Digital Audio Signal Processing

DASP

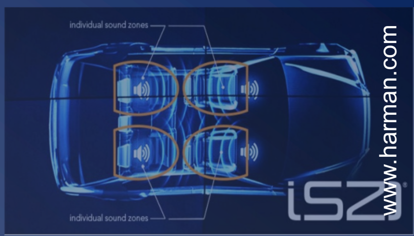
Chapter-6:
Sound Field Control

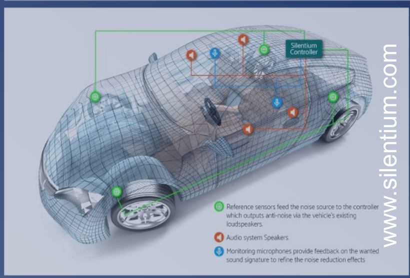
Jesper Brunnström - Marc Moonen
Dept. E.E./ESAT-STADIUS, KU Leuven
marc.moonen@kuleuven.be
homes.esat.kuleuven.be/~moonen/

1

Sound field control - Outline

- **Basics**
  - Signal Model
- **Sound zone control**
  - Pressure Matching
  - Acoustic Contrast Control
- **Active noise control**
  - Feedforward ANC
    Filtered-x LMS
  - Feedback ANC

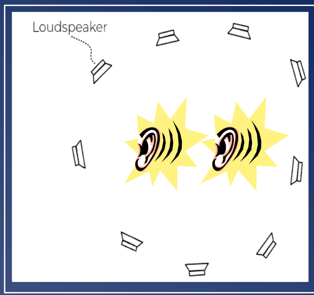




2

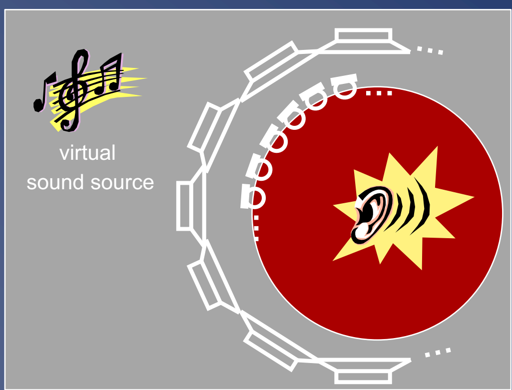
Basics

- Design loudspeaker signal(s) from reference signal(s), to control sound at some point(s) in a room/space
- Reference signal(s) either given or recorded (microphone(s))
- Requires knowledge of acoustic impulses responses from each loudspeaker to each control point (microphone)
  - ➔ Acoustic impulse responses may be measured ahead of time ('calibration')
  - ➔ Calibration can be a difficult in practice, time consuming, or problematic if system is time-varying

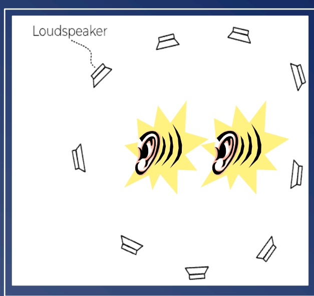


Example: Binaural synthesis

Basics



Example: Sound field synthesis based on Huygens' principle



Example: Binaural synthesis

Basics - Signal model

- Sound pressure at a control point resulting from a single loudspeaker is..

$$\underbrace{p(n) = h}_{\text{Acoustic path}} * \underbrace{y(n)}_{\text{Loudspeaker signal}}$$

Acoustic path is generally modelled as a linear finite impulse response filter

- Loudspeaker signal is reference/input signal filtered by control filter

$$y(n) = \underbrace{w}_{\text{Control filter}} * \underbrace{x(n)}_{\text{Reference signal}}$$

5

Basics - Signal model

- Matrix form: **w** and **h** commute if time-invariant

$$\mathbf{w} = \begin{bmatrix} w(0) & \cdots & w(I-1) \end{bmatrix}^T$$

$$\mathbf{h} = \begin{bmatrix} h(0) & \cdots & h(J-1) \end{bmatrix}^T$$

$$\mathbf{x}(n) = \begin{bmatrix} x(n) & \cdots & x(n-I-J+1) \end{bmatrix}^T$$

$$\mathbf{W} = \begin{bmatrix} w(0) & 0 & 0 & \dots & 0 \\ w(1) & w(0) & 0 & \dots & 0 \\ w(2) & w(1) & w(0) & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & w(I-1) \end{bmatrix}$$

