DIGITAL HEARING AIDS

16 April 2009

James M. Kates

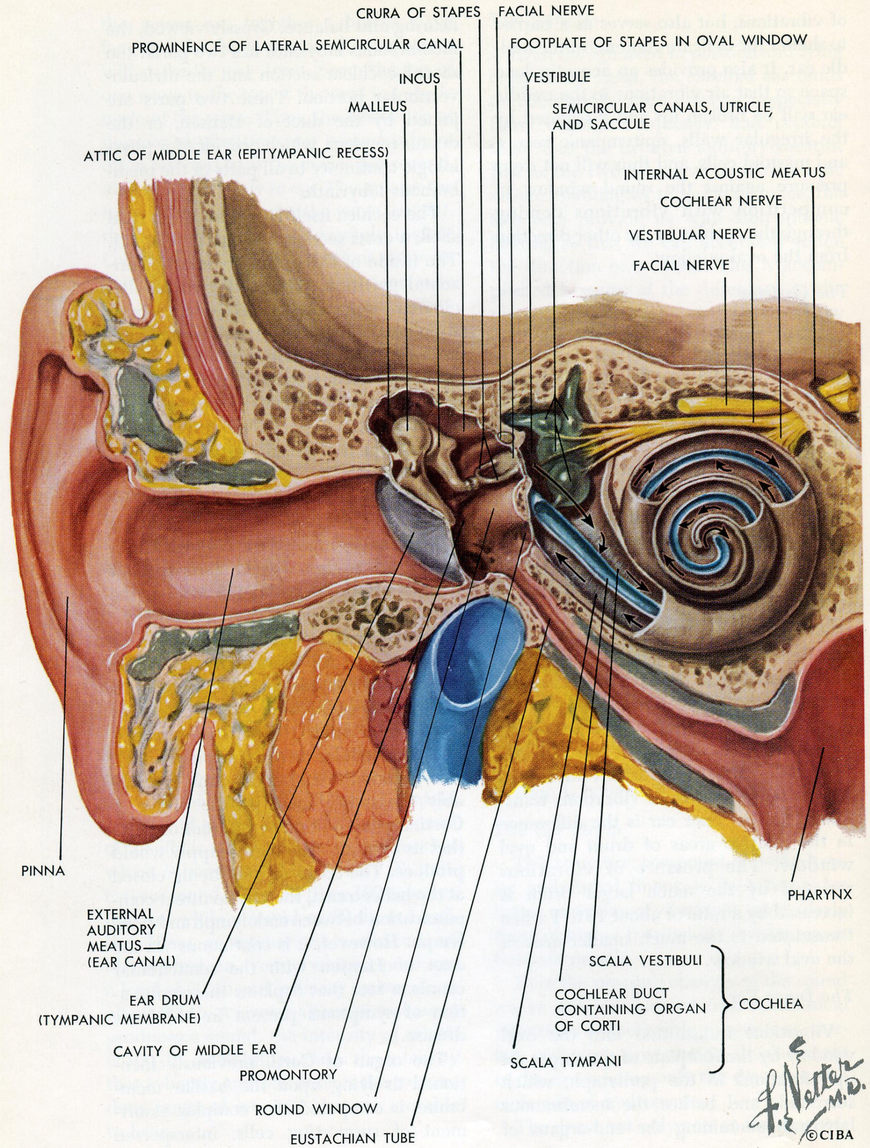
Research Fellow, GN ReSound A/S
Adjoint Faculty, CU Boulder
jkates@gnresound.dk

Contents

Hearing and Hearing Loss

Hearing Loss Classification

Class	Loss, dB	Handicap
Normal	-10 to 26	
Mild	27-40	Difficulty hearing faint or distant speech
Moderate	40-55	Understands speech at a distance of 3-5 feet
Moderately-Severe	55-70	Conversation must be loud, difficulty in group or classroom
Severe	70-90	May hear a loud voice at 1 foot, may distinguish vowels but not consonants
Profound	>90	May hear loud sounds, does not use hearing as primary comm. channel


CRURA OF STAPES
FACIAL NERVE
PROMINENCE OF LATERAL SEMICIRCULAR CANAL
FOOTPLATE OF STAPES IN OVAL WINDOW
INCUS
VESTIBULE
MALLEUS
SEMICIRCULAR CANALS, UTRICLE, AND SACCULE
ATTIC OF MIDDLE EAR (EPITYMPANIC RECESS)
INTERNAL ACOUSTIC MEATUS
COCHLEAR NERVE
VESTIBULAR NERVE
FACIAL NERVE
PINNA
PHARYNX
EXTERNAL AUDITORY MEATUS (EAR CANAL)
SCALA VESTIBULI
COCHLEAR DUCT CONTAINING ORGAN OF CORTI
COCHLEA
EAR DRUM (TYMPANIC MEMBRANE)
CAVITY OF MIDDLE EAR
SCALA TYMPANI
PROMONTORY
ROUND WINDOW
EUSTACHIAN TUBE
F. Netter M.D.
©CIBA

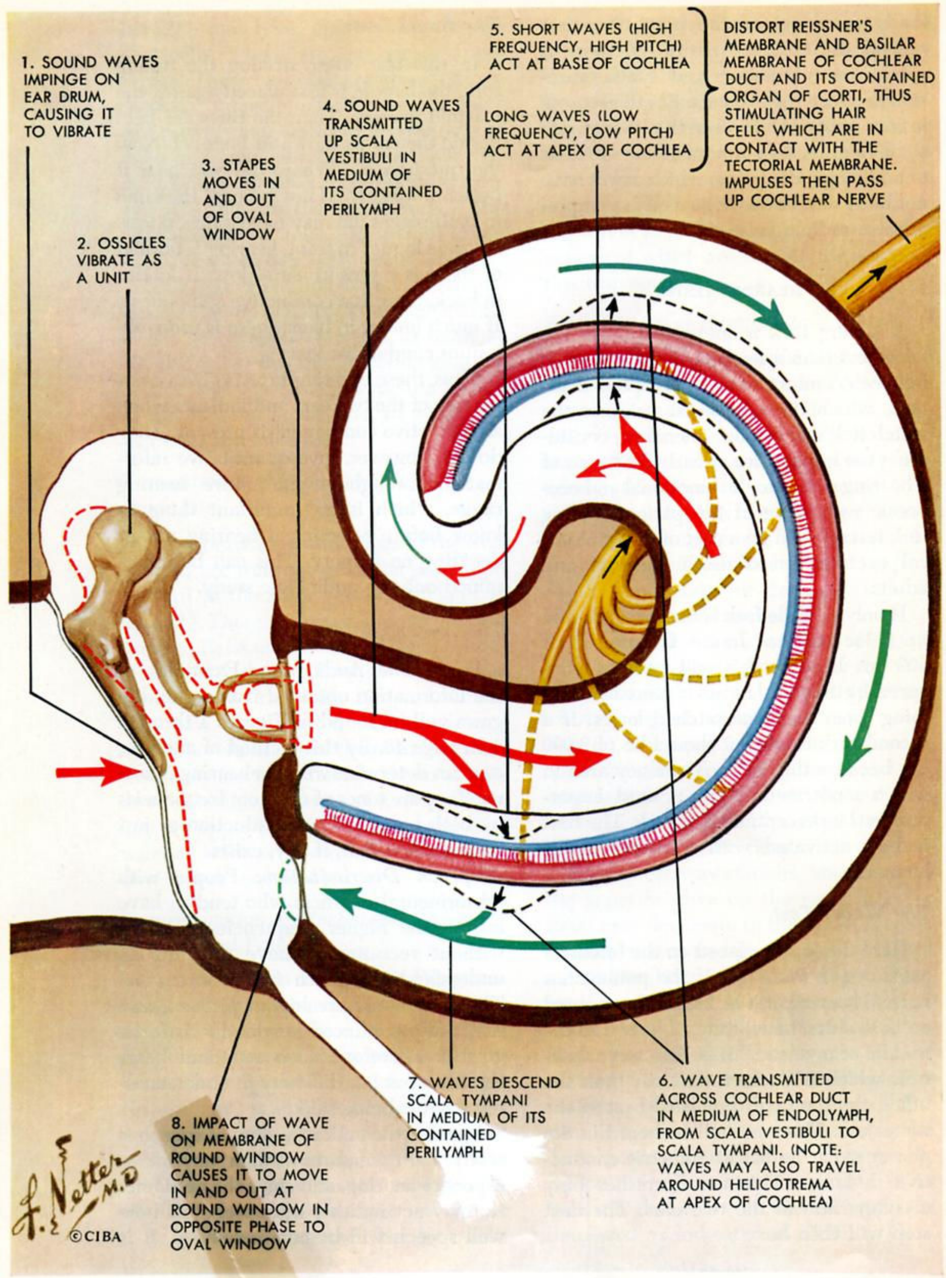


PLATE III TRANSMISSION OF VIBRATIONS FROM DRUM THROUGH THE COCHLEA

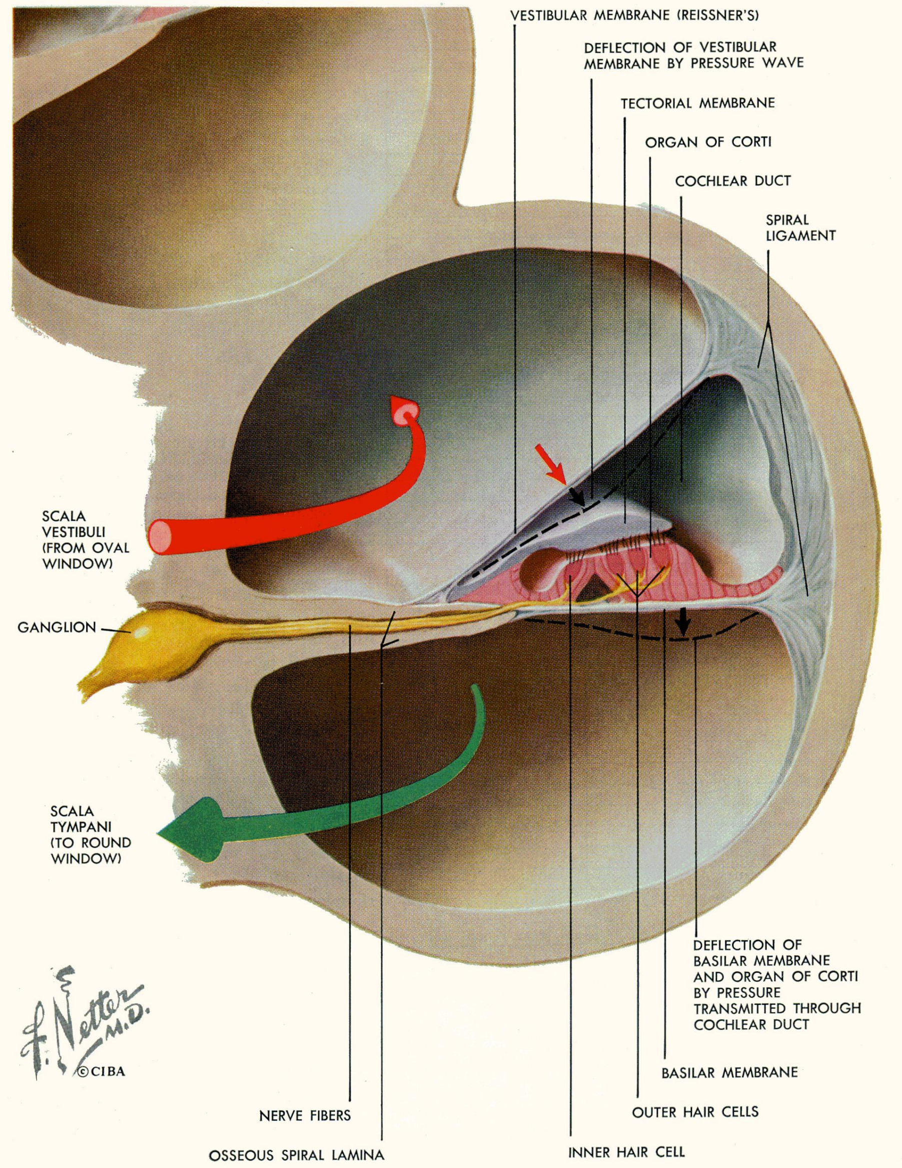
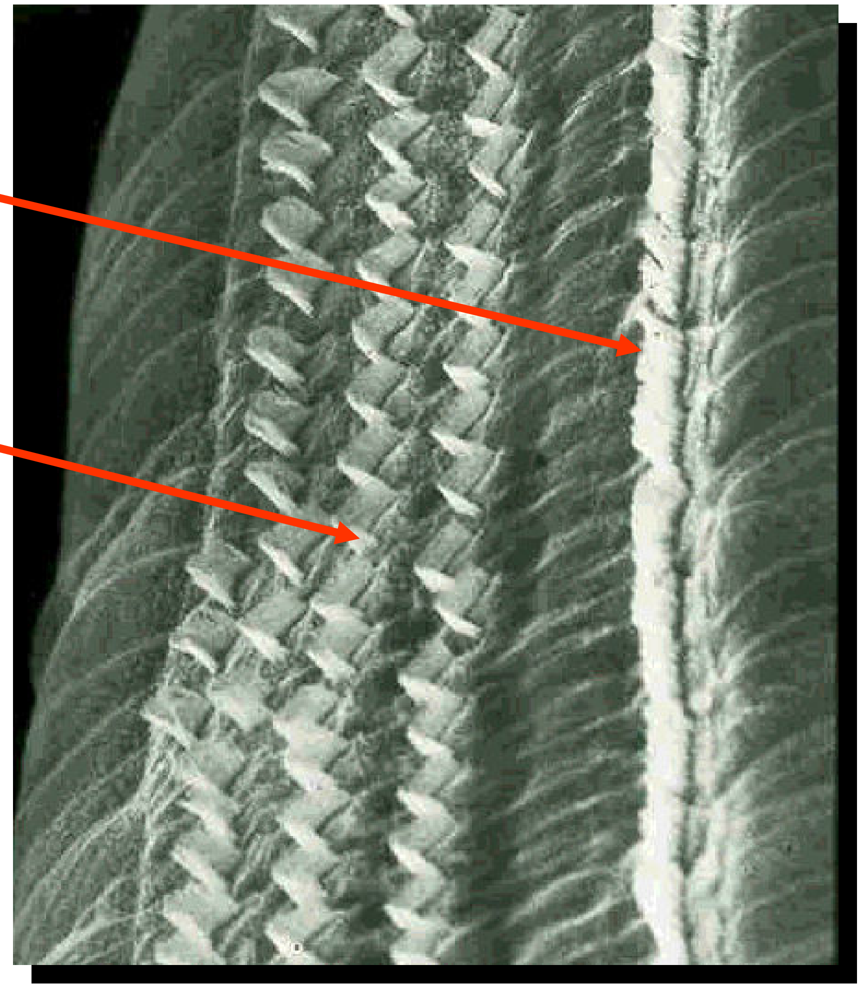


PLATE IV TRANSMISSION OF SOUND ACROSS COCHLEAR DUCT STIMULATING HAIR CELLS

Inner / Outer Hair Cells

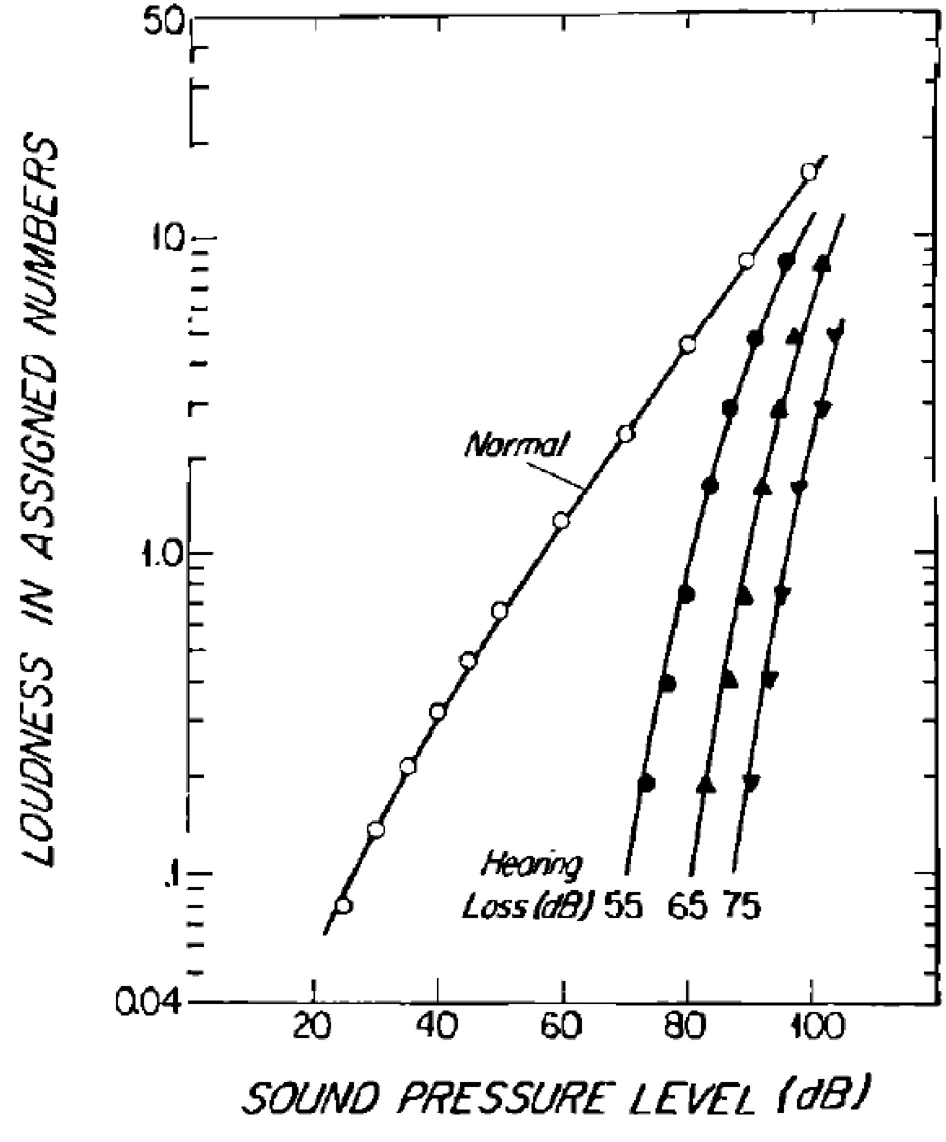
- IHC is the transducer
- OHC enhanced motion:
  - to extent sensitivity for weak sounds
  - sharper filters
- OHC most vulnerable
  - aging
  - noise



