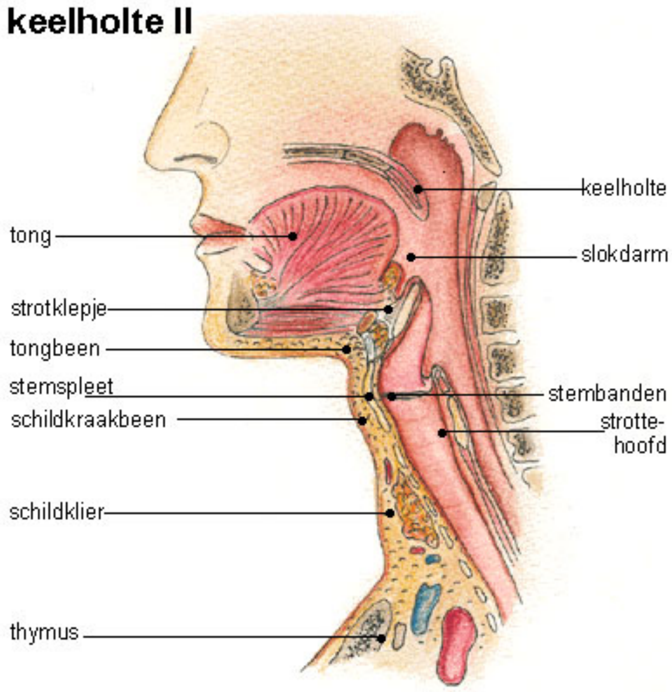
Vocal cords: - vibrate producing pitch (voiced speech)

- don't vibrate (unvoiced speech)

Vocal tract acts as variable filter:

placing of tongue, …

creating "envelope"



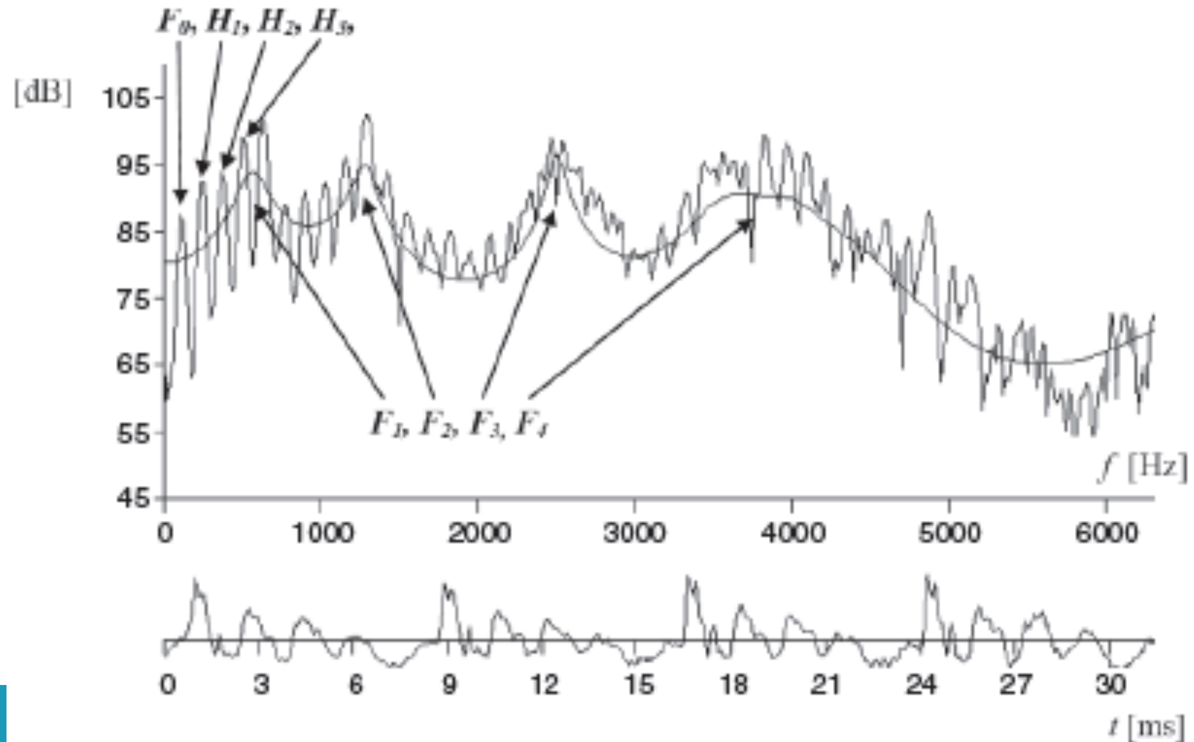
Speech

Speech production (2)