$$p(n) = \mathbf{x}(n)^T.\mathbf{W}.\mathbf{h}$$

$$\mathbf{w} = \begin{bmatrix} w(0) & \cdots & w(I-1) \end{bmatrix}^T$$

$$\mathbf{h} = \begin{bmatrix} h(0) & \cdots & h(J-1) \end{bmatrix}^T$$

$$\mathbf{x}(n) = \begin{bmatrix} x(n) & \cdots & x(n-I-J+1) \end{bmatrix}^T$$

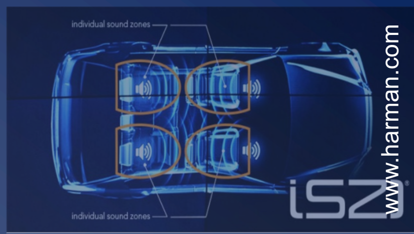
$$\mathbf{H} = \begin{bmatrix} h(0) & 0 & 0 & \dots & 0 \\ h(1) & h(0) & 0 & \dots & 0 \\ h(2) & h(1) & h(0) & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & h(J-1) \end{bmatrix}$$

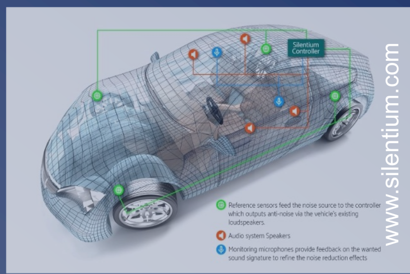
$$p(n) = \mathbf{x}(n)^T.\mathbf{H}.\mathbf{w}$$ = linear function of **w**

6

Sound field control - Outline

- **Basics**
  - Signal Model
- **Sound zone control**
  - Pressure Matching
  - Acoustic Contrast Control
- **Active noise control**
  - Feedforward ANC
    Filtered-x LMS
  - Feedback ANC





7

Sound zone control – Problem statement

- Reproduce different sounds for different listeners in the same room/space, with minimal interference

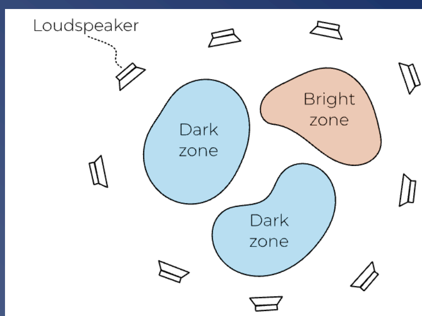
  Can be viewed as 'transmit beamforming' (vs. 'receive beamforming in Chapter 3-4)

- **Solve with superposition**

  Hence can consider problem with only a single input/reference signal, a single 'bright zone', all other zones are 'dark zones'

  Bright zone: Reproduce input signal
  Dark zone: Quiet
  Other spaces: Any sound

  Acoustic impulse responses from all loudspeakers to a dense set of control points (microphones) in all zones assumed given



8

Sound zone control – Signal model

- One input signal x(n) (cfr. superposition)
- One control filter per loudspeaker
- One acoustic path from each loudspeaker to each control point (microphone)

$$p_m(n) = \sum_{l=1}^{L} \boldsymbol{x}^\top(n)\boldsymbol{H}_{ml}\,\boldsymbol{w}_l$$

Sound pressure at *m*-th control point (microphone)

Control filter for *l*-th loudspeaker

Acoustic impulse response from *l*-th loudspeaker to *m*-th control point (microphone)

9

Sound zone control – Signal model

- Matrix form:

$$\boldsymbol{w} = \begin{bmatrix}\boldsymbol{w}_1^\top & \dots & \boldsymbol{w}_L^\top\end{bmatrix}^\top$$

Stack control filters for all loudspeakers

$$\boldsymbol{H}_b = \begin{bmatrix}\boldsymbol{H}_{11} & \dots & \boldsymbol{H}_{1L} \\ \vdots & \ddots & \vdots \\ \boldsymbol{H}_{M_b1} & \dots & \boldsymbol{H}_{M_bL}\end{bmatrix}$$

Acoustic paths from all (L) loudspeakers to all (M_b) microphones in bright zone

$$\boldsymbol{X}_b = \underbrace{\begin{bmatrix}\boldsymbol{x}(n) & & \\ & \ddots & \\ & & \boldsymbol{x}(n)\end{bmatrix}}_{M_b \text{ times}}$$

Block diagonal matrix with duplicated signal vectors

$$\boxed{\boldsymbol{p}_b(n) = \boldsymbol{X}_b^\top(n)\boldsymbol{H}_b\boldsymbol{w}}$$

Sound pressure in M_b bright zone microphones

$$\boxed{\boldsymbol{p}_d(n) = \boldsymbol{X}_d^\top(n)\boldsymbol{H}_d\boldsymbol{w}}$$

Similarly:
Sound pressure in M_d dark zone microphones

10

Sound zone control - Pressure Matching

- Minimize the mean squared error (MSE) between the desired and reproduced signal

$$\mathcal{C} = \mathbb{E}[\|\boldsymbol{p}_b - \boldsymbol{d}(n)\|^2 + \|\boldsymbol{p}_d(n)\|^2]$$

Desired signal in bright zone is usually input signal as reproduced by a virtual source

Desired signal in dark zones is 0

$$\boldsymbol{d}(n) = \tilde{\boldsymbol{h}} * x(n)$$

Optimal filters computed by setting gradient to zero...