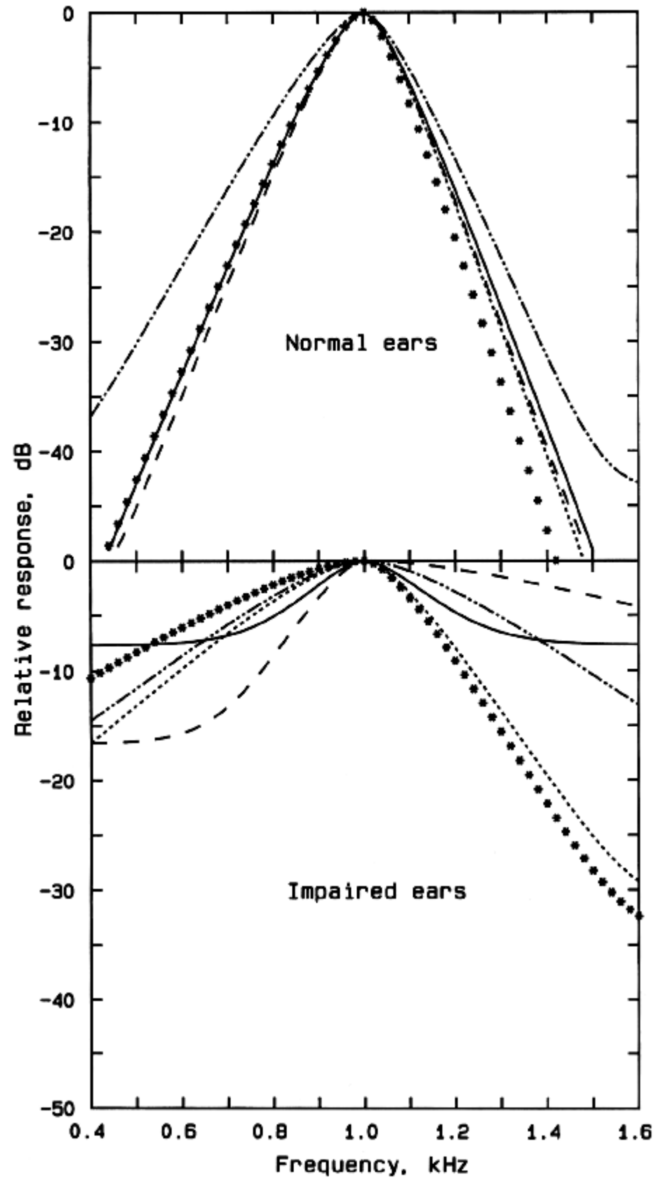
Recruitment

- Reduced dynamic range in impaired ear
  - Healthy cochlea has instantaneous compression
  - Loss of active OHC gain mechanism in impairment
  - Reduced cochlear gain, increased filter bandwidth
  - Impaired ear is more linear than healthy ear
  - Recruitment = abnormal rapid increase in loudness
  - Loudness at 100 dB SPL approx equal
- Compression
  - Compensate for OHC damage
  - High face validity
  - Benefits in practice are mixed

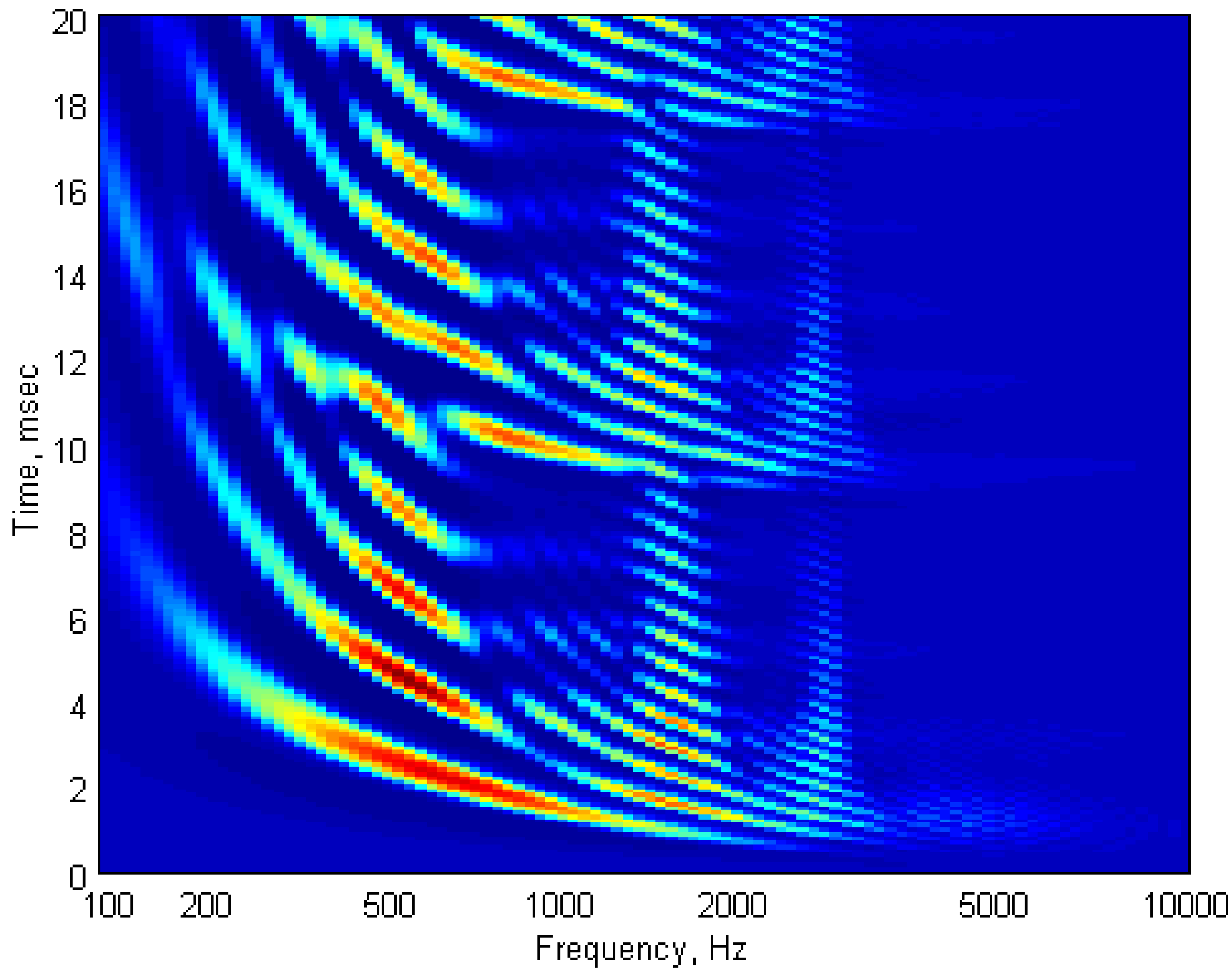
Recruitment Example



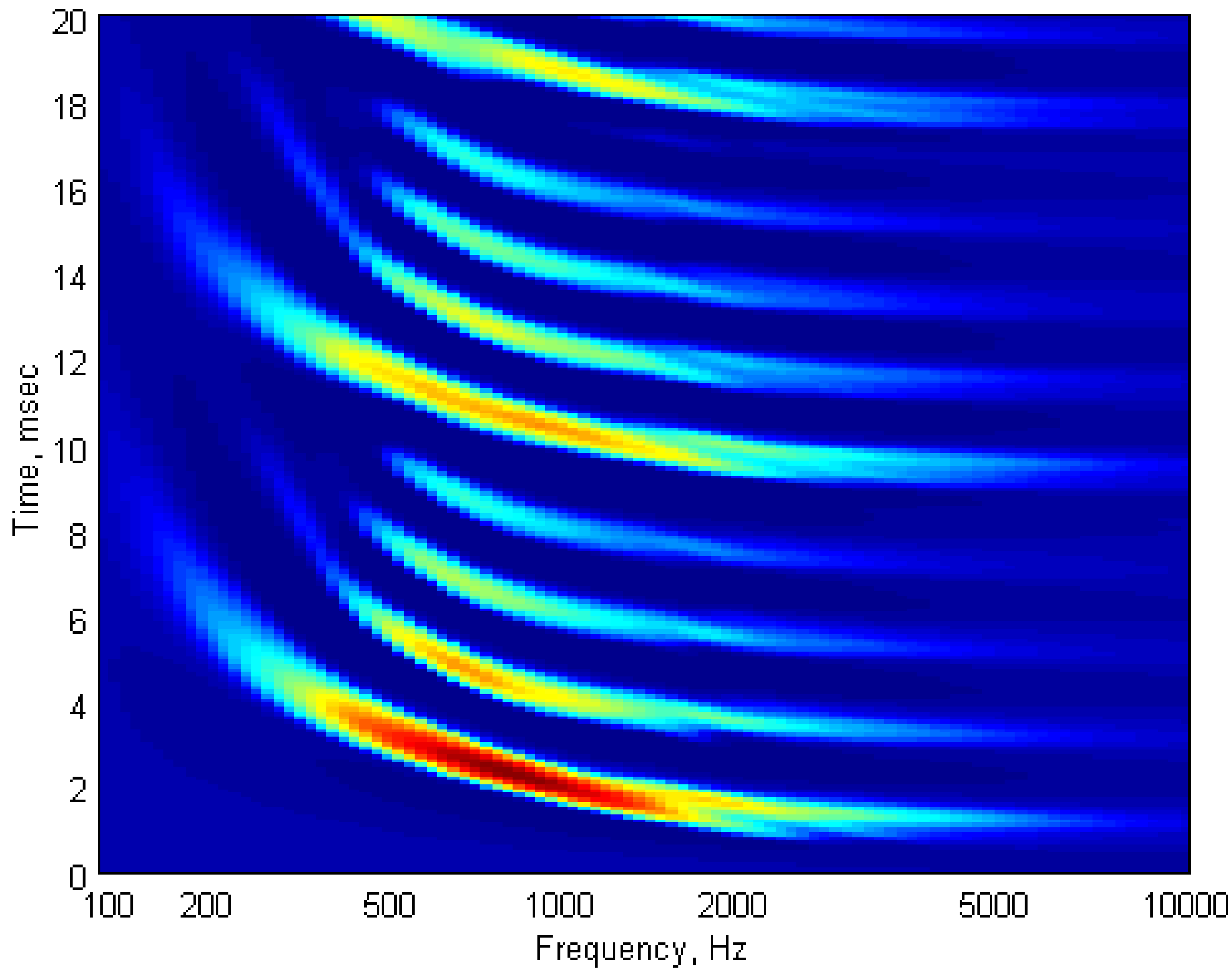
Perceptual Auditory Filters



Simulated Firing Rate /da/, Normal Ear



Simulated Firing Rate /da/, Impaired Ear

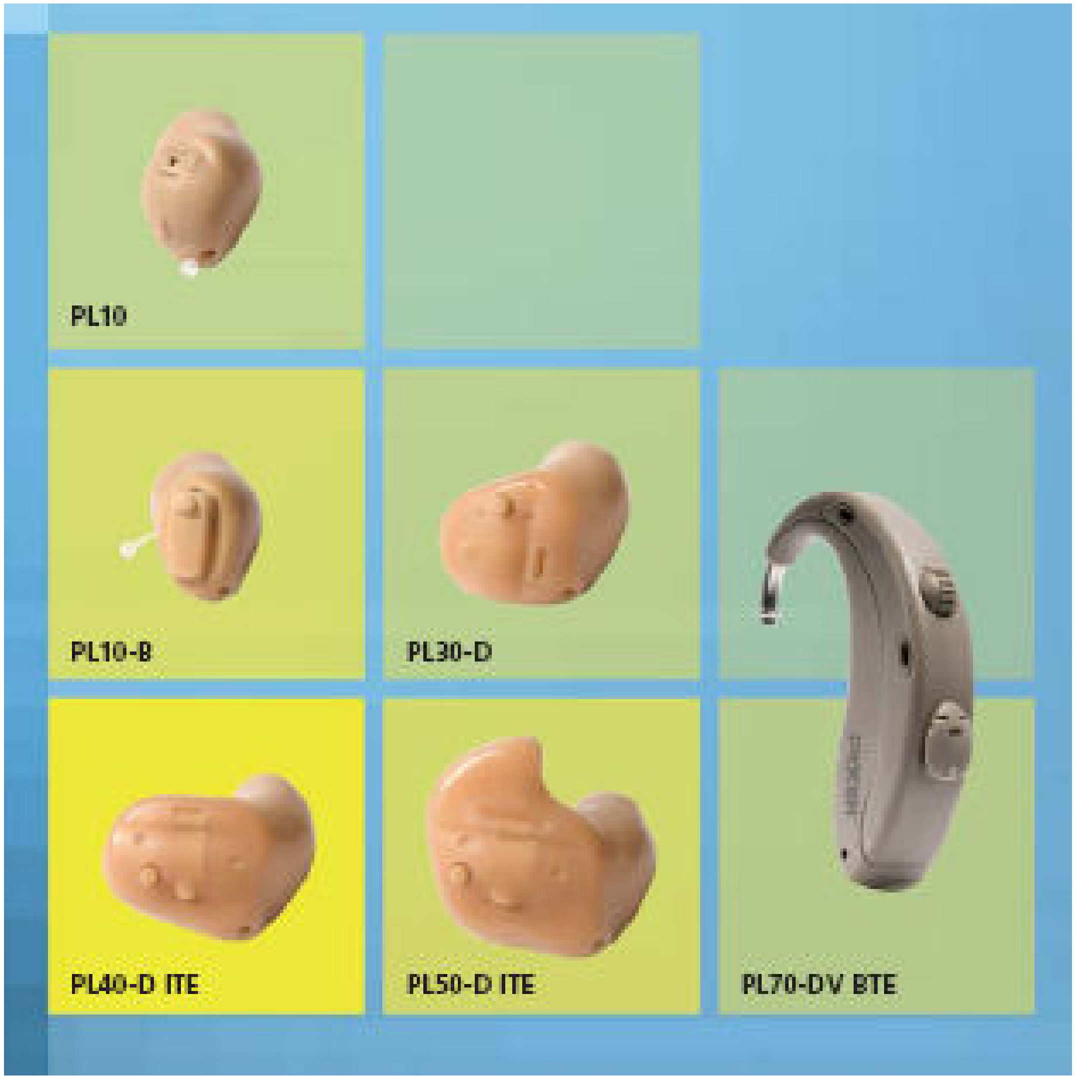


Hearing Loss Conclusions

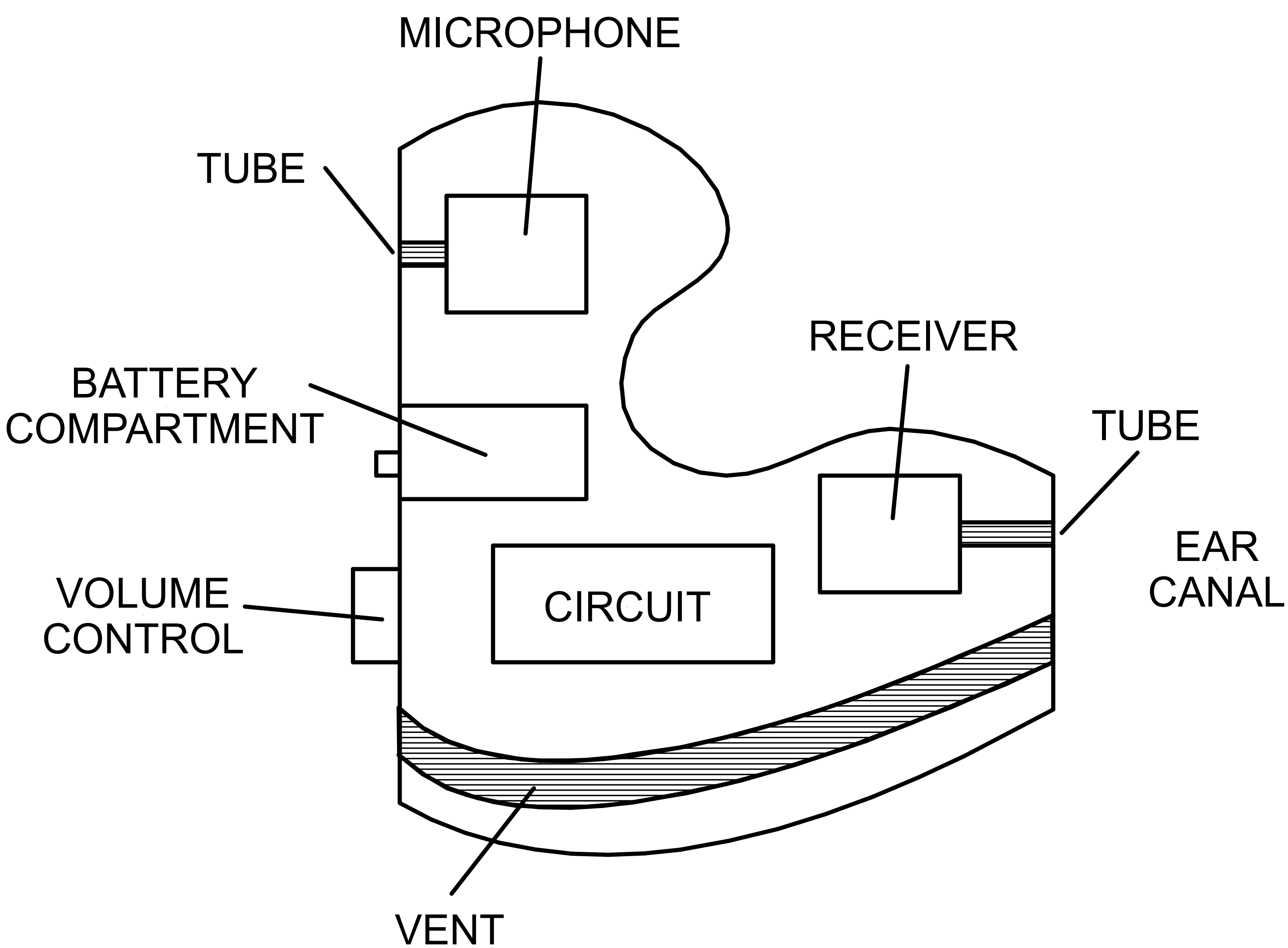
- Outer hair-cell damage
  - Shift in auditory threshold
  - Recruitment: Impaired system is more linear
  - Broader auditory filters
- Inner hair-cell damage
  - Shift in auditory threshold
  - “Dead regions” with no response
- Can not perceive low-intensity speech sounds
- Difficulty in noise and reverberation