- Spectral speech characteristics:
  - F_0: pitch frequency
  - H_n: *n*th harmonic of pitch frequency
  - F_1 to F_M: formants

By trembling of vocal cords

Filtering by vocal tract



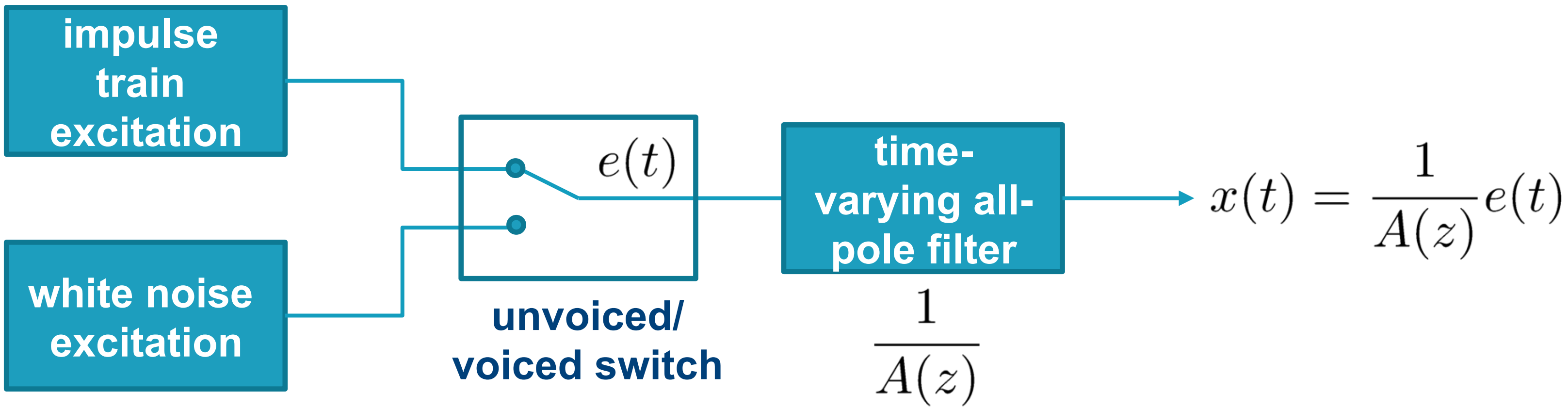
KU LEUVEN

LP modeling/coding of speech

- Speech production
- LP modeling of speech
- LP coding (LPC) of speech

LP modeling of speech (1)

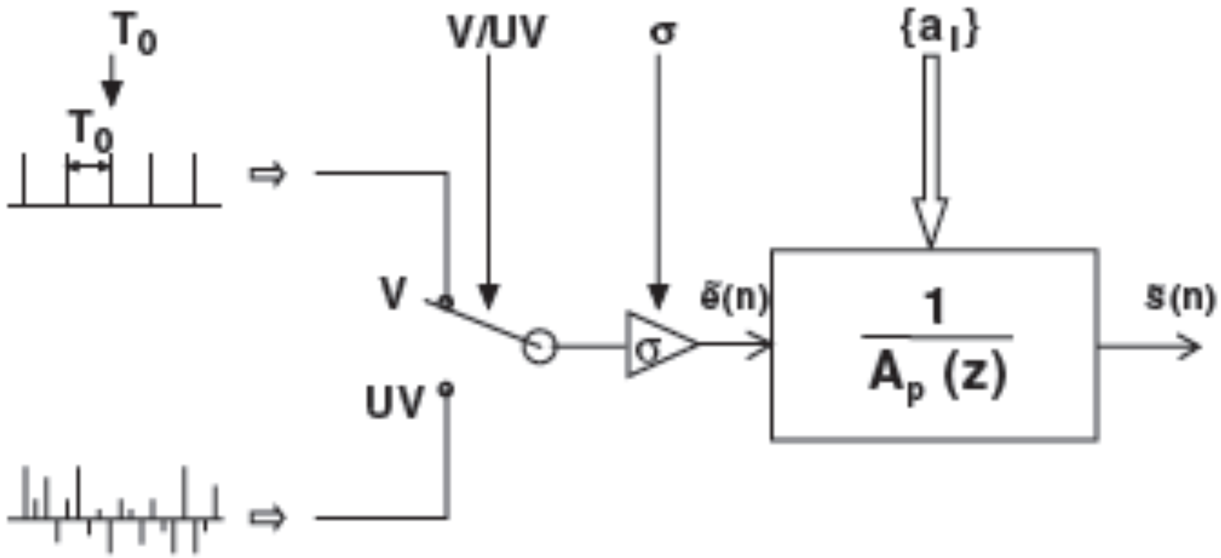
- Correspondence between AR source-filter model and human speech production system:



  - glottal air flow represented as broadband noise signal (white noise excitation for unvoiced speech)
  - vocal cords shape glottal air flow into periodic signal (impulse train excitation for voiced speech)
  - vocal tract behaves as time-varying all-pole filter (spectral shaping filter applied to excitation)

LP modeling of speech (2)

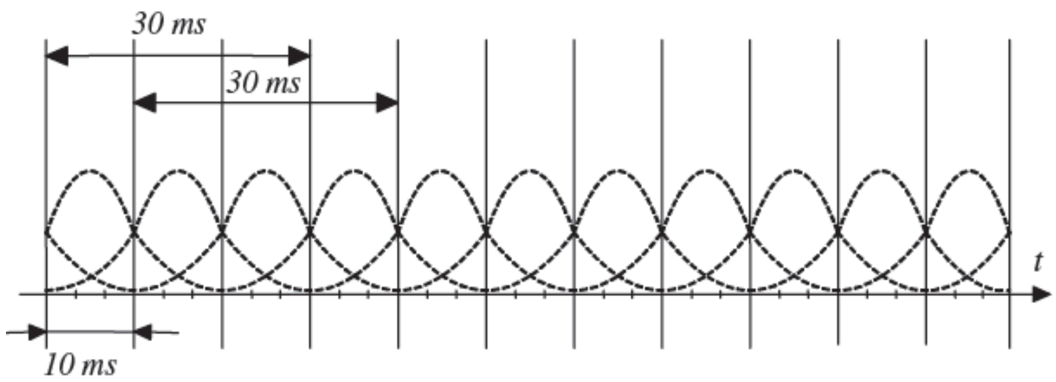
- Different components of LP speech model:
  - Voiced/unvoiced (V/UV) decision
  - Pitch period T_0
  - Prediction error standard deviation σ
  - *P*th order prediction error filter $A_P(z)$



LP modeling of speech (3)

- User choices:
  - **length of speech signal segment** (*N*): compromise between
    - accurate estimation of autocorrelation function: $N \nearrow$
    - speech stationarity throughout signal segment: $N \searrow$

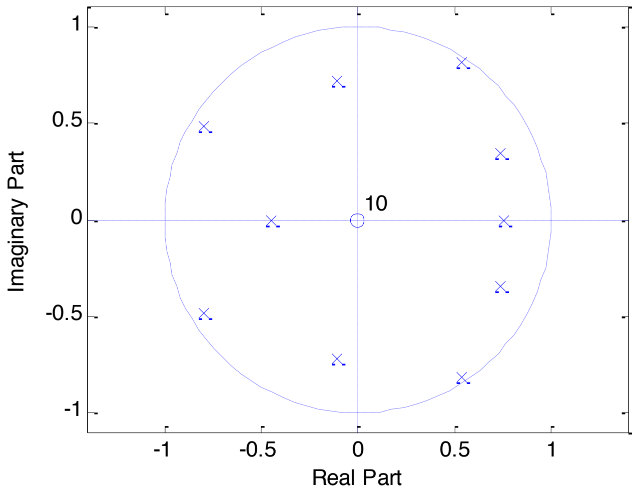
    rule of thumb: N ~ 30 ms (e.g., $N = 240$ at $f_s = 8$ kHz)
  - split signal in <u>overlapping</u> segments of length *N*
  - apply <u>window</u> to each segment (e.g., Hann)
  - calculate the LP coefficients for each frame