Sound zone control - Pressure Matching

- MSE-optimal control filters are

$$\boldsymbol{w} = (\boldsymbol{R}_b + \boldsymbol{R}_d)^{-1}\boldsymbol{r}_b$$

with

$$\boldsymbol{R}_b = \mathbb{E}[\boldsymbol{H}_b^\top \boldsymbol{X}_b(n)\boldsymbol{X}_b^\top(n)\boldsymbol{H}_b]$$

$$\boldsymbol{R}_d = \mathbb{E}[\boldsymbol{H}_d^\top \boldsymbol{X}_d(n)\boldsymbol{X}_d^\top(n)\boldsymbol{H}_d]$$

$$\boldsymbol{r}_b = \mathbb{E}[\boldsymbol{H}_b^\top \boldsymbol{X}_b(n)\boldsymbol{d}(n)]$$

Correlation matrices are defined in terms of known signal and acoustic impulse responses, hence can be computed ahead of operation

Sound zone control - Pressure Matching

- Same principles as for GEVD-based multichannel Wiener filter can be applied

  Simultaneous diagonalization of matrix pair $\{R_b, R_d\}$ is

$$\boldsymbol{U}^\top \boldsymbol{R}_b \boldsymbol{U} = \boldsymbol{\Lambda}$$
$$\boldsymbol{U}^\top \boldsymbol{R}_d \boldsymbol{U} = \boldsymbol{I}$$

then

$$\boldsymbol{w} = (\boldsymbol{R}_b + \boldsymbol{R}_d)^{-1} \boldsymbol{r}_b$$
$$\boldsymbol{w} = \boldsymbol{U}(\boldsymbol{\Lambda} + \boldsymbol{I})^{-1} \boldsymbol{U}^\top \boldsymbol{r}_b$$

Skip this slide

T. Lee, J. K. Nielsen, J. R. Jensen, and M. G. Christensen, 'A unified approach to generating sound zones using variable span linear filters', *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Canada, Apr. 2018, pp. 491–495.

13

Sound zone control – Acoustic Contrast Control

- Approximate with the R largest eigenvalues/eigenvectors

$$\boldsymbol{w} = \sum_{r=1}^{R} \frac{\boldsymbol{u}_r^\top \boldsymbol{r}_b}{\lambda_r + 1} \boldsymbol{u}_r$$

r-th column of U

r-th diagonal element of Λ

Trade-off...

- Full rank gives lowest error in the bright & dark zone (cfr. MSE)
- Rank-1 gives highest 'acoustic contrast', i.e. ratio of signal power in bright zone to signal power in dark zones (no proof)

  = Acoustic Contrast Control (ACC)

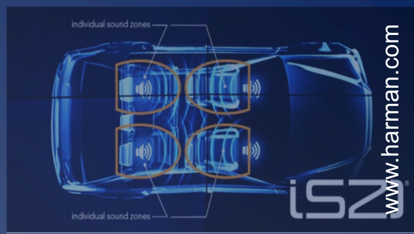
Skip this slide

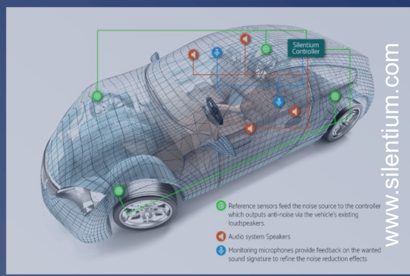
T. Lee, J. K. Nielsen, J. R. Jensen, and M. G. Christensen, 'A unified approach to generating sound zones using variable span linear filters', *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Canada, Apr. 2018, pp. 491–495.