Hearing Aid Types and Processing Constraints

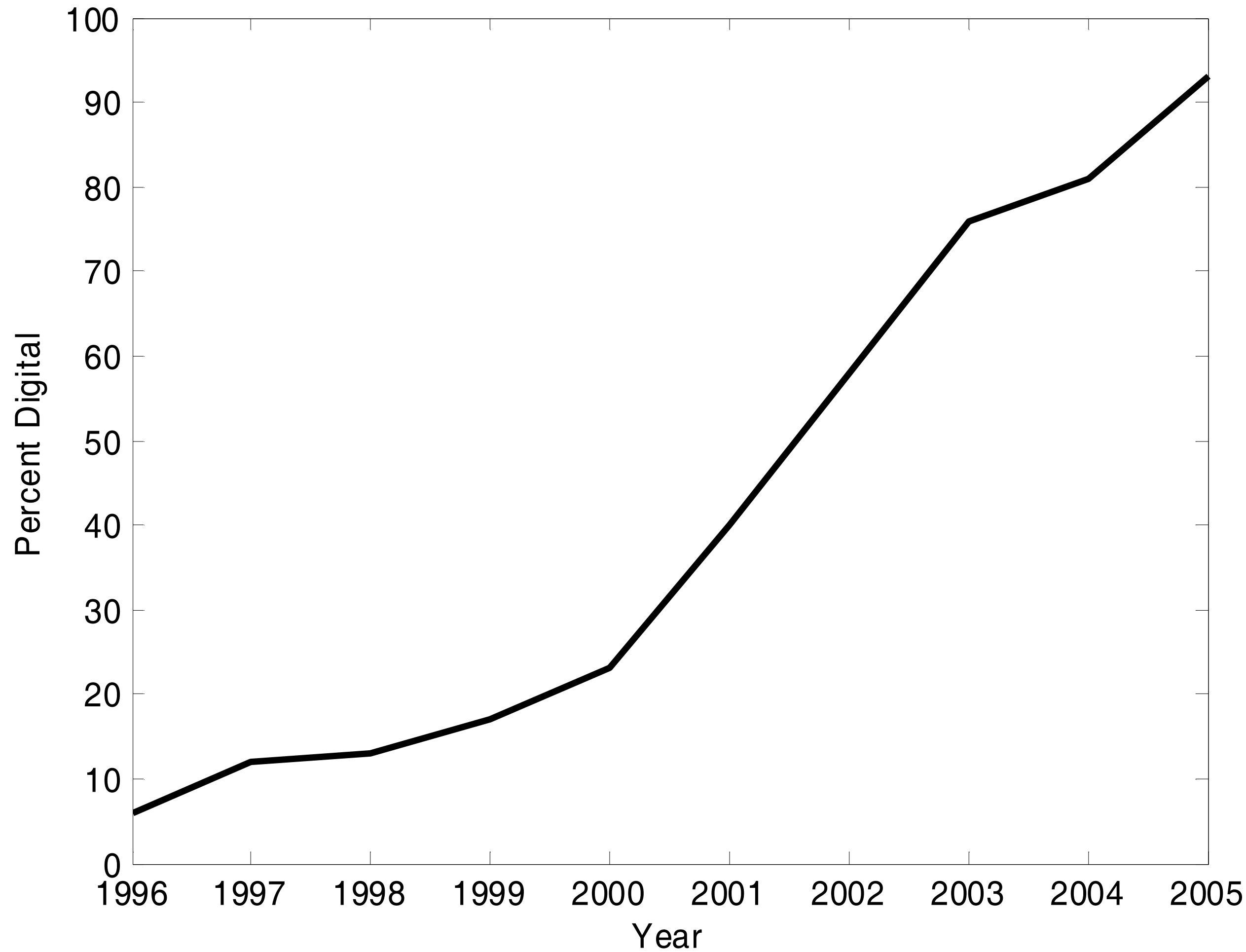
Styles of Hearing Aids



ITE Hearing Aid



Digital Hearing Aid Sales, USA



Factors Affecting Power Use

- Number of Transistors
  - Transistor size decreasing over time
  - Circuit complexity increasing
- Word Size
  - Fixed point vs. floating point
  - Long digital word gives better sound quality
  - Adder and multiplier increase in complexity
  - Memory proportional to word size
- Clock Rate
  - Slow rate reduces battery drain
  - Constrains algorithm complexity
- Voltage
- Algorithms

Digital Processor Comparison

- Personal Computer
  - 6400 MIPS (Xbox 360)
  - 1 to 2 GB memory
  - Draws 20 W
  - Fits on circuit board
- Hearing Aid
  - 4 to 8 MIPS
  - 4 to 8 kW memory
  - Draws 1 mW
  - Fits inside hearing aid

Hardware Conclusions

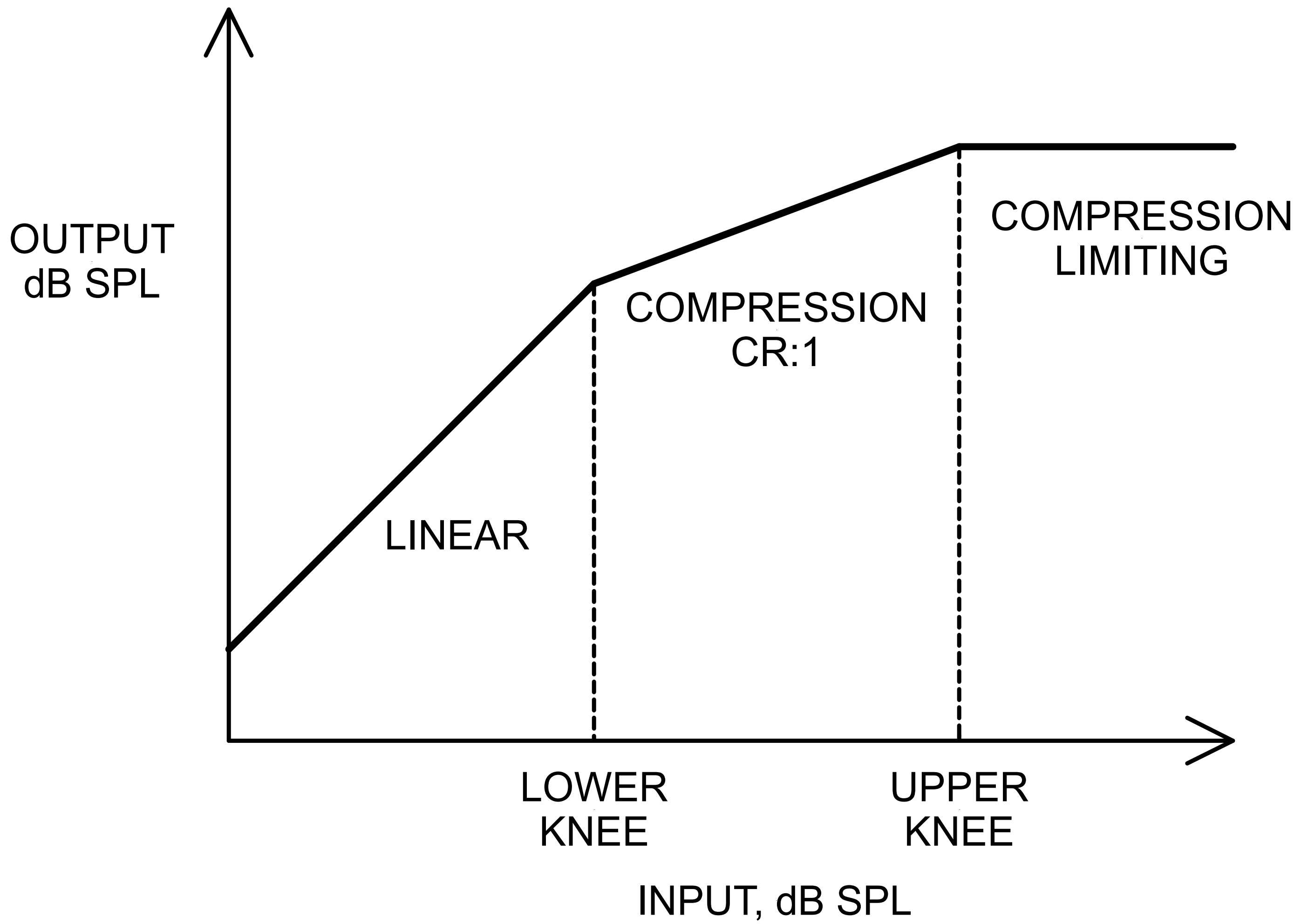
- Hearing Aids Have Become Digital Devices
- Digital Advantages
  - Programmable processor
  - Algorithm complexity
  - Processing features
  - Fitting flexibility
- Digital Limitations
  - Power consumption and battery life
  - Circuit size
  - Cost of circuit and specialized DSP chip development

Dynamic-Range Compression

Compression Rules

- Processing steps
  - Detect signal level
  - Convert to dB SPL
  - Compute gain using input/output relationship
  - Multiply signal by gain
- Piece-wise linear function
  - Kneepoints where slope changes
  - Linear at low levels: Minimize amplification of noise
  - Compression limiting at high levels: Avoid UCL
- Gain = output level – input level in dB

Compressor I/O Function



Envelope Detection

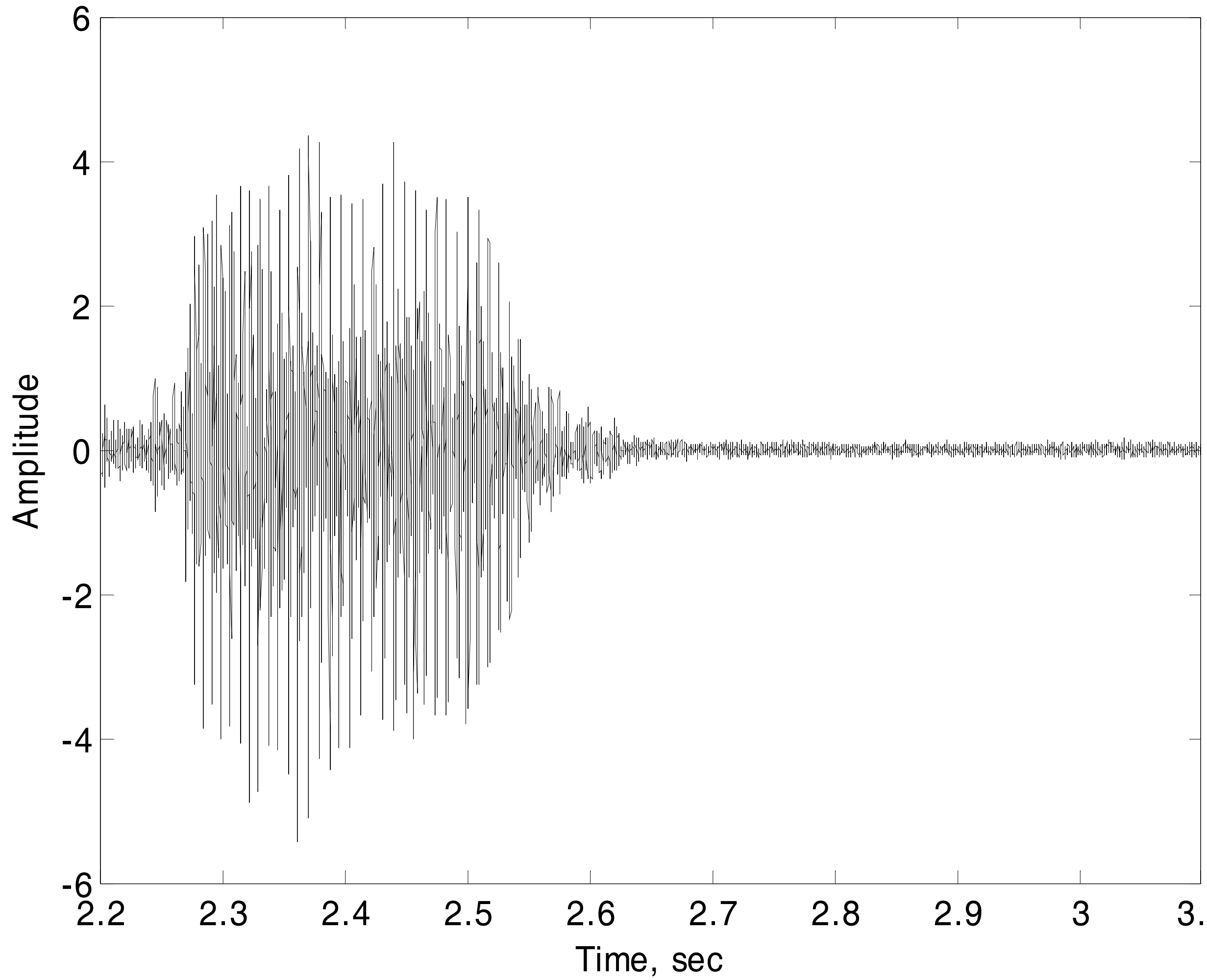
- Track incoming signal level
- Response depends on sign of signal changes
  - Rapid response to increases in signal level (attack)
  - Slower response to decreases (release)
  - Defined by attack and release time constants
- Fast attack
  - Signals tend to increase more rapidly than decrease
  - Prevent over-amplification of large sudden increase
- Slow release
  - Rapid gain changes cause audible modulation
  - Hold gain relatively constant during syllables

Peak Detection

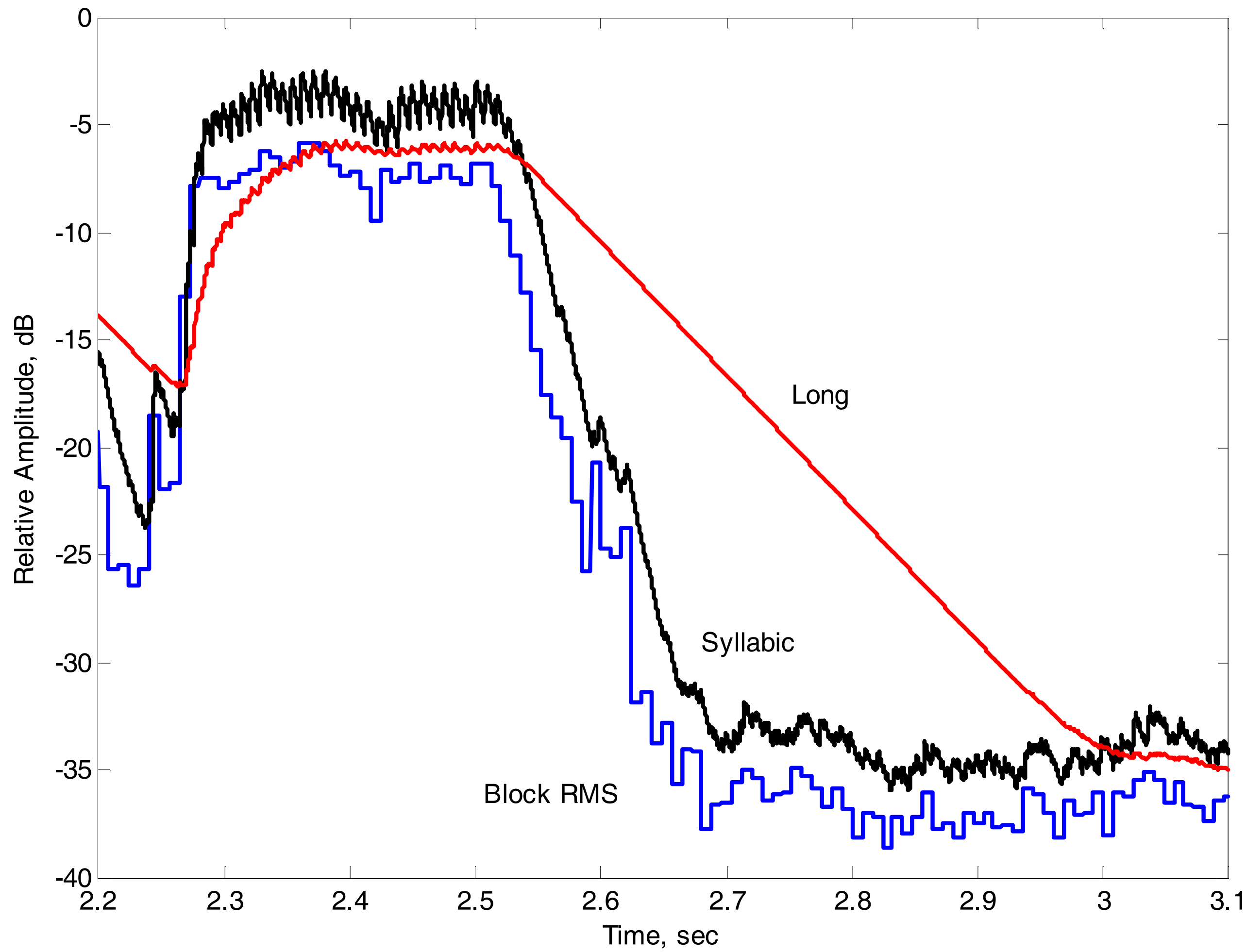
- Follow increase with fast time constant
- Follow decrease with slow time constant

$$
\begin{aligned}
&if\ \left|x(n)\right| \geq d(n-1) \\
&\qquad d(n) = \alpha d(n-1) + (1-\alpha)\left|x(n)\right| \\
&else \\
&\qquad d(n) = \beta d(n-1) \\
&end
\end{aligned}
$$