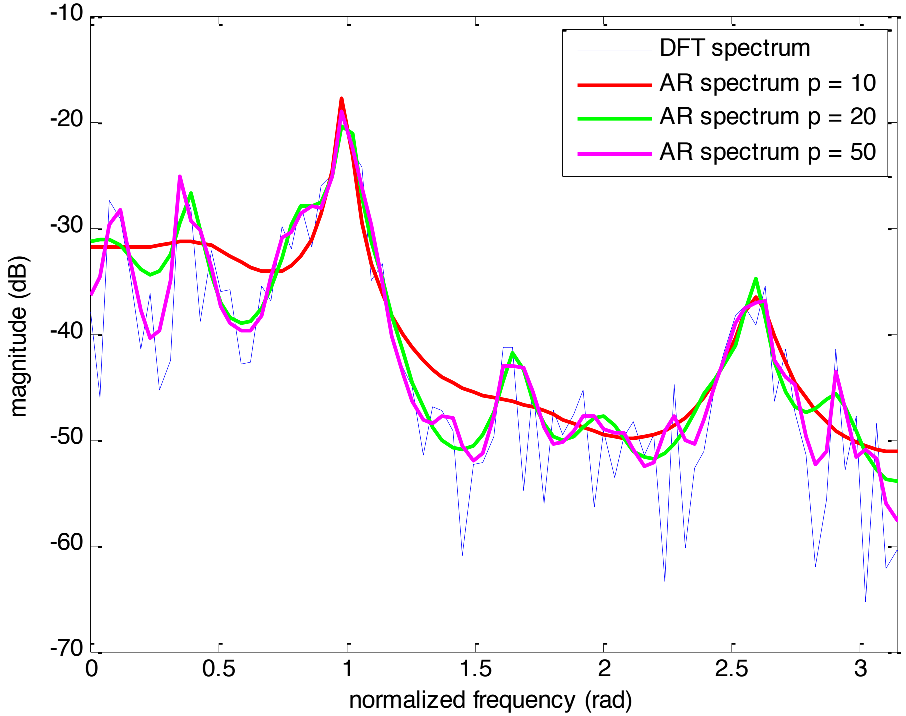


LP modeling of speech (4)

- User choices:
  - **model order** (*P*) = compromise between
    - model accuracy: $P \nearrow$
    - model complexity: $P \searrow$

    rule of thumb: $P = 10 - 20$
  - speech example:





LP modeling of speech (5)

- Pitch prediction:
  - periodicity of voiced speech (originating at vocal cords) cannot be modeled using (low-order) AR model

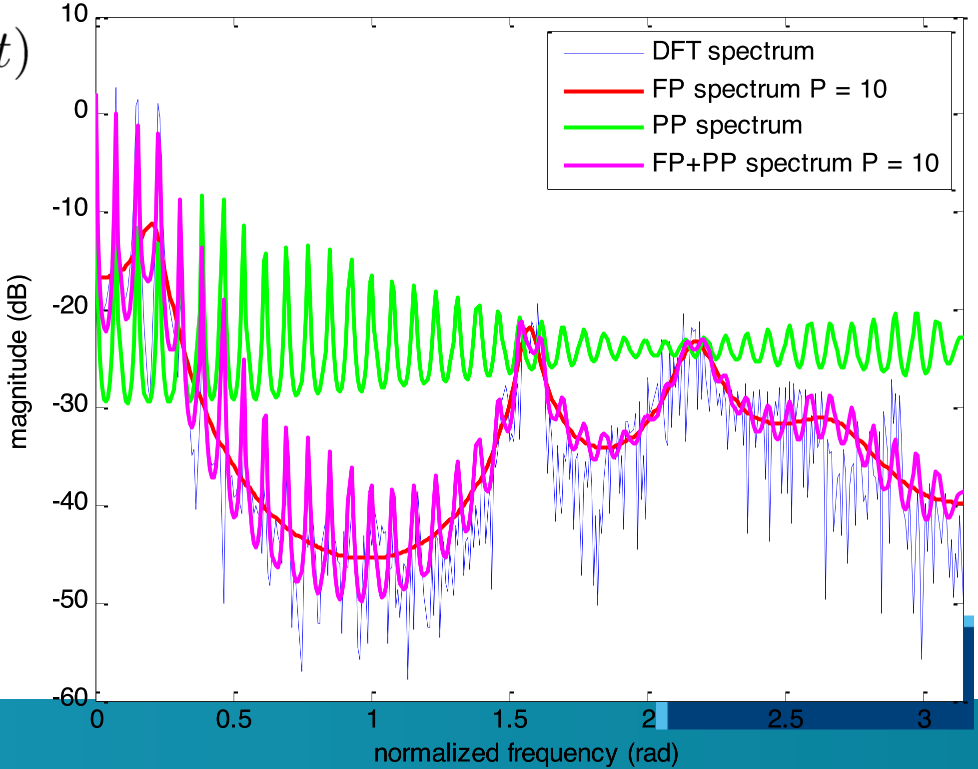
    (AR model only represents signal autocorrelation up to lag *P*)
  - cascade of two AR models (formant predictor + pitch predictor)