14

Sound field control - Outline

- **Basics**
  - Signal Model
- **Sound zone control**
  - Pressure Matching
  - Acoustic Contrast Control
- **Active noise control**
  - Feedforward ANC
    Filtered-x LMS
  - Feedback ANC





Active Noise Control - Intro

- **Passive** noise control : sound absorbers, …
  works well only for high frequencies (`centimeter-waves’)
- **Active** noise control : for low frequencies (e.g. 100 Hz ➔ λ=3,4m)
  - General set-up

Aim: generate `quiet’ at error microphone

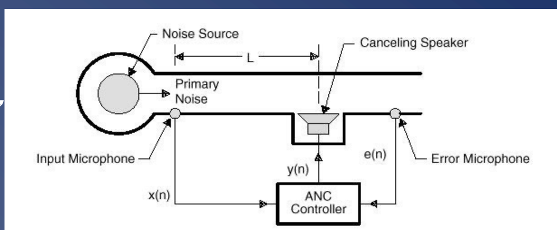


Fig 2. Single-Channel Broadband Feedforward ANC System in a Duct

ANC works on the principle of ***destructive interference*** between the sound field generated by the `primary’ (noise) source and the sound field due to secondary source(s), whose output can be controlled

Active Noise Control - Intro

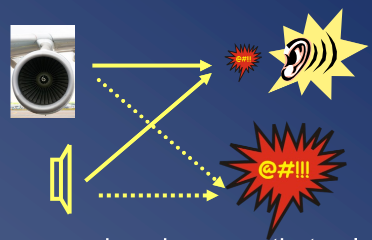
PS: Destructive interference relies on superposition & <u>linearity</u>

At higher volume, loudspeaker can exhibit non-linear behaviour

After destructive interference at main frequency, harmonics generated by loudspeaker may become distinctly audible

Non-linearity not considered here

PS: Destructive interference at one point, may imply <u>constructive interference</u> at other point

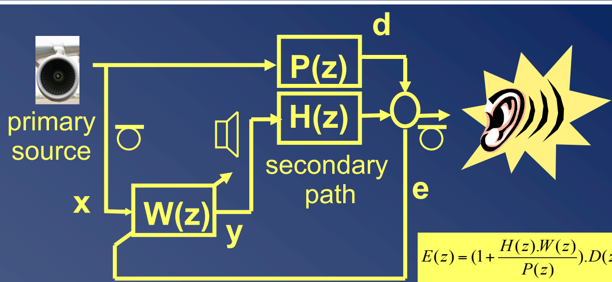


Secondary source to be placed close to error microphone, so that only modest secondary signal is required, and hence points further away from secondary source are not affected. Produce `zone of quiet' near the error microphone (e.g. 10dB reduction in zone approx λ/10)

17

Feedforward ANC

- **Basic set-up**



$$E(z) = (1 + \frac{H(z).W(z)}{P(z)}).D(z)$$

– P(z) = primary path = unknown

– H(z) = secondary path = acoustic path from loudspeaker (=secondary source) to error microphone.
H(z) can be modeled/identified (=calibration)

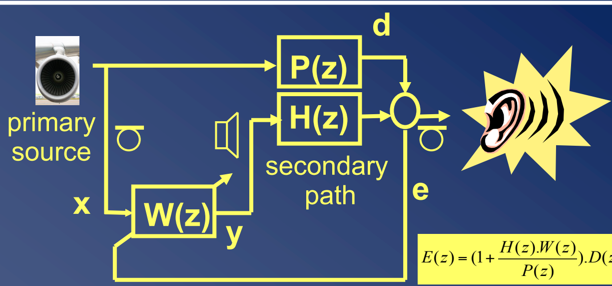
– Secondary signal at error microphone is ideally equal to –d(n)

– However (unlike in sound zoning):
d(n) is not known/observed, only e(n) is observed…

18

Feedforward ANC

- **Basic set-up**
  (continued)



$$E(z) = (1 + \frac{H(z).W(z)}{P(z)}).D(z)$$

  - However (unlike in sound zoning):
    d(n) is not known/observed, only e(n) is observed…
  - Hence need **<u>adaptive filter</u>** controlled by e(n)
  - Design adaptive filter W(z) to minimize MSE
    (compare to AEC!)

Feedforward ANC

- **Signal model**

$$e(n) = p(n) + d(n)$$

Error microphone signal | Secondary signal | Primary (unwanted) signal

Reference signal is input to the control filter
(must be correlated with the primary signal)

$$p(n) = h * w * x(n) = \mathbf{x}(n)^T.\mathbf{H}.\mathbf{w}$$

Secondary path | Control filter | Reference signal

See page 5
= correct if w & h are time-invariant
= approximation if w is (slowly) time-varying (cfr infra)