“Air” Male Talker



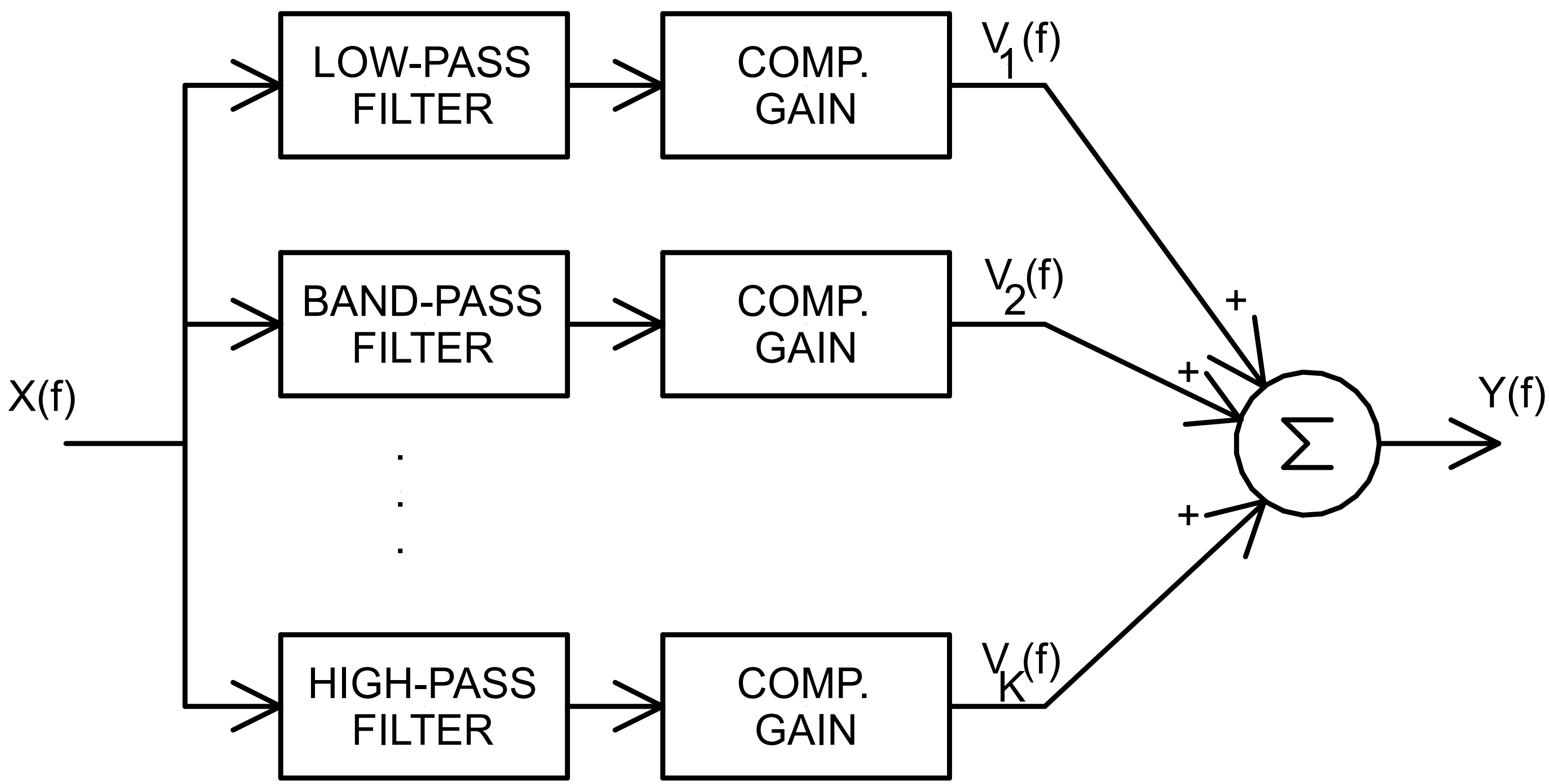
Peak Detector Output



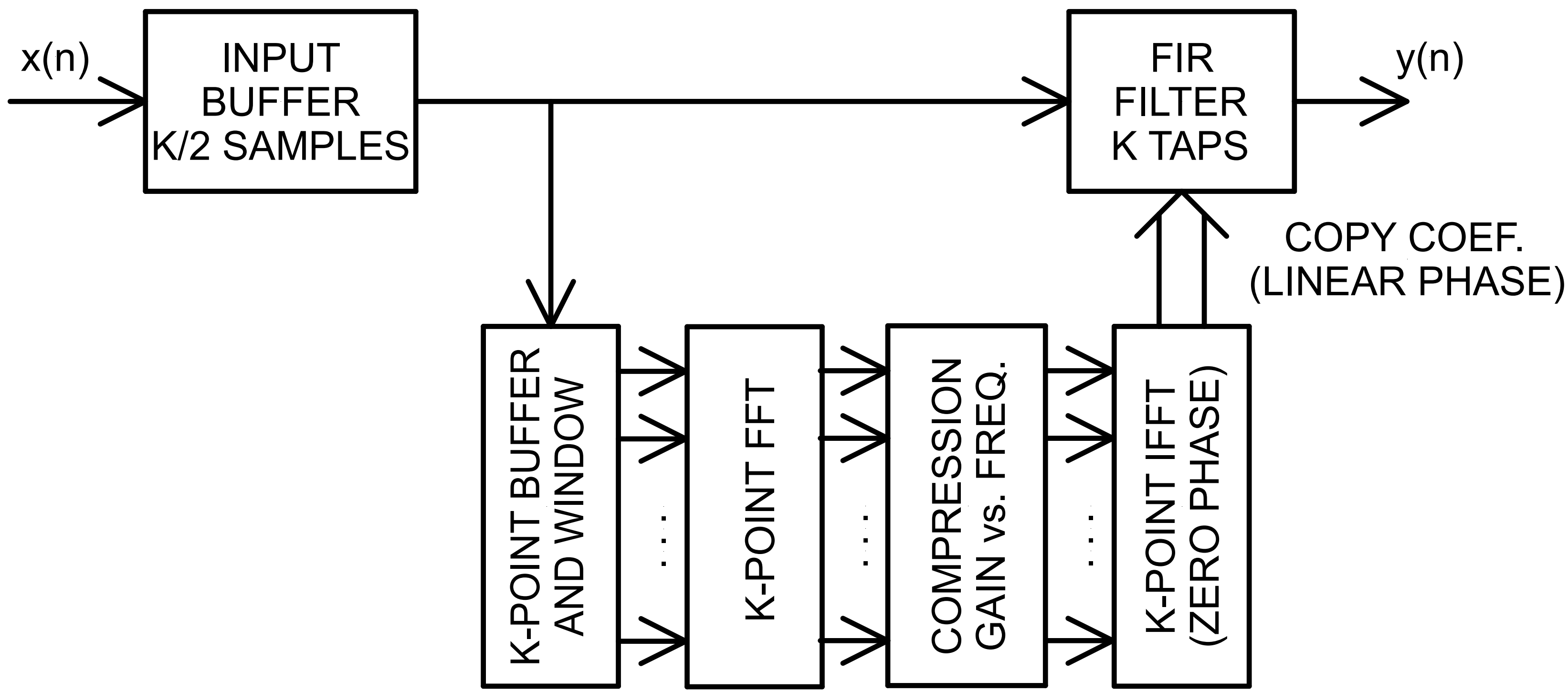
Multichannel Compression

- Frequency analysis
  - Filter bank
  - FFT
- Filter bank
  - Auditory frequency spacing
  - Independent compression in each frequency band
  - Gain set in response to preceding signal level
  - Response to signal change depends on time constants

Multichannel Compressor



Side-Branch Compressor



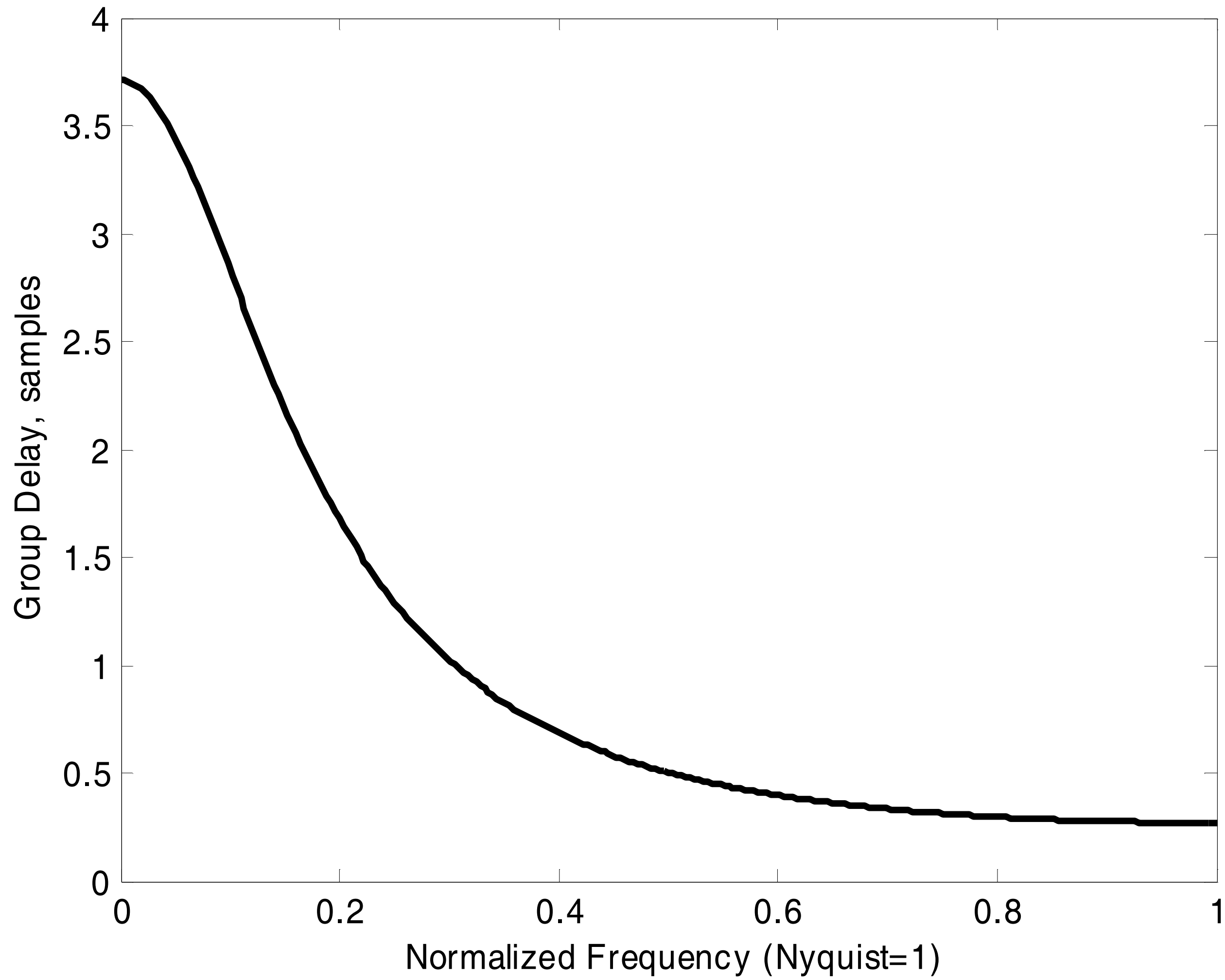
Frequency Warping

- FFT problems
  - Uniform frequency spacing
  - Resolution at low frequencies is poor
  - Need long delay to get good low-frequency analysis
- Goals of frequency warping
  - Auditory frequency analysis
  - Reduced group delay
  - Reasonable computational requirements
- Want group delay < 10 msec
- Used in GN ReSound products

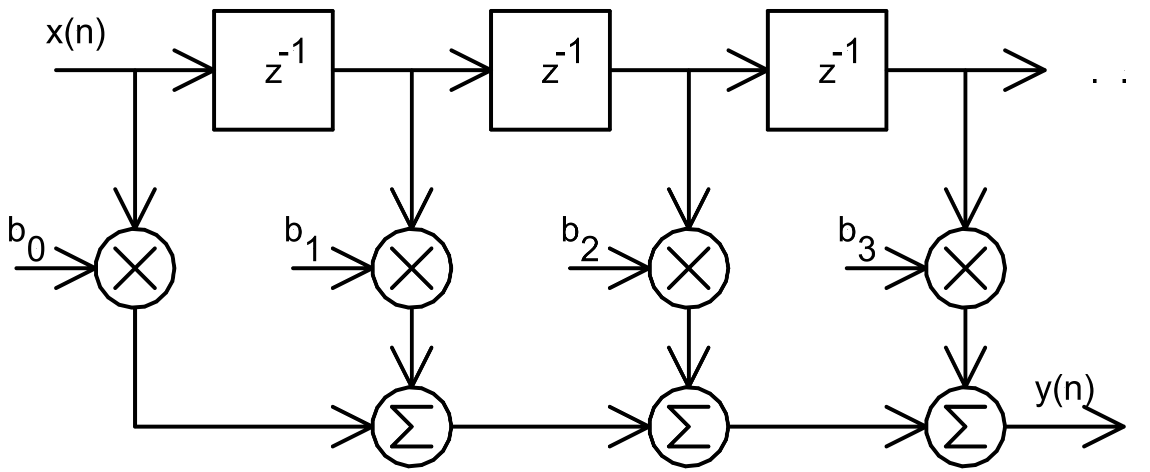
Warped Filter Structure

- Replace unit delays with all-pass filters
  - FIR filter sums outputs at different delays
  - Cascade of all-pass filters for cascade of unit delays
  - Warped FIR sums outputs of all-pass filters
- Effects of the all-pass filters
  - Low frequencies delayed from filter to filter
  - Separation between filtered samples at low frequencies > unit sample
  - Separation at high frequencies < unit sample

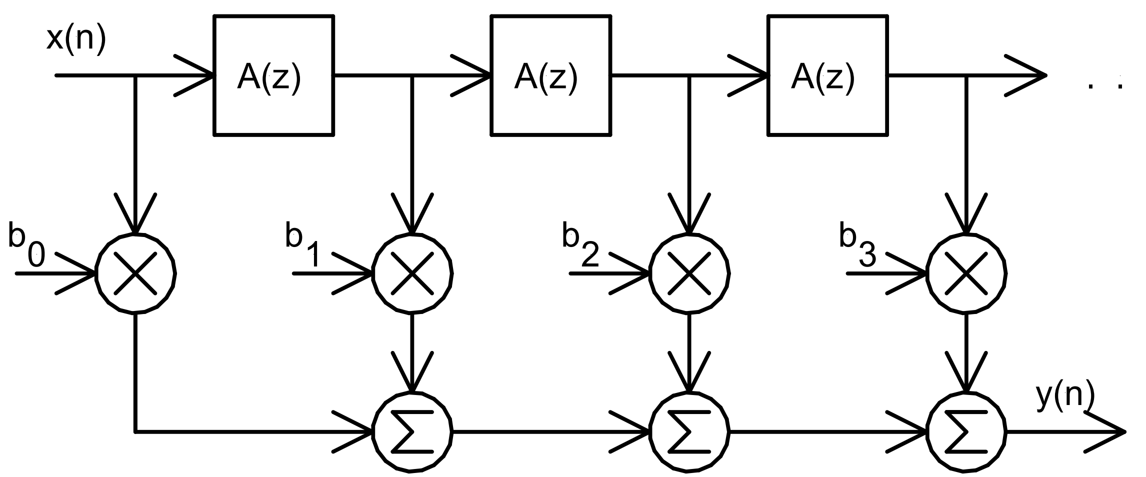
Group Delay of AP Filter



Warped FIR Filter

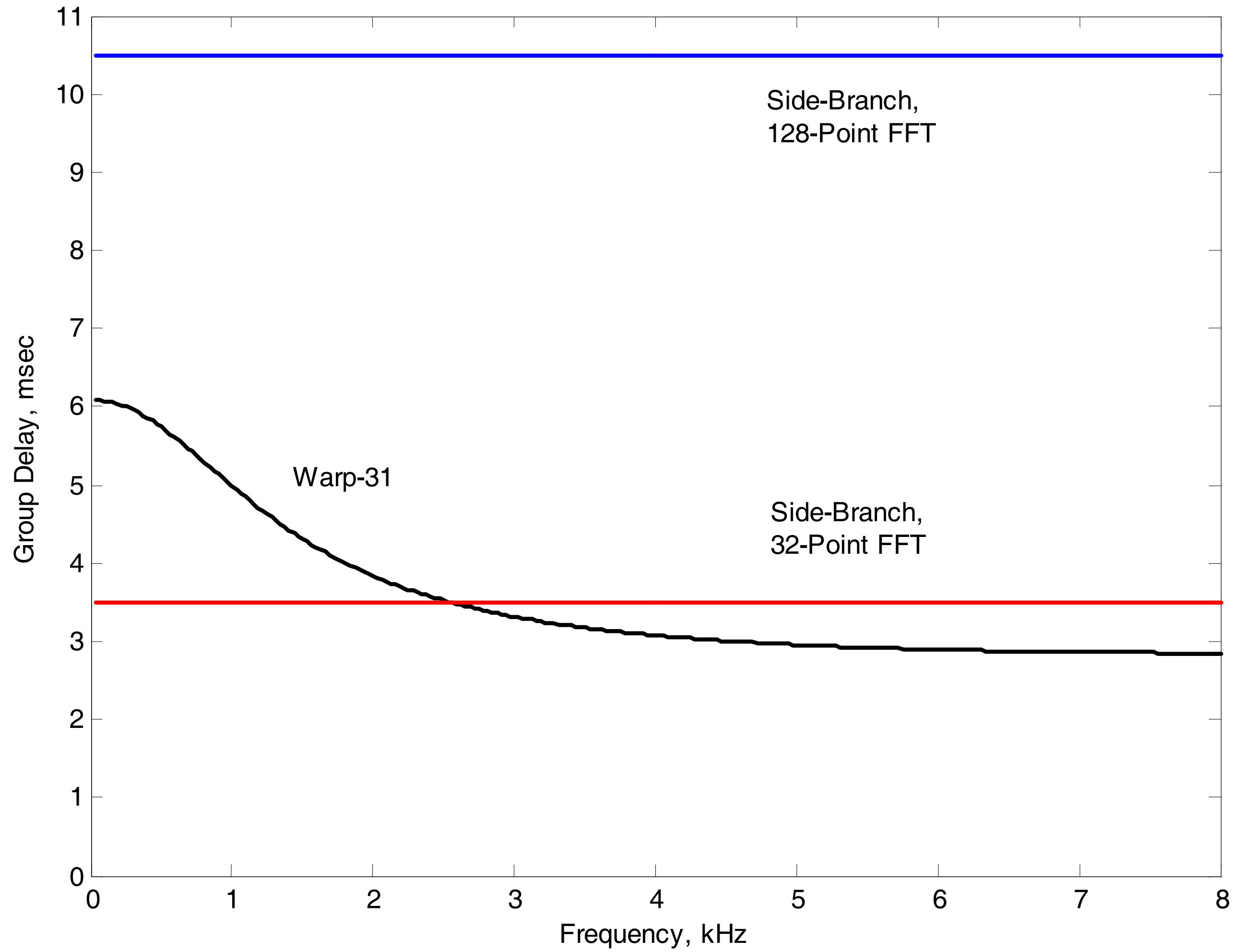


a)



b)

Group Delay Comparison



Warped Compressor Example

- Input
- Sampling rate 22.05 kHz
- 16 bits

- 23-band compressor
- Att time=5 msec
- Rel time=70 msec
- CR=2:1 all bands
- Input at 65 dB SPL

Compression Conclusions

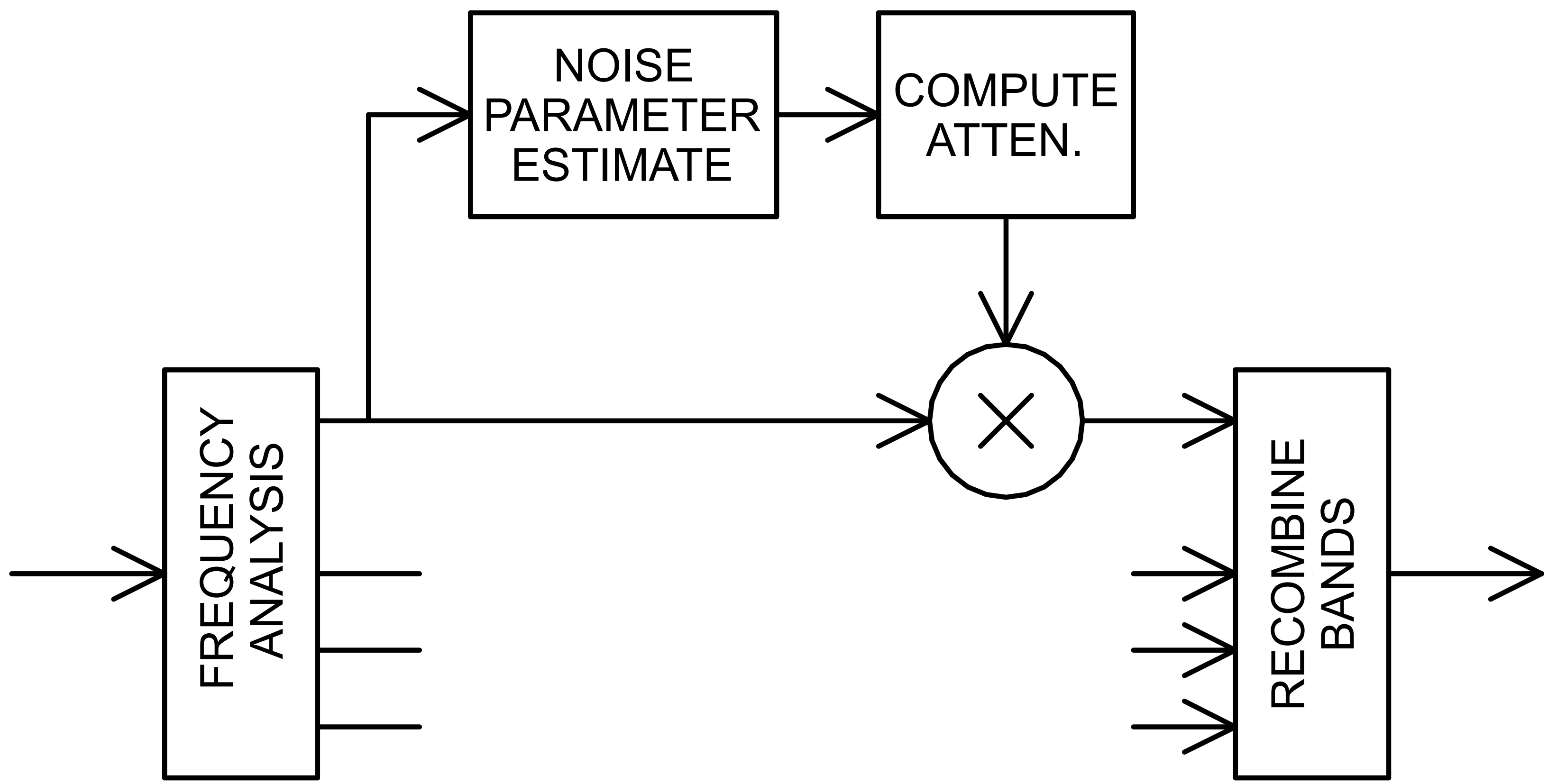
- Helps at low signal levels
- Compression ratio
  - Generally prefer CR < 2:1
  - Want lower CR as noise level increases
  - Comp preferred if residual dynamic range < 30 dB
- Number of Channels
  - No clear benefit to increasing the number of bands
  - One channel shows small benefit compared to multi
  - Compressor co-modulates noise to match speech