$$x(t) = \frac{1}{A_{\mathrm{FP}}(z)A_{\mathrm{PP}}(z)}e(t)$$

$$\boxed{A_{\mathrm{PP}}(z) = 1 - z^{-K}} \quad K = T_0 f_s$$

  - pitch lag estimation:
    - scalar YW equation
    - exhaustive search

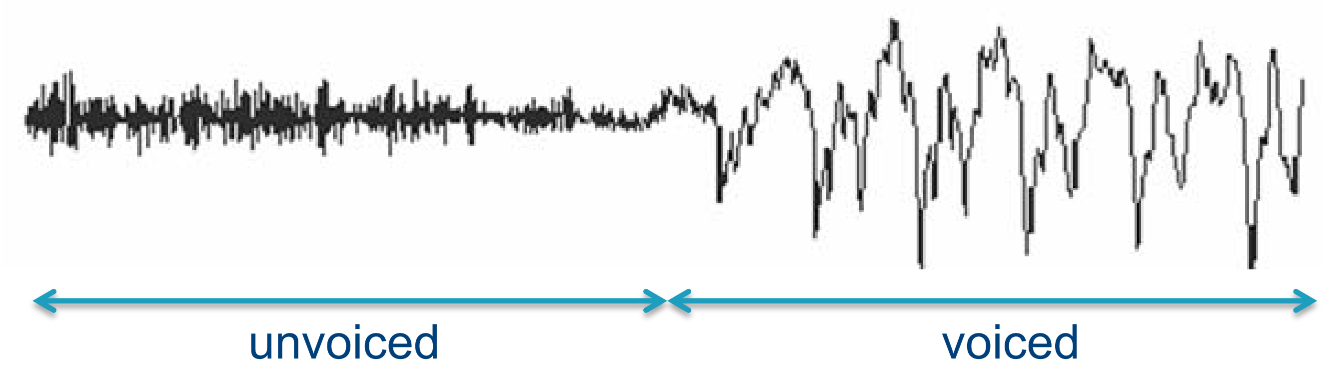
      to find optimal K
  - pitch lag: $20 \leq K \leq 160$
  - comb filter behavior



LP modeling of speech (6)

- Voiced/unvoiced (V/UV) decision
  - detection or binary classification problem

    (cf. Les 5: Detectieproblemen)
  - based on signal features such as:
    - zero crossing rate
    - short-term power
    - spectral flatness of LP residual
    - …