Feedforward ANC

- **Minimize MSE cost function**

$$\mathcal{C} = \mathbb{E}[e(n)^2]$$

$$e(n) = \mathbf{x}(n)^T.\mathbf{H}.\mathbf{w} + d(n)$$

$$\nabla\mathcal{C} = 2\mathbb{E}[\boldsymbol{H}^\top \boldsymbol{x}(n)e(n)]$$

Correlation between the error signal e(n)
and the reference signal x(n) filtered through the secondary path ***h,*** i.e.

$$\boldsymbol{x}^f(n) = \boldsymbol{H}^\top \boldsymbol{x}(n)$$

Feedforward ANC – Filtered-x LMS

- **Gradient descent adaptive filter**

Gradient is the direction in which the function increases fastest
Change the control filter in small steps in the opposite direction

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \mu\nabla\mathcal{C}$$

Step size

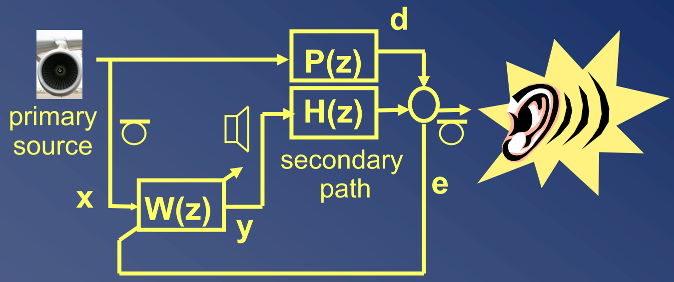
- **Stochastic gradient adaptive filter**

uses instantaneous estimate of gradient (= LMS (Least Mean Squares))

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \mu\boldsymbol{H}^\top \boldsymbol{x}(n)e(n)$$

This is referred to as 'Filtered-x LMS'
..as the input signal x(n) is first filtered by the secondary path ***h***

PS: Interpretation



– Straightforward application of LMS :

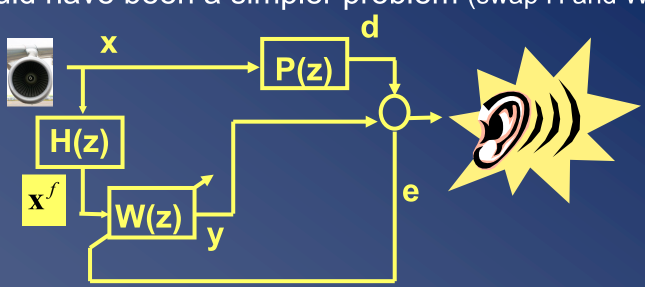
$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu.\mathbf{x}(n).e(n)$$

…does not work here

(ex.: H(z)=-1, then steepest descent turned into steepest ascent)

PS: Interpretation

– This would have been a simpler problem (swap H and W)…



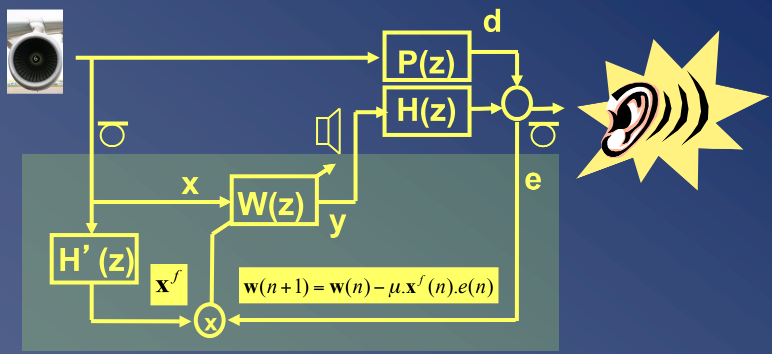
...allowing for straightforward application of LMS, with filtered x-signal

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu.\mathbf{x}^f(n).e(n)$$

– PS: Again, only <u>time-invariant</u> linear systems commute, hence (implicit) swapping will require <u>slow</u> adaptation of W(z)

PS: Interpretation

– Filtered-X LMS scheme : swapping of H and W only in adaptation path (not in filtering path)



...with in practice H' (z) an estimate of H(z)