Noise Suppression

Spectral Subtraction

- Estimate clean power spectrum
  - Start with power spectrum of the noisy speech
  - Estimate the noise power spectrum
  - Clean spectrum approximated by noisy speech spectrum minus estimated noise spectrum
  - Goal: envelope of noisy speech matches clean speech
- Problems
  - Can not extract actual noise signal, only its statistics
  - Signal reconstructed using noisy phase
  - Gain changes from segment to segment, causing “musical noise”

Generic Noise Suppression



Noise Estimation

- Spectral subtract needs accurate noise estimate
- Voice activity detection
  - Monitor signal segment by segment
  - Determine if segment is speech or noise
  - Update noise statistics during noise segment
  - Hold statistics constant during speech segment
- Continuous noise estimation
  - Adjust noise statistics on every segment
  - Update rate based on probability that segment is noise
  - Reduced computation better match to hearing aid

Hirsch-Ehrlicher Algorithm

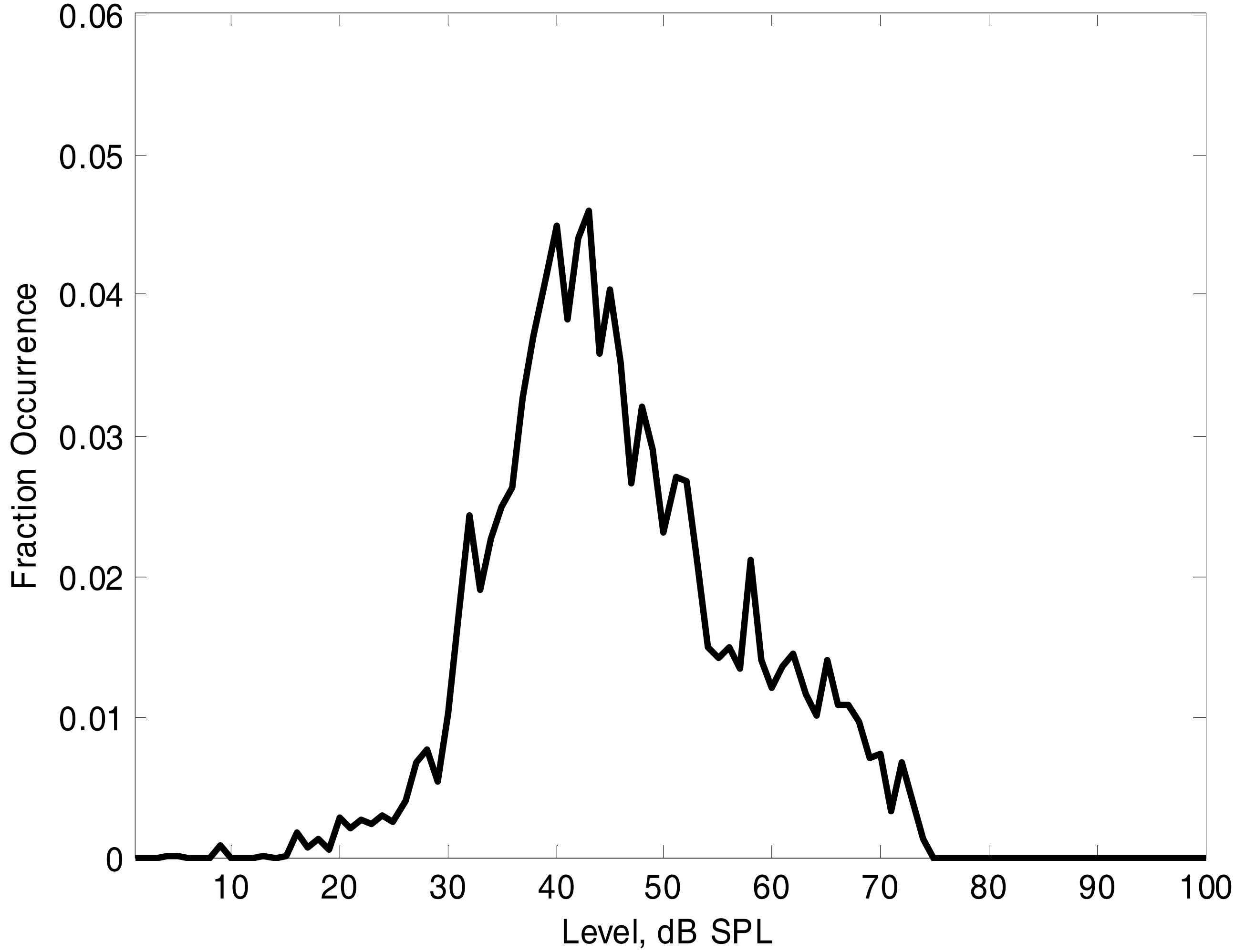
- Incoming noisy signal: $|X(k,m)|$
- Noise estimate: $|N(k,m)|$
- Update the noise estimate

$$
\begin{aligned}
&if\ |X(k,m)| > b|N(k,m-1)| \\
&\quad |N(k,m)| = |N(k,m-1)| \\
&else \\
&\quad |N(k,m)| = a|N(k,m-1)| + (1-a)|X(k,m)| \\
&end
\end{aligned}
$$

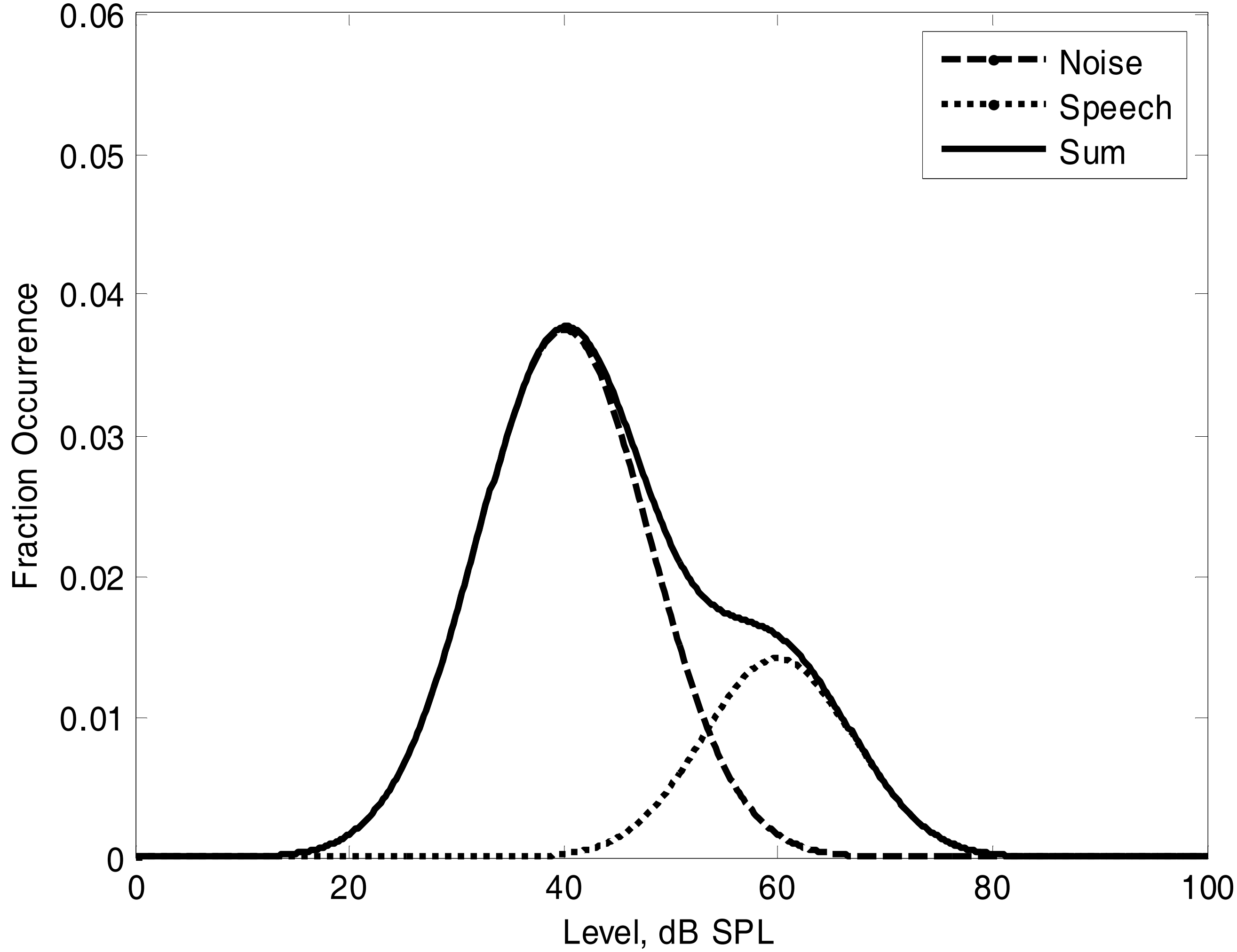
Histogram Estimation

- Sample each noisy speech segment
- Assign segment to closest histogram bin
  - Histogram bins contain count of number of segments at that intensity level
  - Decay contents of all bins
  - Increment contents of assigned bin
- Find peak of noise distribution
  - Model as two Gaussian distributions
  - Assume noise lower distribution
  - Search for peak below the mean

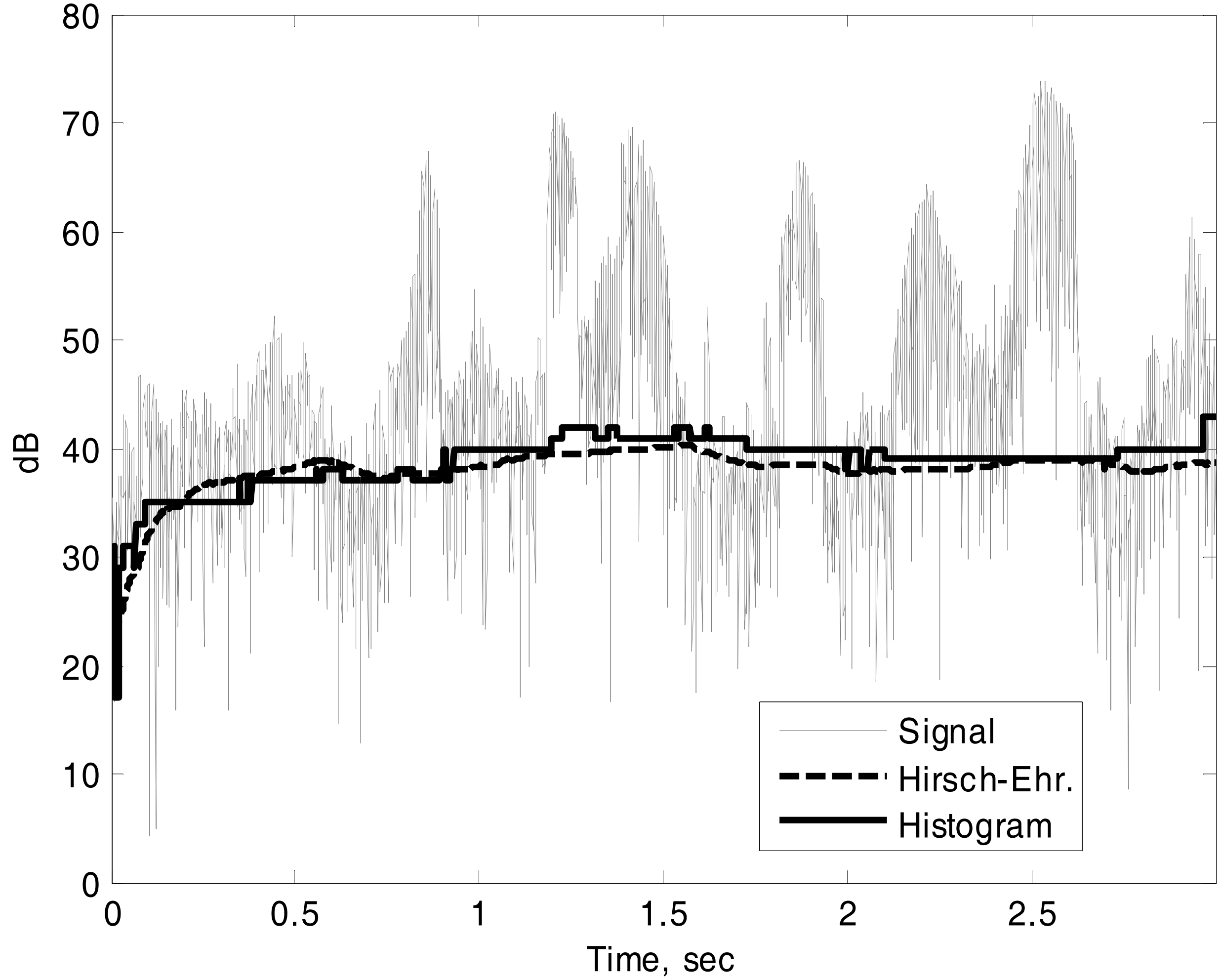
Log-Level Histogram



Gaussian Fit to Histogram



Noise Estimation, SNR=20dB



Wiener Filter Calculation

- Noisy signal $x(n)=s(n)+d(n)$, output $y(n)$
- Min MSE between output and clean input

$$\varepsilon = \sum_n \left[s(n) - y(n) \right]^2$$

- Optimal filter:

$$G(f) = \frac{|S(f)|^2}{|S(f)|^2 + |D(f)|^2}$$

- Approx from noisy speech and estimated noise:

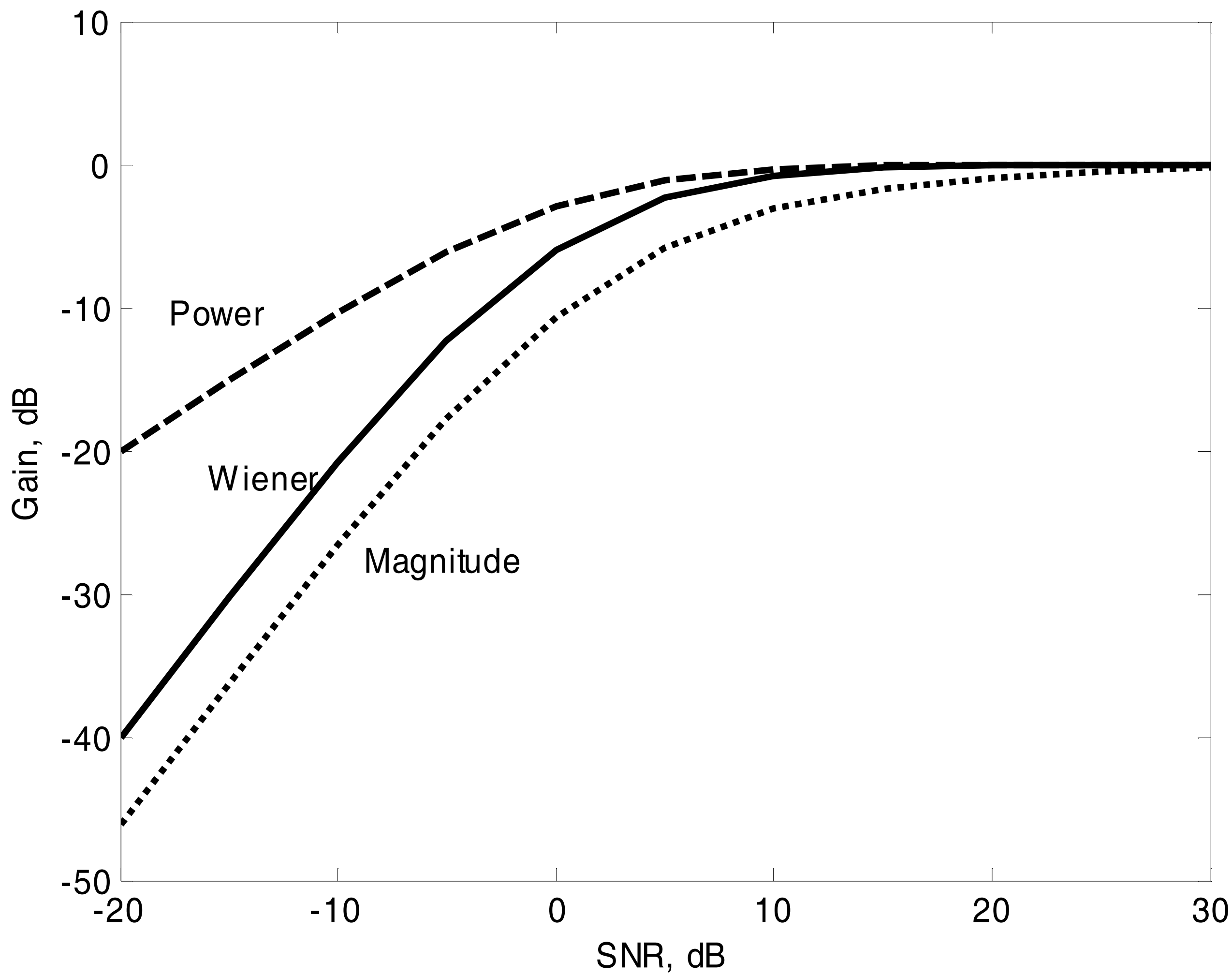
$$G(f) \approx \frac{|X(f)|^2 - |N(f)|^2}{|X(f)|^2}$$

Adaptive Wiener Filter

- Time-varying signal and noise estimates
  - Signal spectrum: Replace $X(f)$ with $X(k,m)$
  - Noise spectrum: Replace $N(f)$ with $N(k,m)$
- Time-varying gain $G_W(k,m)$ in band k, block m

$$G_W(k,m) = \frac{|X(k,m)|^2 - |N(k,m)|^2}{|X(k,m)|^2} = 1 - \frac{|N(k,m)|^2}{|X(k,m)|^2}$$