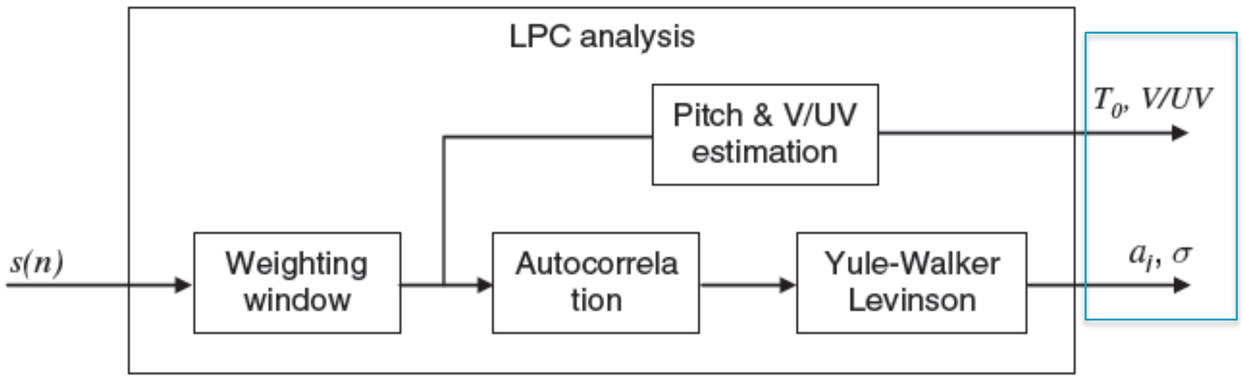
KU LEUVEN

LP modeling/coding of speech

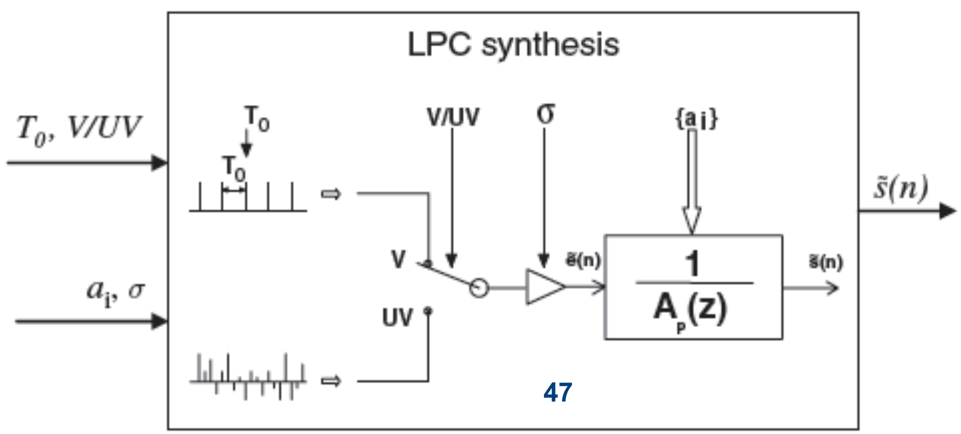
- Speech production
- LP modeling of speech
- LP coding (LPC) of speech

LP coding (LPC) of speech (1)

- LP analysis at transmitter side (coding):

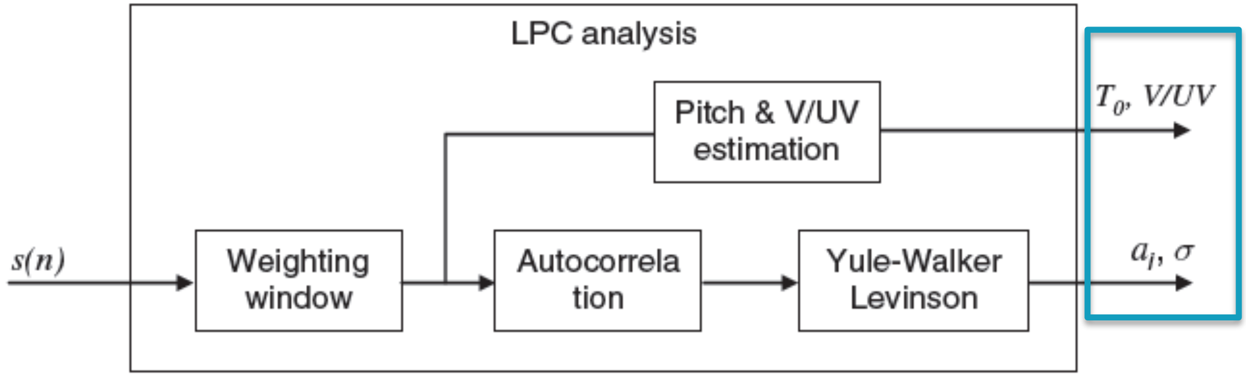


- LP synthesis at receiver side (decoding):



LP coding (LPC) of speech (2)

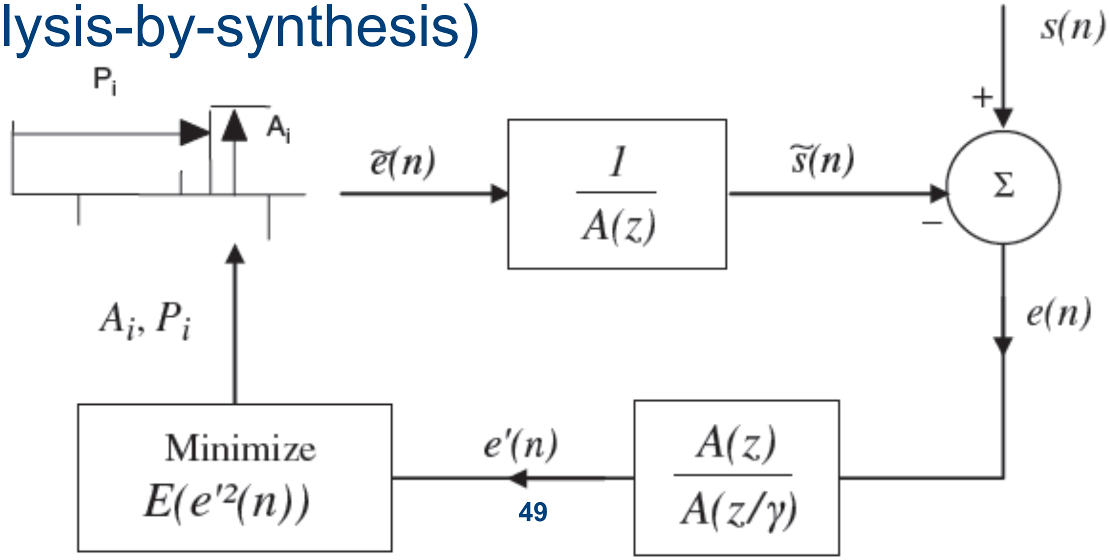
- LP analysis at transmitter side (coding):