Feedforward ANC – Filtered-x LMS

– Convergence is dependent on secondary path. Step size bounds assuming white reference signal [1]

$$0 < \mu < \frac{2}{(I + 2\Delta_{eq})P_{x^f}}$$

Power of filtered reference signal

Length of control filter *Equivalent delay* of secondary path

hence (long) H(z) implies smaller μ & slower convergence

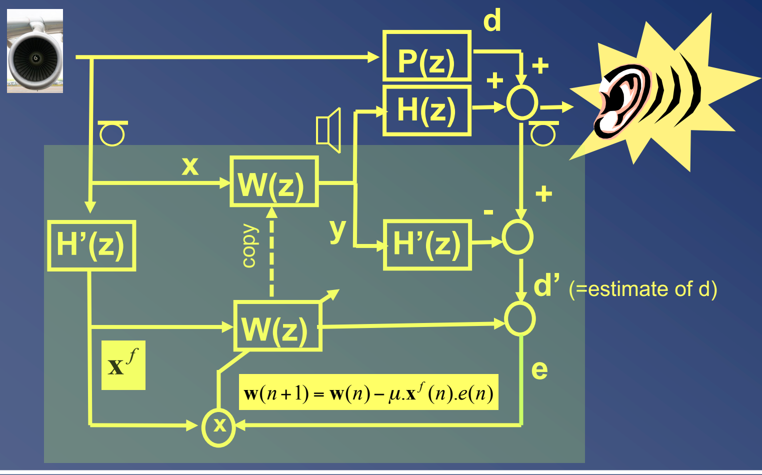
– Stability also affected by the accuracy of secondary path estimate H' (z). Filtered-X LMS found to be robust to errors in H' (z)... (details omitted)

– Note that adaptation in "first-H'(z)-then-W(z)" (as in p.24) relies on error signal (e) from "first-W(z)-then-H(z)", which is inconsistent...

[1] I. Tabatabaei Ardekani and W. H. Abdulla, 'Theoretical convergence analysis of FxLMS algorithm', *Signal Processing*, vol. 90, no. 12, pp. 3046–3055, Dec. 2010.

Feedforward ANC – Filtered-x LMS

- **Improved version is 'Modified Filtered-x LMS'**



Feedforward ANC – Filtered-x LMS

- **Improved version is 'Modified Filtered-x LMS'**
  (continued)

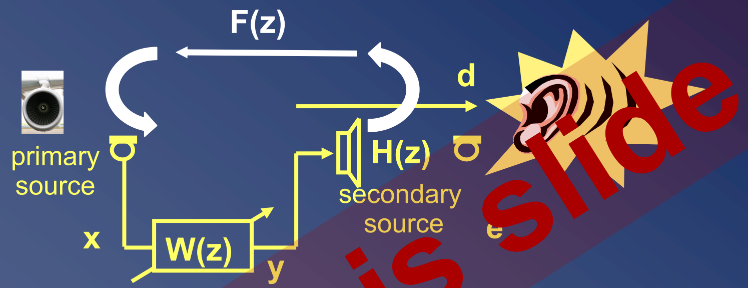
  Note that (if H'(z)=H(z), hence d'=d) adaptation in "first-H'(z)-then-W(z)" now relies on error signal (e) from "first-H'(z)-then-W(z)", which is consistent

  Hence improved convergence (can use larger step size)

  Hence can also replace LMS by RLS (='Modified Filtered-x RLS')

Feedforward ANC

- Additional problem: Feedback (previously assumed to be negligible)



Feedback from secondary source (loudspeaker) into ref. microphone
This is an acoustic echo cancellation/feedback problem :

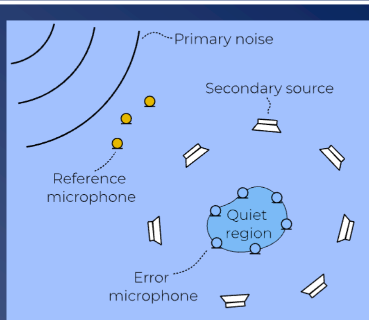
- Fixed AFC based on given F(z) (obtained through calibration) is easy
- Adaptive AFC is problematic (combination of 2 adaptive systems)

Skip this slide

29

Feedforward ANC - MIMO

- **Extension:**
  Multiple reference signals
  Multiple secondary sources
  Multiple error signals



(a.k.a. 'MIMO ANC', multiple-input multiple-output ANC)

- Applications: airplane/car cabin noise control, active vibration control,...
- Needs generalization of Filtered-x algorithms, where coefficients of control filters are adapted to minimize the sum of the mean square values of all the error signals.