Spectral Subtraction Gain



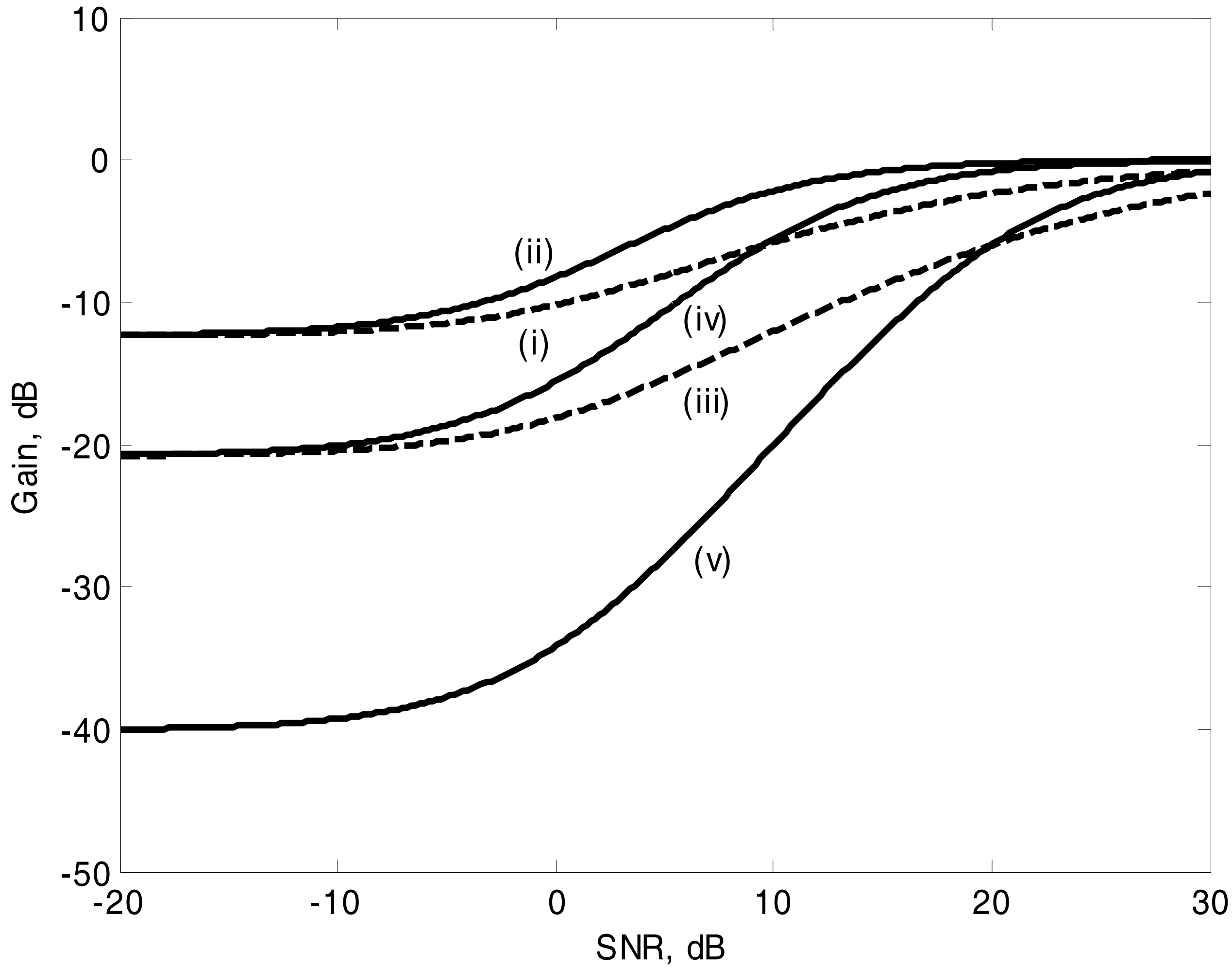
TMK and Related Algorithms

- Gain based on signal and estimated noise

$$G(k,m) = \frac{1}{1 + \nu \left[\frac{\left| N(k,m) \right|}{\left| X(k,m) \right|} \right]^{\gamma}}$$

- Parameters
  - Increasing ν increases maximum attenuation
  - Increasing γ increases function slope with SNR

TMK Spectral Subtraction



Spectral Subtraction Example

- Input
- Sampling rate 22.05 kHz
- 16 bits
- Stationary speech-shaped noise
- SNR=10dB

- 23-band warped system
- TMK algorithm
- 12-dB maximum attenuation

Noise Suppression Conclusions

- Little improvement in intelligibility
  - Depends on algorithm
  - TMK and Ephraim-Malah most effective
- Some improvement in quality
  - TMK shows benefit for NH and HI listeners
  - Strongest effect in range of 10 to 20 dB SNR
- Greatest benefit in stationary noise

Feedback Cancellation

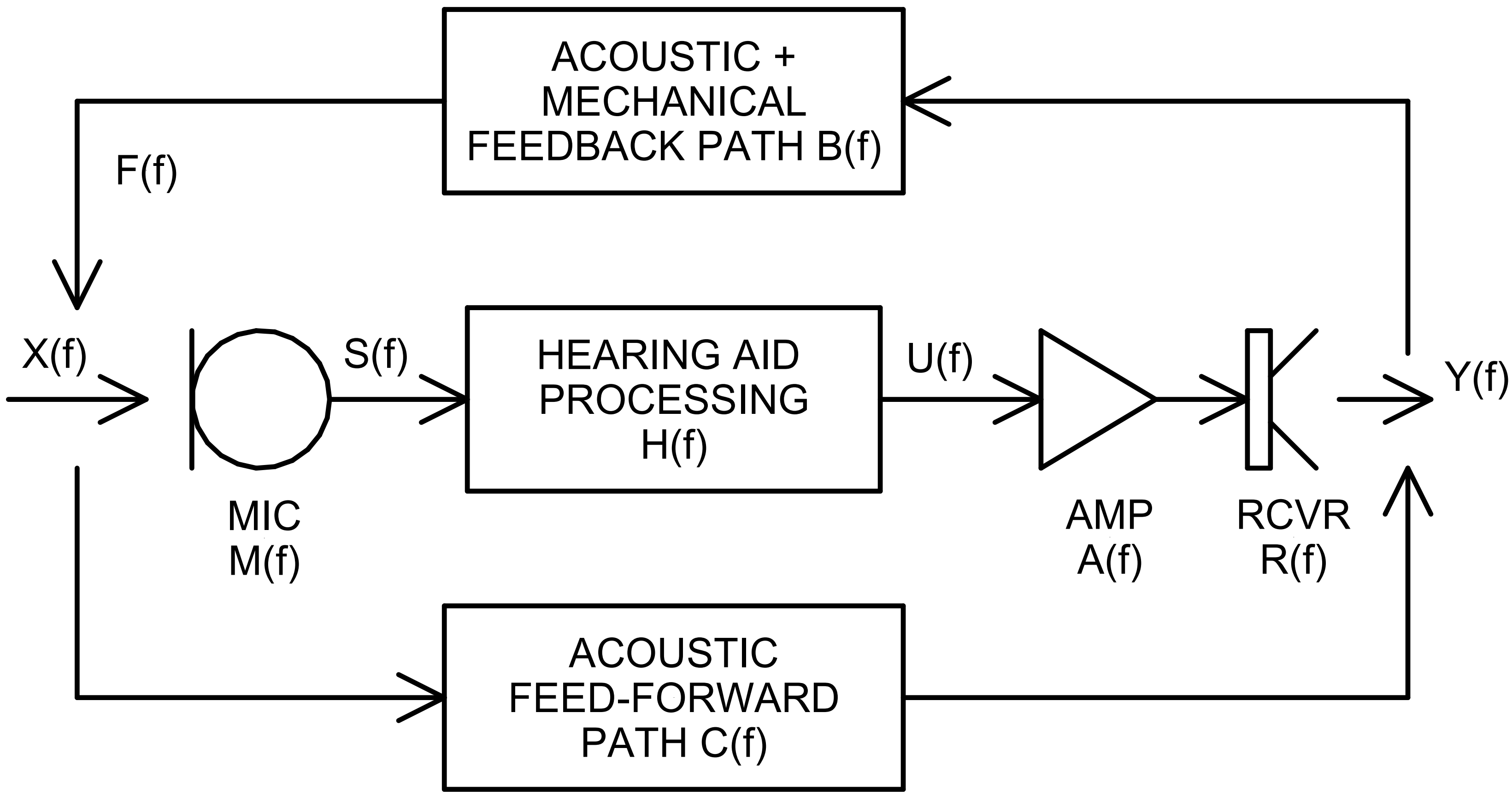
The Problem of Feedback

- Primary sources of feedback
  - Acoustic from vent and leaks
  - Mechanical from receiver vibration
- Effect on hearing aid
  - Instability causes "whistle" or high-frequency tone
  - Ringing after sound stops
  - Hearing aid amplifier saturates => distortion
- Maximum stable gain
  - Maximum gain for operation without whistles
  - Problem most acute at high frequencies

Algorithm Goals

- Increased gain
  - Achieve fitting targets
  - Improved speech intelligibility
- No processing artifacts
  - Hearing aid always stable
  - No whistles due to feedback
  - No chirps, momentary tones, or audible gain changes
- Computationally efficient