  - LP speech model parameters T_0, *VU/V*, a_i, σ need to be quantized = represented by finite number of bits
  - results in certain bit rate for speech codec, e.g.
    - NATO LPC10 codec: 2400 bit/s (satellite telephony)
    - EFR codec: 11.2 kbit/s (GSM telephony)
    - …

LP coding (LPC) of speech (3)

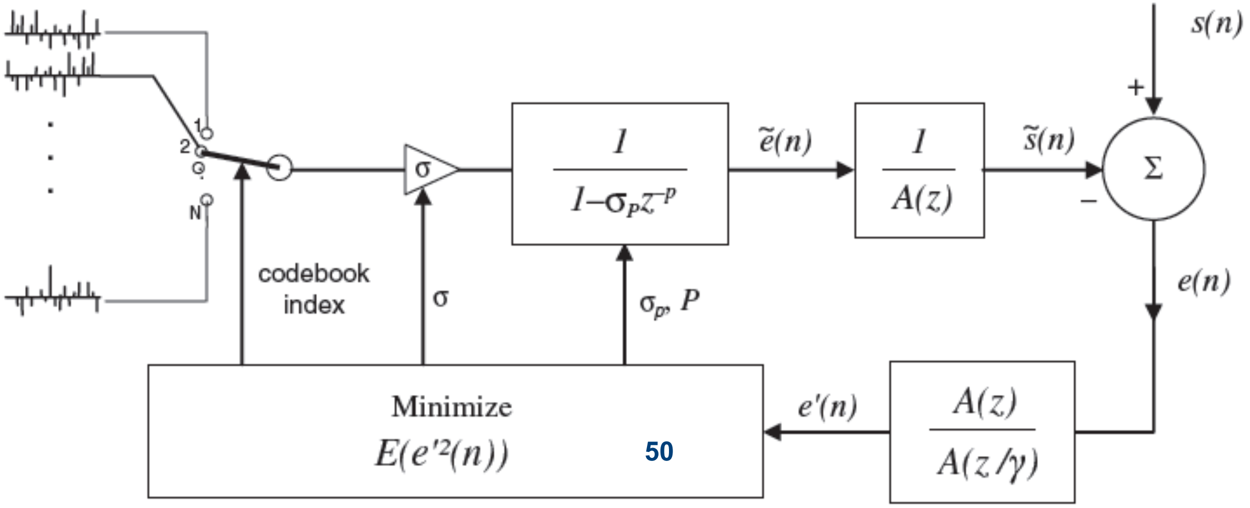
- Problem:
  - LP analysis method only estimates LP parameters, not excitation signal
- Solution = Multipulse Excited (MPE) LP
  - represent excitation signal by limited number pulses/frame
  - estimate pulse positions & amplitudes using feedback loop (= analysis-by-synthesis)



KU LEUVEN

LP coding (LPC) of speech (4)

- Problem:
  - the real excitation is far more complex than what V/UV detector & pitch frequency can represent
- Solution: Code Excited LP (CELP)
  - use real excitations out of a database (or codebook)
  - analysis-by-synthesis to find best code & gain (vector quantization)



KU LEUVEN

Exercise / Homework

- **MATLAB exercise**:
  speech analysis & synthesis with linear prediction

  T. Dutoit, *Applied Signal Processing*
  - Ch. 1, “How is speech processed in a cell phone conversation?”
    - **Section 1.2.1 – 1.2.6**
    - **see Matlab script ASP_cell_phone.m on Toledo !**