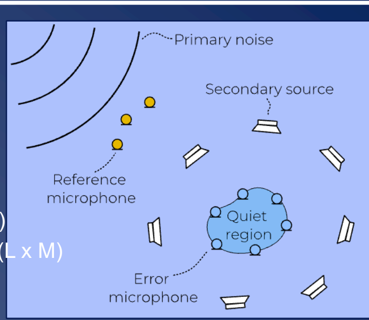
30

Feedforward ANC - MIMO

Loudspeaker signal is sum of contributions from all reference signals r

$$y_l(n) = \sum_{r=1}^{R} \boldsymbol{w}_{rl}^{\top} \boldsymbol{x}_r(n)$$

One control filter per ref.micr.-loudspeaker pair (R x L)
One secondary path per loudspeaker-error micr. pair (L x M)



R reference microphones
L loudspeakers
M error microphones

Signal at microphone m is...

$$e_m(n) = p_m(n) + d_m(n)$$

$$p_m(n) = \sum_{l=1}^{L} \sum_{r=1}^{R} \boldsymbol{x}_r^{\top}(n) \boldsymbol{H}_{ml} \boldsymbol{w}_{rl}$$

Feedforward ANC - MIMO

Minimize MSE (summed over error microphones)

$$\mathcal{C} = \mathbb{E}\left[\sum_{m=1}^{M} e_m(n)^2\right]$$

Leads to MIMO Fx-LMS

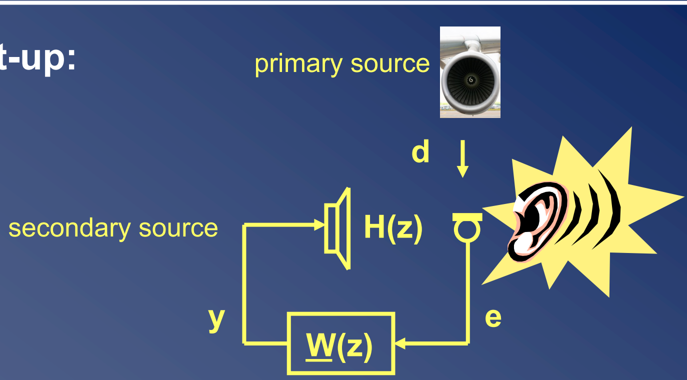
$$\boldsymbol{w}_{rl}(n+1) = \boldsymbol{w}_{rl}(n) - \mu \sum_{m=1}^{M} \boldsymbol{H}_{lm}^{\top} \boldsymbol{x}_r(n) e_m(n)$$

Every reference signal is filtered by every secondary path
i.e. number of filtering operations is R x L x M (!)

If secondary paths are high-order FIR, leads to (extremely) high computational complexity

Feedback ANC

Basic set-up:

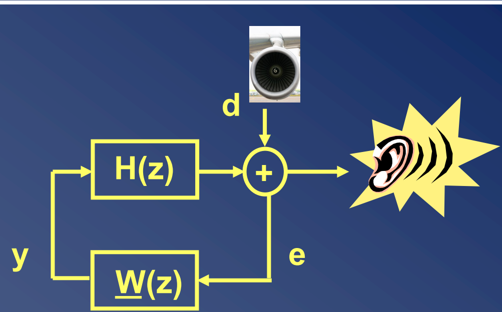


– H(z) = secondary path
– 1 microphone instead of 2 microphones
– Applications : active headsets, ear defenders
  (e.g.10-15dB reduction achieved for frequencies 30-500Hz)

33

Feedback ANC

Design problem:



- If H(z) is given design W(z) (=feedback control) such that E(z) is `minimized'

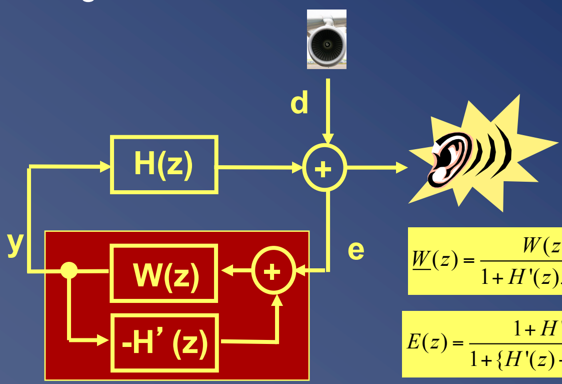
$$E(z)=\frac{1}{1-H(z).\underline{W}(z)}.D(z)$$

- For `flat' H(z)=C^{nt} : W(z)=-A for large A (as in an opamp)
- For general H(z) : see control courses

34

Feedback ANC

An interesting feedback controller is formed as follows :



$$\underline{W}(z) = \frac{W(z)}{1 + H'(z).W(z)}$$

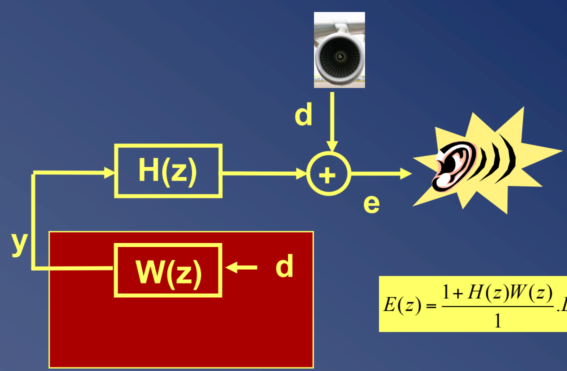
$$E(z) = \frac{1 + H'(z)W(z)}{1 + \{H'(z) - H(z)\}.W(z)}.D(z)$$

…with H’ (z) an estimate of H(z) and W(z) yet to be defined.