Feedback System



System Output

- Output = Processed + Direct

$$Y = X \frac{C + H[MAR]}{[1 - BC] - H[MARB]}$$

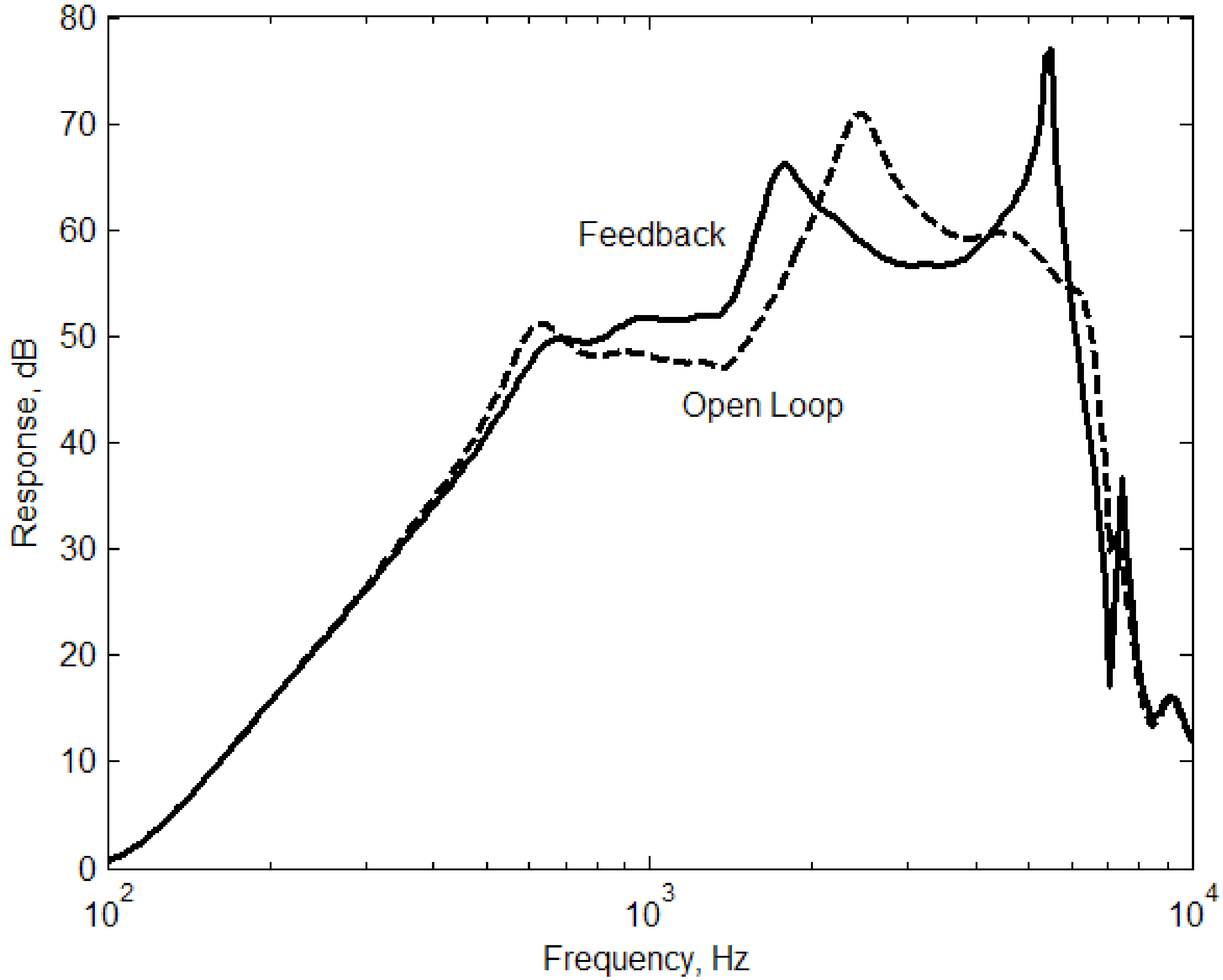
- Approximate solution for $|BC| << 1$

$$Y \cong X \frac{C + H[MAR]}{1 - H[MARB]}$$

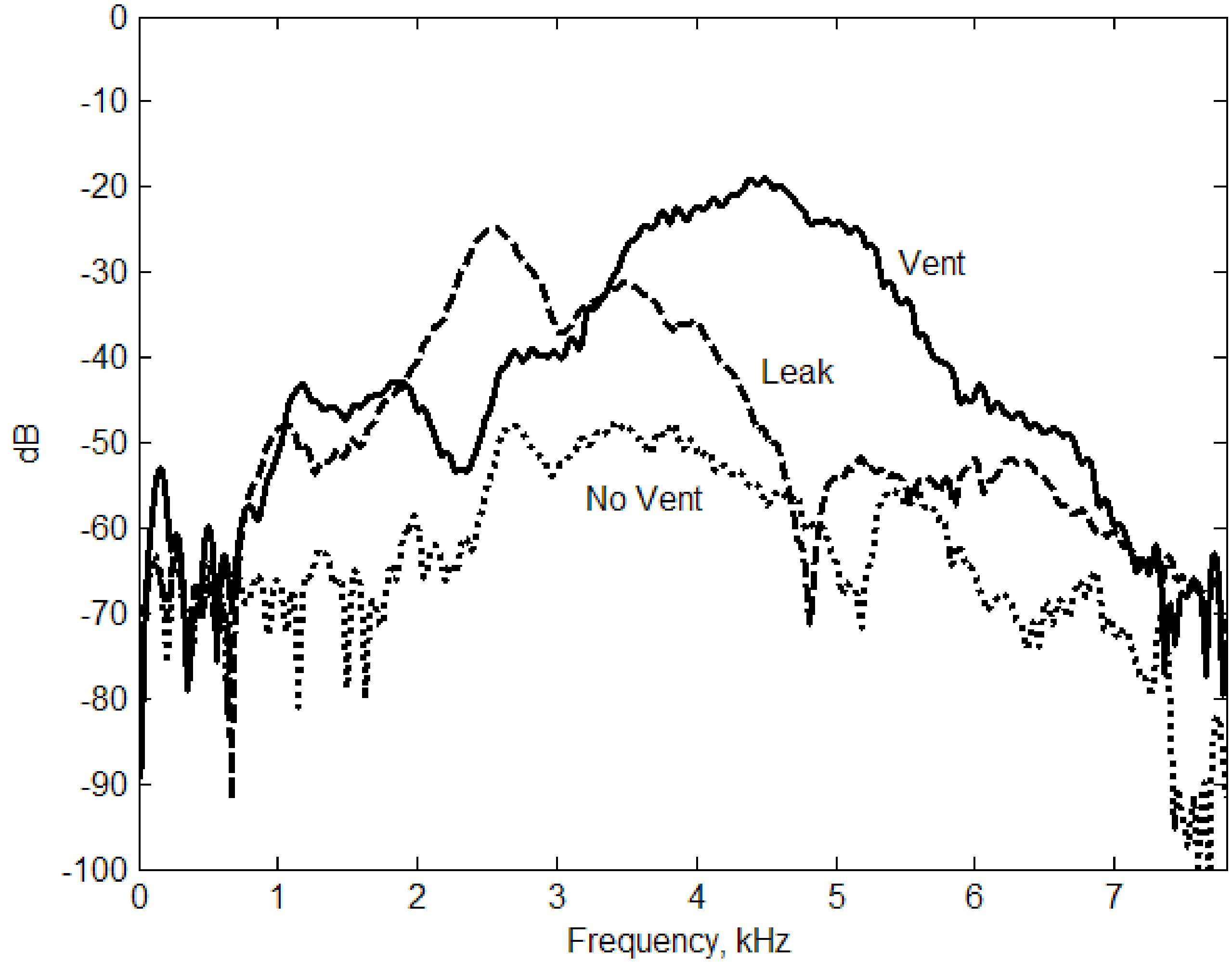
- Stability guaranteed for

$$|H[MARB]| < 1$$

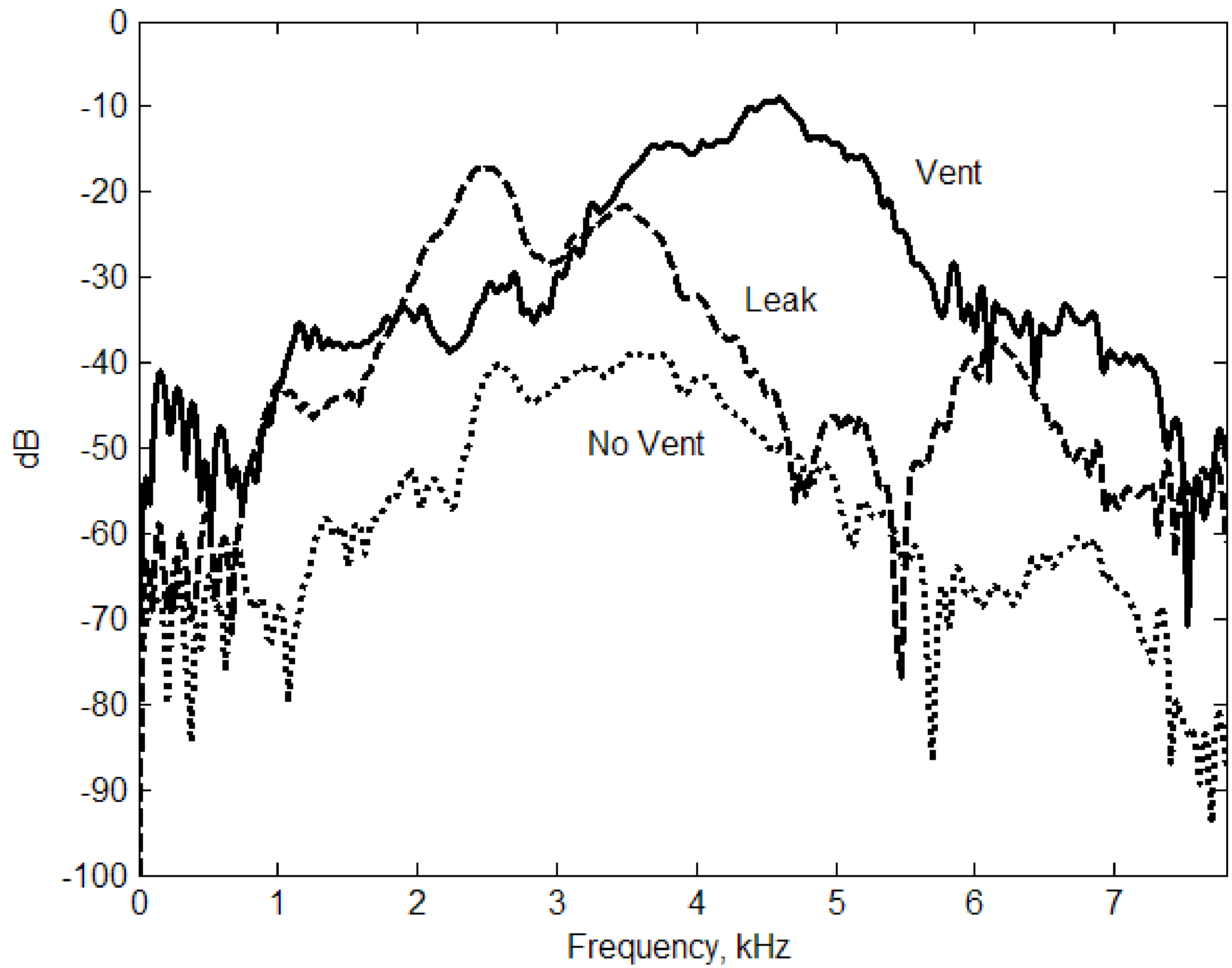
Effects of ITE Vent Feedback



BTE Feedback Path



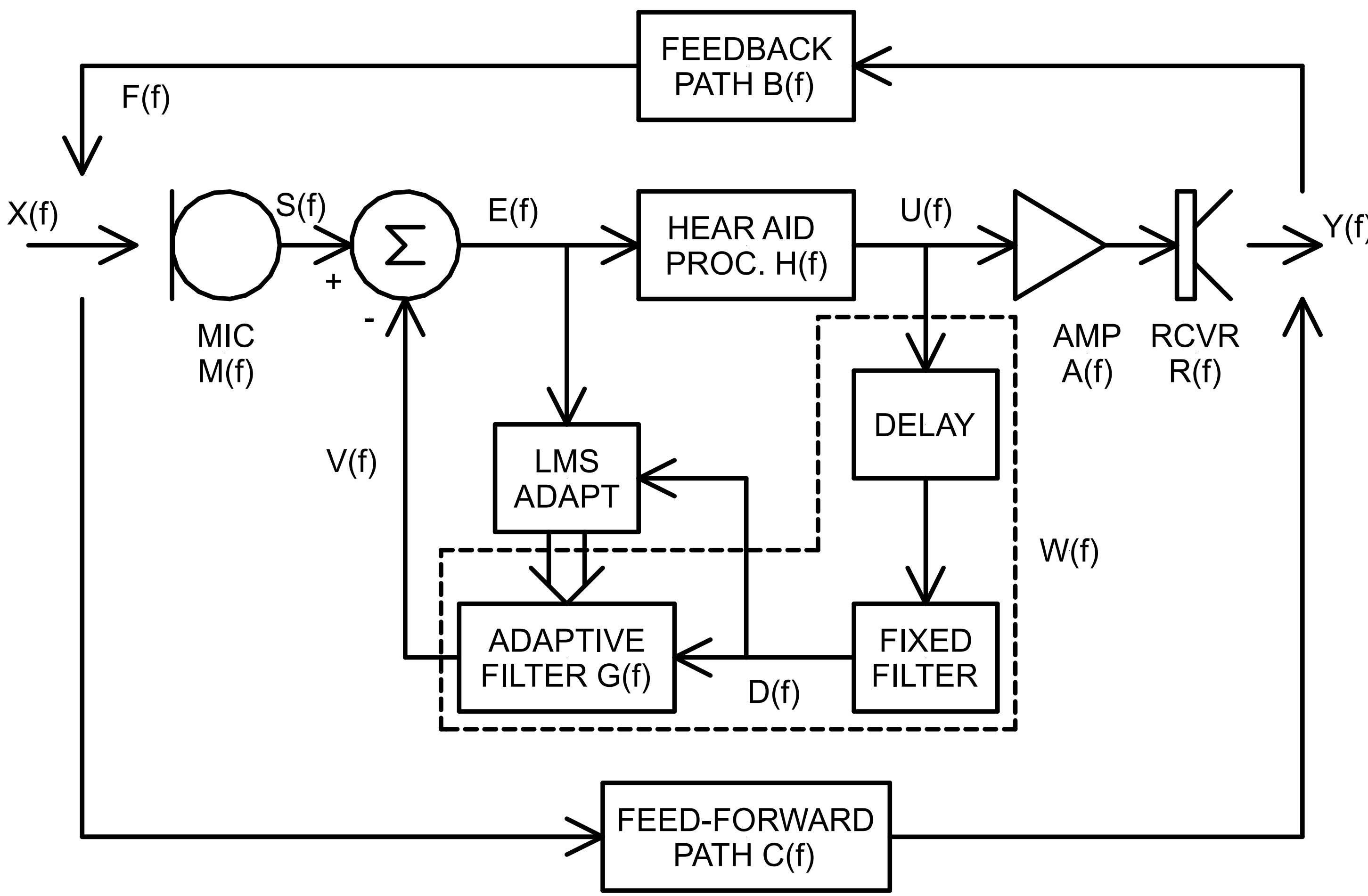
BTE + Telephone Handset



Adaptive Feedback Cancellation

- Model feedback path
  - Adaptive model to minimize system power
  - Subtract model output from microphone signal
- Why adaptive?
  - Feedback path can change over time
  - Track changes
- Limitations
  - Resolution of the feedback path model
  - Ability to track rapid changes
  - Cancellation of tonal inputs
  - Room reverberation

Adaptive FB Cancellation



Feedback Path Model

- Electroacoustics vary slowly
  - Microphone
  - Amplifier
  - Receiver
- Acoustics can change rapidly
  - Telephone handset
  - Hat
  - Earmold shift in the ear canal (jaw movement)
- Model = Fixed in series with adaptive
  - Fixed delay + IIR filter for computational efficiency
  - Short adaptive FIR filter to model changes

System Output

- Output = Processed + Direct

$$Y = X\frac{C + H[WC + MAR]}{[1 - BC] - H\{MARB - W[1 - BC]\}}$$

- Approximate solution for $|BC| << 1$

$$Y \cong X\frac{C + H[WC + MAR]}{1 - H[MARB - W]}$$

- Stability guaranteed for

$$|H[MARB - W]| < 1$$

LMS Adaptation

- Adapt FIR filter coefficients $\{g_k\}$
- Minimize error $\varepsilon(n) = e^2(n)$
- Compare error with processed
  - Error = microphone output - feedback path model
  - Processed = HA output => delay & fixed filter
- Weight update

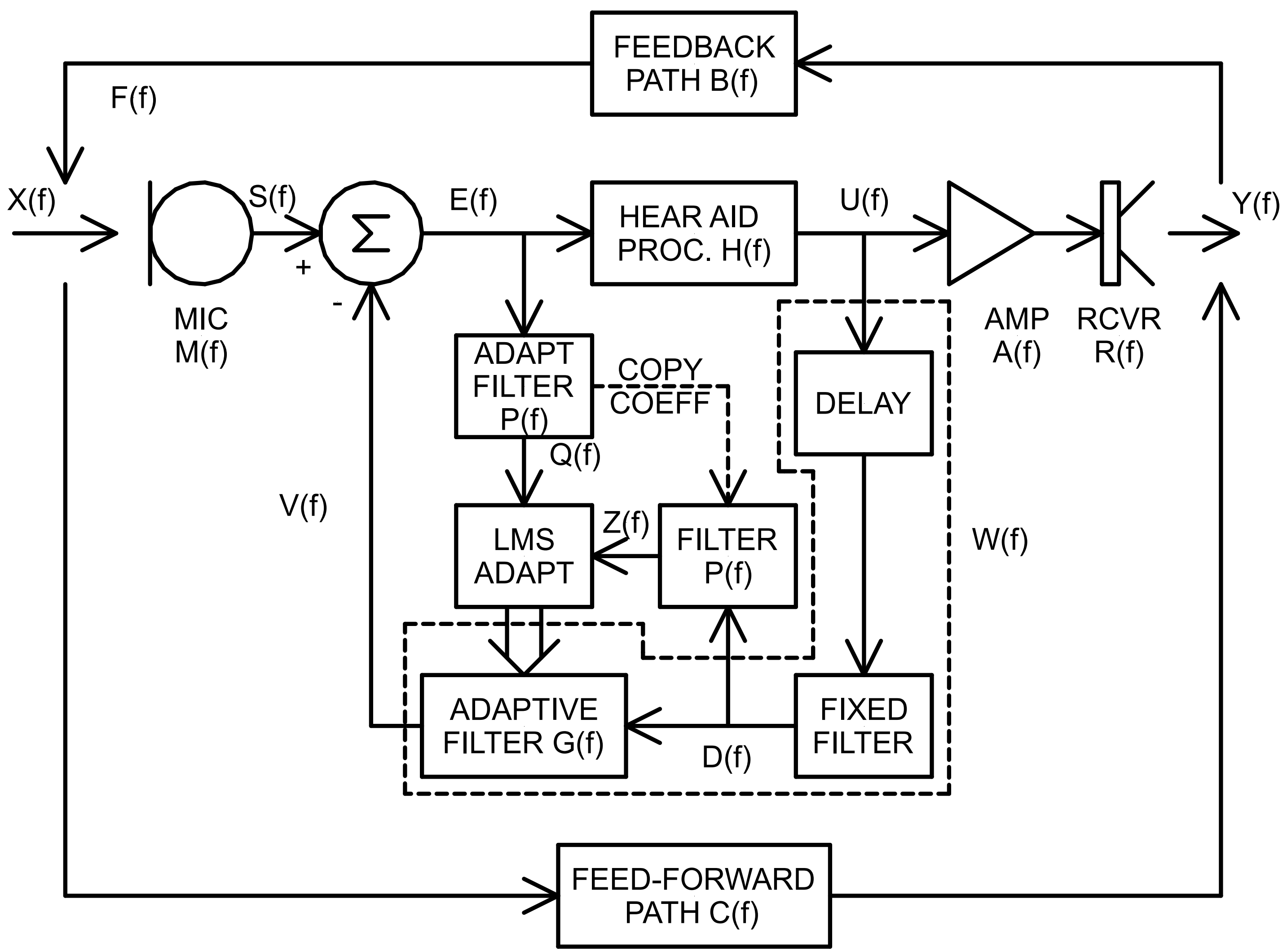
$$g_k(n+1) = g_k(n) + \frac{2\mu}{\sigma_d^2} e(n) d(n-k)$$

  - Proportional to error
  - Trade-off between fast time constant and accuracy

Filtered-X Algorithm

- Objective
  - LMS adaptation works best for white noise input
  - Track changes in the feedback path
  - Prevent response to sinusoidal inputs
- Approach
  - Goal is to decorrelate (whiten) the inputs to the LMS cross-correlation
  - Filter inputs to remove sinusoids
  - Fixed filter based on average signal spectrum
  - Adaptive filter to detect and remove sinusoids

Fig 7.15 Adaptive Filtered-X



Feedback Conclusions

- Feedback cancellation
  - Model the feedback path
  - Subtract model output from microphone signal
  - LMS adaptive filter update
  - Filtered-X approach most effective
- Processing effectiveness
  - Generally get 10 - 15 dB headroom increase
  - Increased gain gives increased audibility
  - Increased stability improves sound quality
- Problems remain
  - LMS tries to cancel tones and narrowband inputs
  - Reverberation
  - Clients don't like initialization probe signal
  - Linear model affected by distortion (e.g. power aids)

Microphones and Arrays

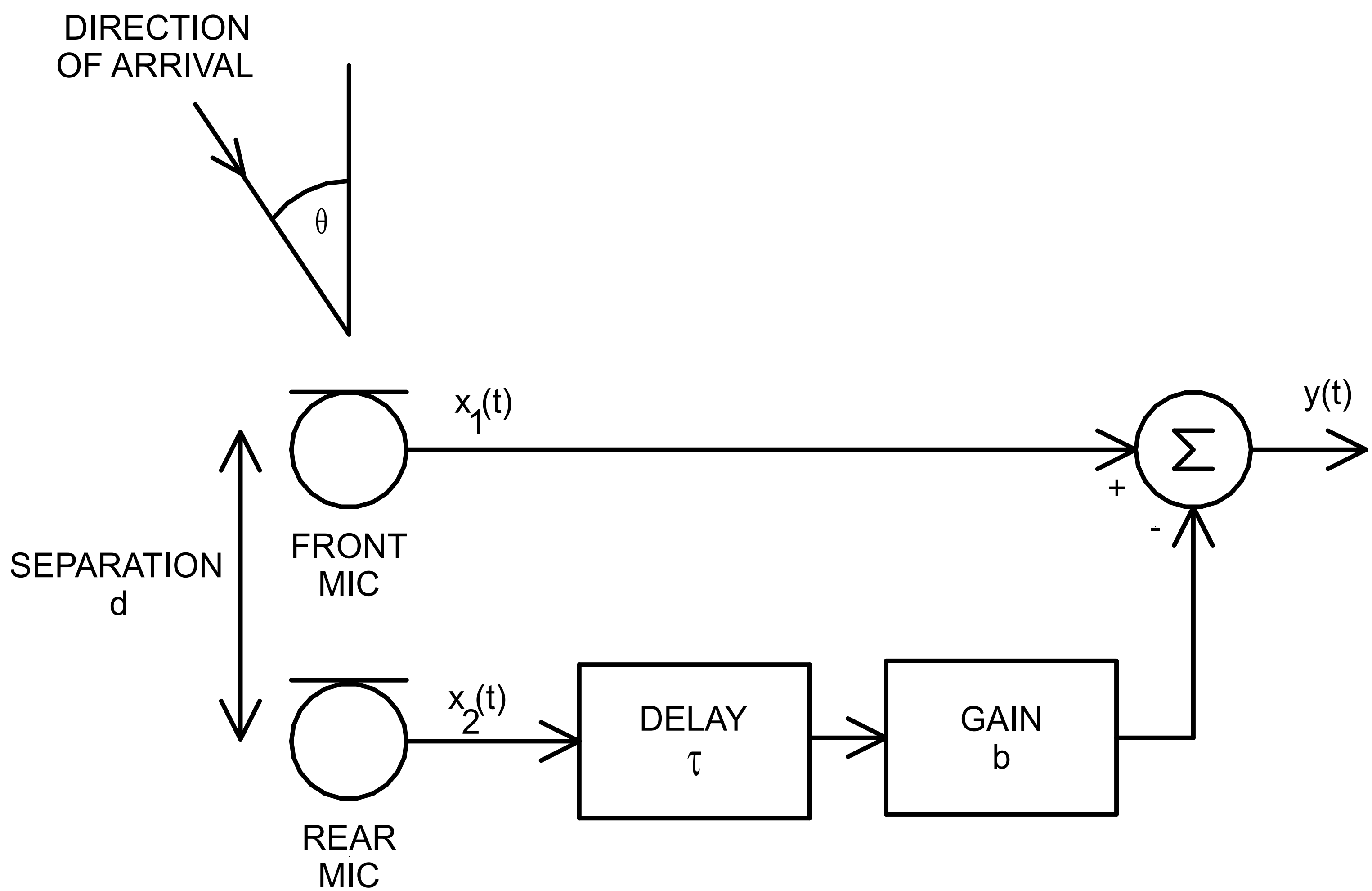
Omnidirectional Microphone

- Single Surface
  - Measure pressure at a point in space
  - Proportional to membrane displacement
- Microphone Technology
  - Moving coil
  - Electret
  - Silicon

Directional Microphone

- Measure pressure at two nearby points
  - Pressure difference
  - Proportional to velocity
- Implementation
  - Pressure difference across a single membrane
  - Difference between adjacent omnidirectional mics
- Control the directional pattern
  - Microphone separation
  - Time delay of rear microphone
  - Relative amplitude of rear microphone

Spatial Signal Processing

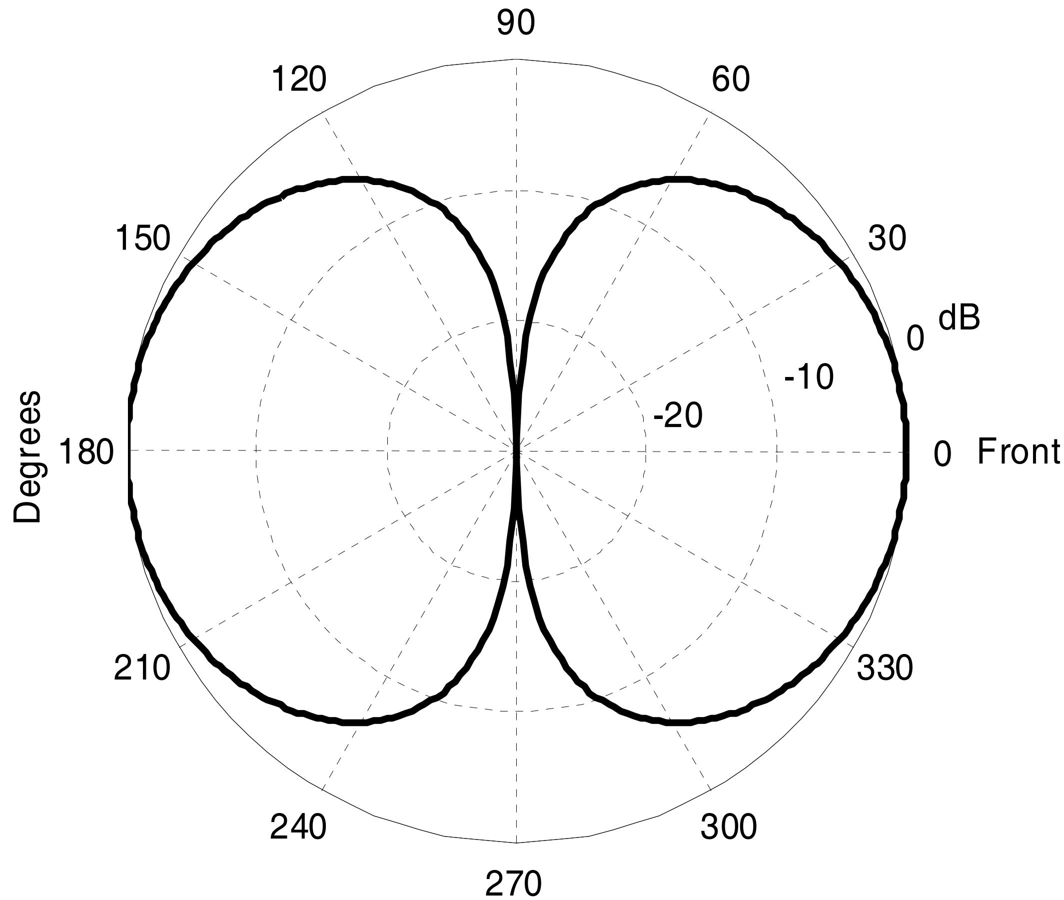


Directional Responses

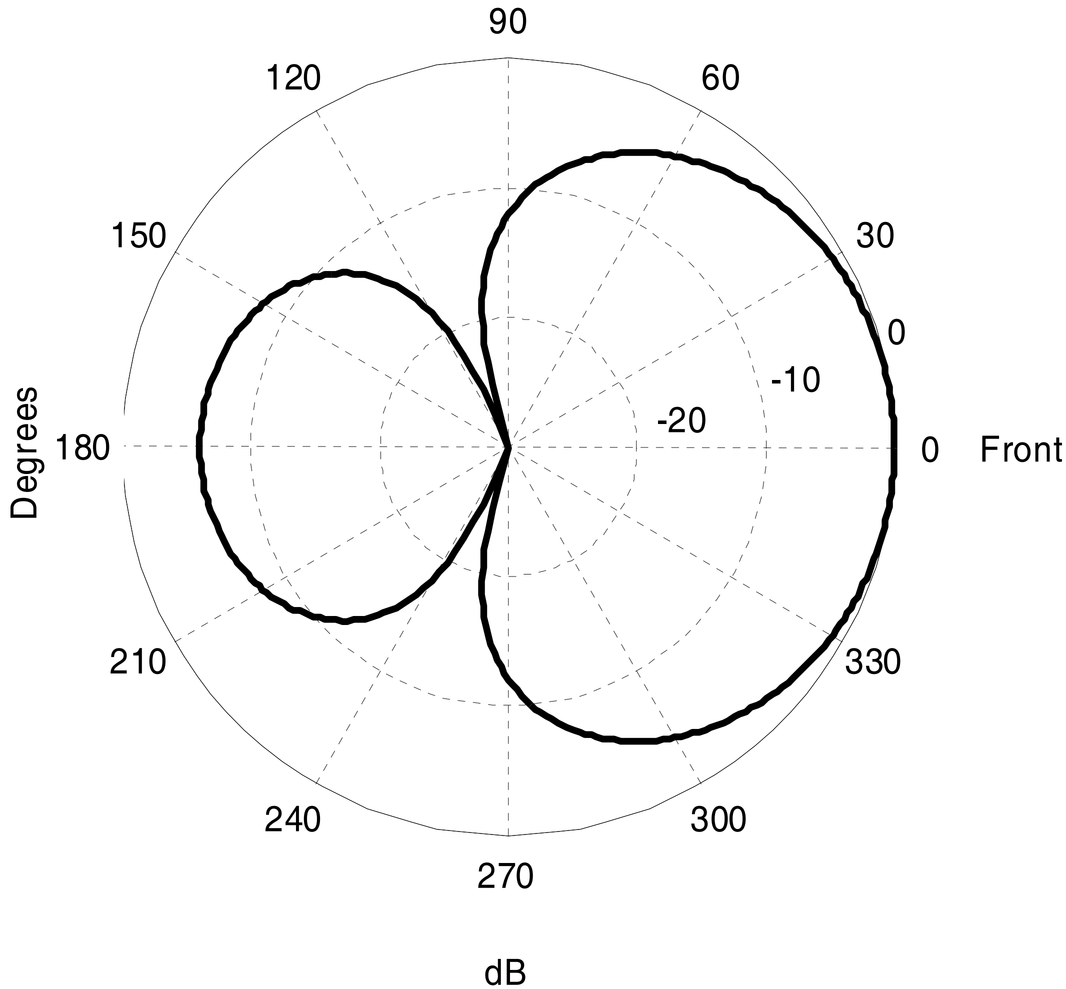
Pattern	Rear Mic Delay τ	a	DI, dB
Omnidirectional	--	1	0
Cosine	0	0	4.8
Hypercardioid	$d/3c$	0.25	6.0
Supercardioid	$2d/3c$	0.4	5.7
Cardioid	d/c	0.5	4.8

The directional response is $R(\theta) = a + (1-a)\cos(\theta)$.