Note that if H’ (z)=H(z), then W(z) is fed by d(n), i.e. …

Feedback ANC

Note that if H’ (z)=H(z), then W(z) is fed by d(n), i.e. …

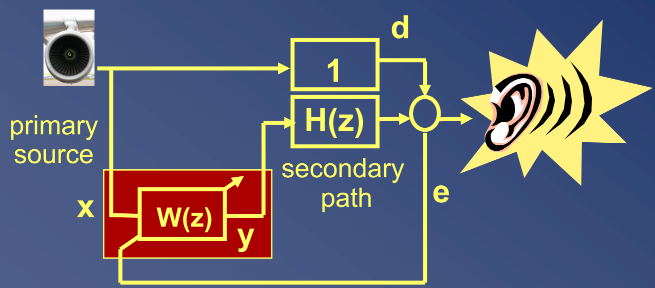


$$E(z) = \frac{1 + H(z)W(z)}{1}.D(z)$$

…which means the feedback system has been transformed into a <u>feedforward</u> system,..

Feedback ANC

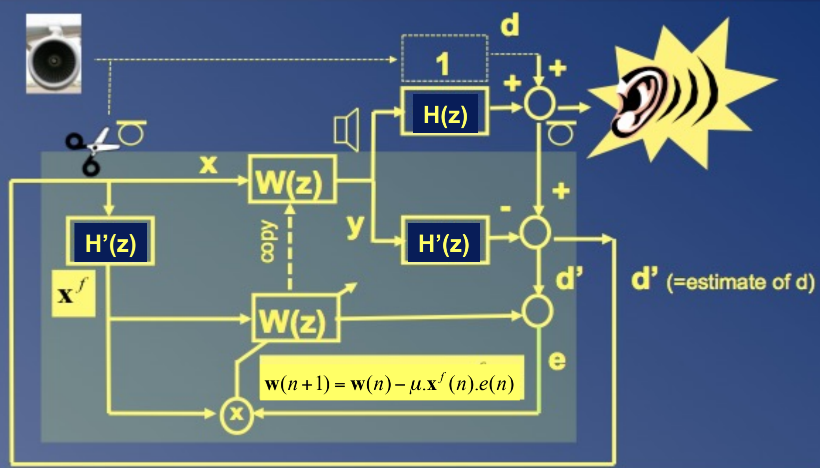
In the set-up of p.18, this is …



– With P(z) =1, and for instance for H(z) containing pure delay, this means W(z) must act as a predictor for d
– Adaptation of W(z) based on (modified) filtered-X algorithm

Feedback ANC

With H’(z)≠H(z) and modified Fx-LMS algorithm, this is...



Feedback ANC

With H' (z)≠H(z) and modified Fx-LMS algorithm, this is...

Note that the error microphone signal (to the right) is used to replace the (non-existing) noise reference microphone signal (to the left)

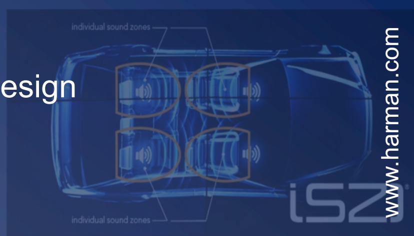
However, this error microphone signal has a contribution from the secondary loudspeaker signal (through H(z)), which is then as if the input microphone were affected by feedback (through H(z))

The subtraction of the filtered loudspeaker signal (filtered by H'z) then also acts as a feedback cancellation.

Conclusion

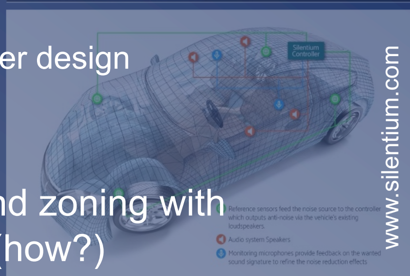
- **Sound zone control**

  Considered <u>fixed</u> control filter design
  - Pressure Matching
  - Acoustic Contrast Control



- **Active noise control**

  Considered <u>adaptive</u> control filter design
  - Feedforward ANC
  - Feedback ANC



PS: Can also consider sound zoning with <u>adaptive</u> control filters (how?)