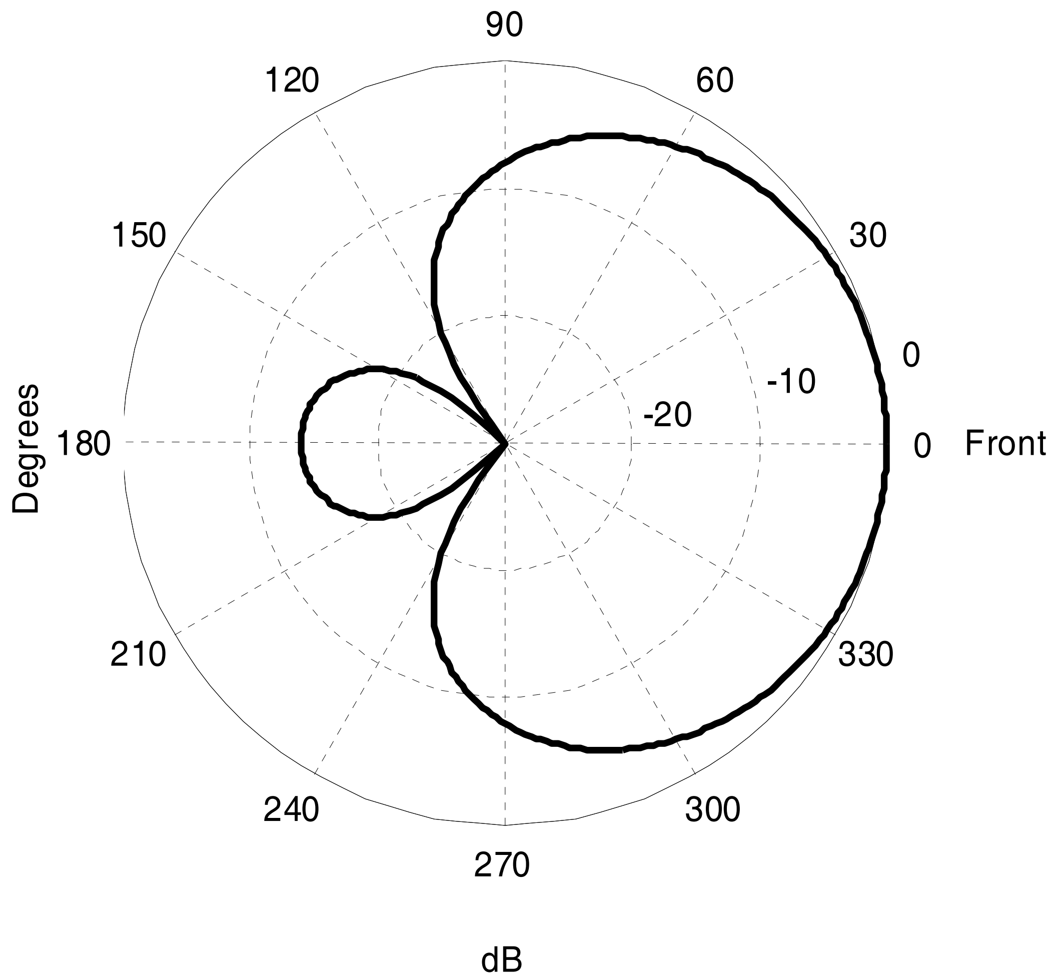
Cosine Pattern



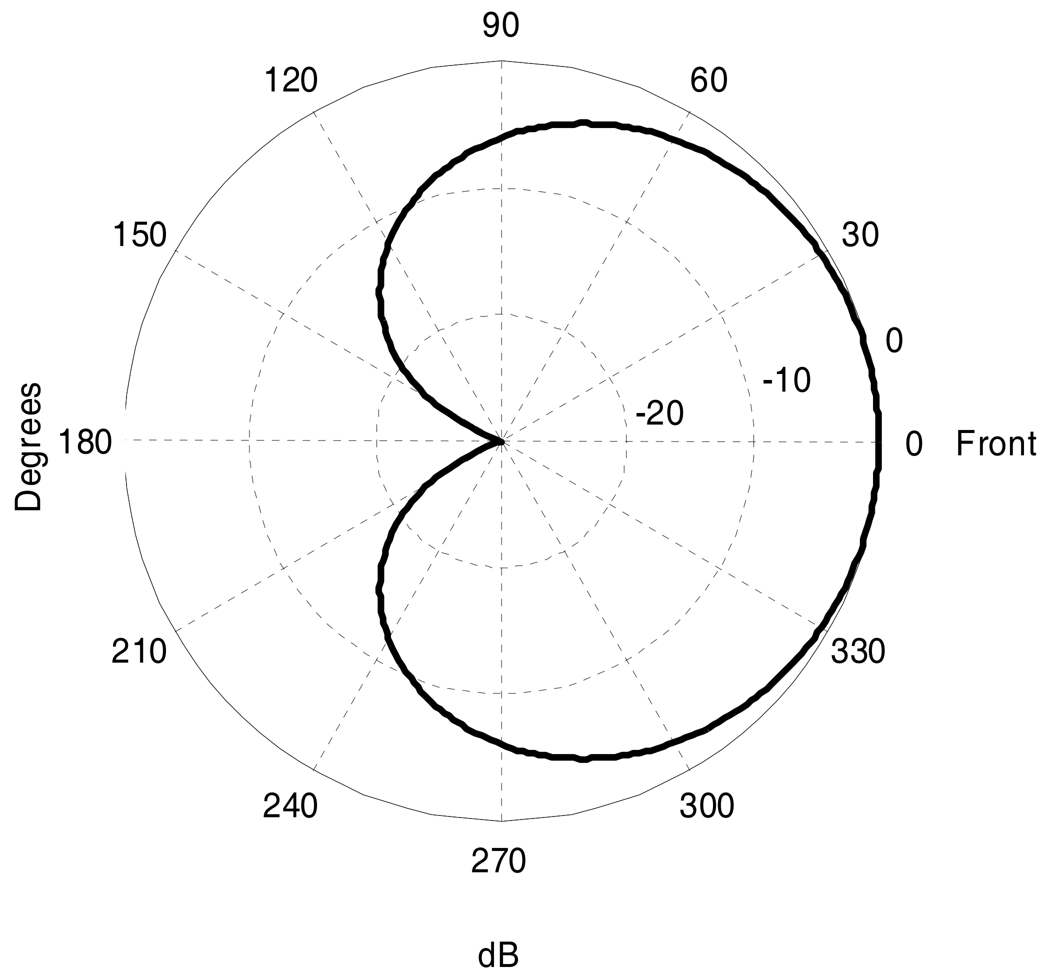
Hypercardioid Pattern



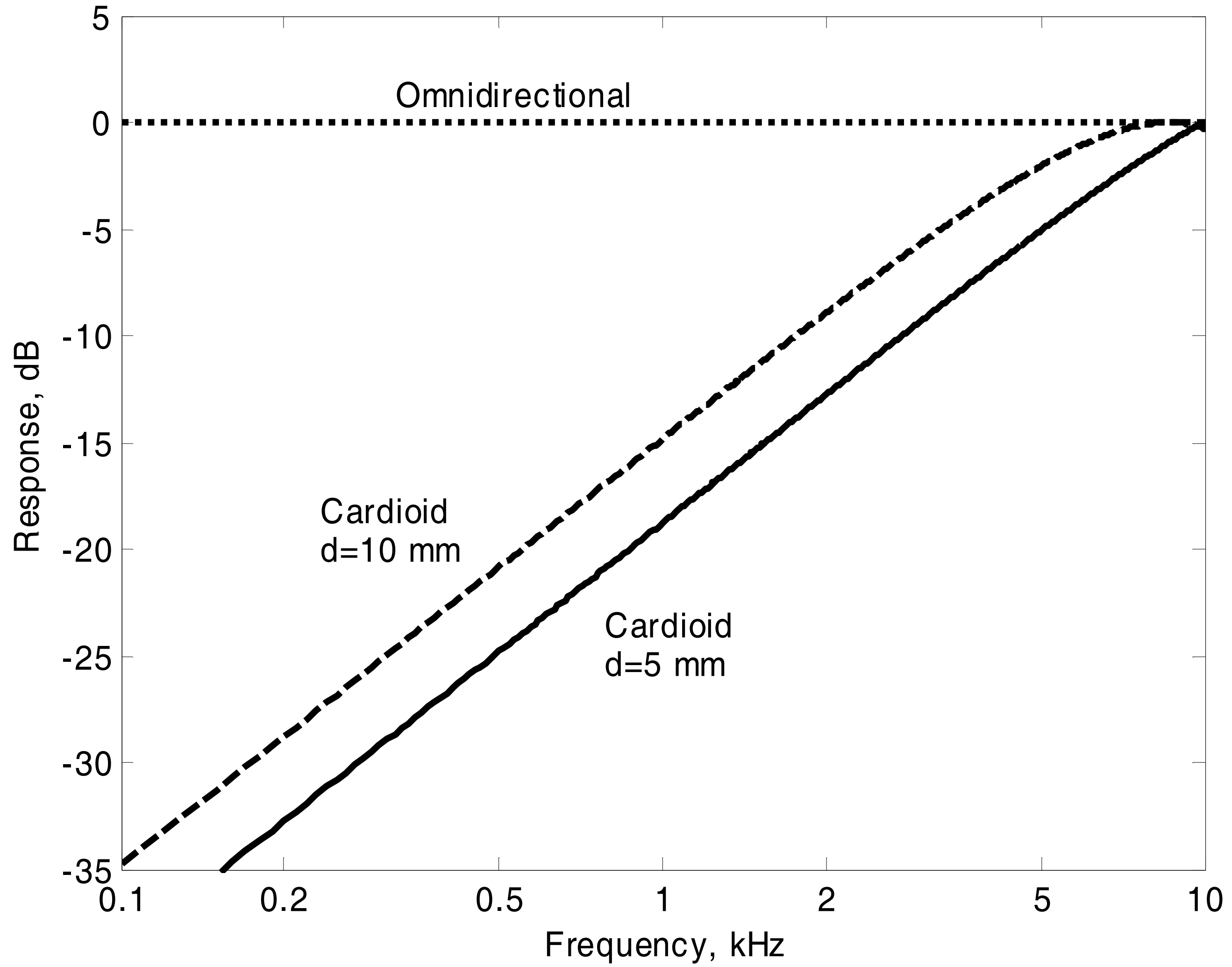
Supercardioid Pattern



Cardioid Pattern



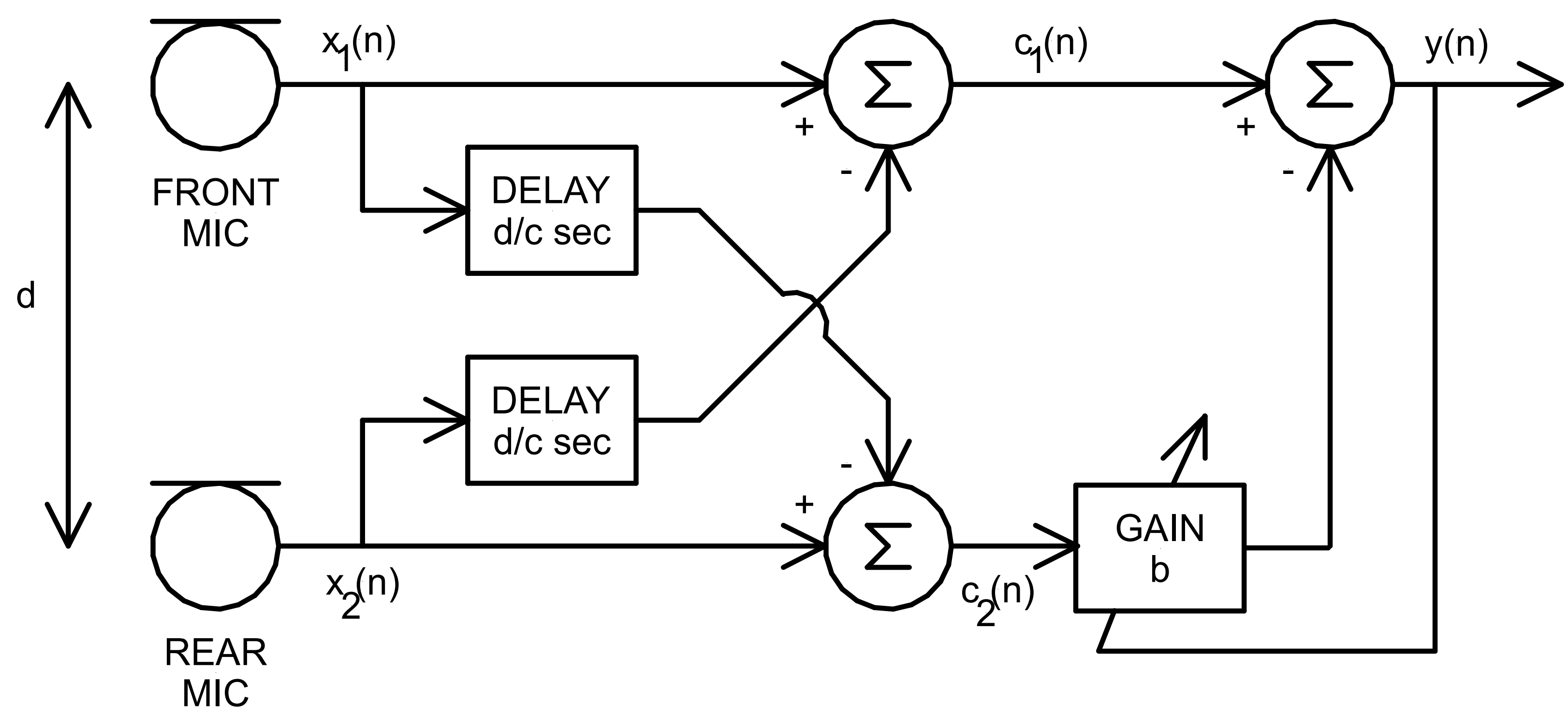
Cardioid Response



Adaptive Microphone Arrays

- Start with 2 microphones
  - Extension of directional microphone technology
  - Can be built into hearing-aid case
- Multi-microphone arrays
  - Combine outputs from multiple microphones
  - Improved directivity (DI) compared to 2-mic array
  - Fixed combination of outputs
  - Adaptive combination of outputs

Cardioids with Adaptive Gain



Adaptive Gain

- Combine mics to give cardioid responses
  - One response faces forward, null at 180 deg
  - Second response faces rear, null at 0 deg
- Cardioid outputs $c_1(n)$ and $c_2(n)$
  - Also give directional patterns

$$c_1(\theta) = \frac{1}{2}[1 + \cos(\theta)] \qquad c_2(\theta) = \frac{1}{2}[1 - \cos(\theta)]$$

  - Front output $c_1(n)$ primarily desired signal
  - Rear output $c_2(n)$ primarily interference

Combined Cardioid Outputs

- Sum cardioid signals: $y(n) = c_1(n) - bc_2(n)$
- Sum directional patterns: $y(\theta) = c_1(\theta) - bc_2(\theta)$
- Null direction depends on *b*
  - Set *b*=0: front cardioid pattern

    $y(n) = c_1(n) \qquad y(\theta) = 1/2[1 + \cos(\theta)]$
  - Set *b*=1: cosine pattern

    $y(n) = c_1(n) - c_2(n)$

$$y(\theta) = c_1(\theta) - c_2(\theta) = 1/2[1 + \cos(\theta)] - 1/2[1 - \cos(\theta)] = \cos(\theta)$$

- Adjust 0 ≤ *b* ≤ 1 to steer null 180 to 90 deg

LMS Adaptive Update

- Adjust b based on rear signal and output

$$b(n+1) = b(n) + \frac{\mu}{\sigma_y^2} c_2(n) y(n)$$

- Interference from side
  - $c_2(n)$ similar to $y(n)$, interference in both signals
  - b is driven towards 1, cancel interference at 90 deg
- Interference from rear
  - Average of $c_2(n)$ times $y(n)$ tends negative
  - b is driven negative
  - Pattern moves towards forward cardioid

Microphone Conclusions

- Summary from Walden *et al* (2003):
  - Directional mics test better than real-world results
  - Environment often limits directional benefit
  - Benefit should be expected only when the signal source is in front of and relatively near the listener and the interference is spatially separated
  - Omnidirectional mode should be the default

Additional Processing Areas

Additional Areas, I

- Electro-acoustic interactions
  - Head and external ear modify microphone input
  - Vent reduces system low-frequency response
  - Occlusion effect: own voice louder
- Wind noise
  - Wind flow over microphone generates turbulence
  - Directional mics especially sensitive
- Spectral enhancement
  - Broader auditory filters reduce spectral contrast
  - Modify short-time spectrum to increase contrast

Additional Areas, II

- Sound classification
  - Change parameters for different environments
  - Identify environments to set optimum processing
- Binaural processing
  - Link the hearing aids at the two ears
  - Low-power short-range communication link
  - Synchronize device parameter settings
  - Compare signals for compression and noise suppression

Conclusions

Conclusions

- Hearing aid = perceptual engineering
  - Designs apply criteria such as MMSE or Max Likelihood
  - These criteria ignore perception
  - Need to consider the human listener as well
- Problems remain
  - Have not solved the basic problem of speech intelligibility in noise
  - Most algorithms are not effective despite having a strong intuitive appeal

- April 2008
- Plural Publishing, San Diego