
MUS424 / EE367D:

Signal Processing Techniques for

Digital Audio Effects

Course Notes

Jonathan S. Abel and David P. Berners
`{abel,dpberner}@ccrma.stanford.edu`

March 30, 2011

CCRMA, Stanford University, Stanford, CA, Spring, 2011

© Copyright 2011 Jonathan S. Abel and David P. Berners. All rights reserved worldwide.

Catalog Description

MUSIC 424/EE 367D. Signal Processing Techniques for Digital Audio Effects—Digital signal processing methods for audio effects used in music mixing and mastering. Topics: dynamic range compression, reverberation and room impulse response measurement, equalization and filtering, delay effects and distortion; digital emulation of analog processors and implementation of time varying effects. Single-band and multiband compressors, limiters, noise gates, de-essers, convolutional and feedback delay network reverberators, parametric and linear-phase equalizers, wah-wah and envelope-following filters, flanging and phasing, distortion. Students develop effects algorithms of their own design.

Prerequisites

An exposure to digital signal processing, including familiarity with the sampling theorem, digital filtering and the Fourier Transform at the level of Music 320 or EE 102B is required. An understanding of digital signal processing at the level provided by Music 420 or EE 264 is helpful. Familiarity with the use of audio effects in mixing and mastering, such as presented in Music 192 is also of benefit. Only a modest amount of Matlab or C programming experience is required for the homework and laboratory exercises.

Meeting Time and Place

Tuesdays and Thursdays, 1:15–2:30 PM, Knoll Classroom.

Instructors

Jonathan S. Abel, abel@ccrma.stanford.edu
David P. Berners, dpberner@ccrma.stanford.edu
Office hours after class and by appointment.

Teaching Assistant

Jorge Herrera, jorgeh@ccrma.stanford.edu
Office hours to be announced.

Grading and Credit

The course is given for three units credit; students successfully completing an optional project will receive one unit additional credit. Projects may be proposed in the first two weeks of May, and are due the day of the final exam.

Grading is based on performance on problem sets, laboratory exercises and midterm and final exams, weighted as follows:

30% Problem Sets and Labs

30% Midterm Exam

40% Final Exam

Collaboration on problem sets and labs is encouraged, but students must write up their submissions individually. Problem set and lab solutions will typically be available at the first class meeting after the due date. The TA will set policy with respect to problem sets and labs, including accepting late problem sets and labs and grading.

Course Materials

Course notes are provided for the photocopying cost, and will be available for purchase at the second lecture. Lecture notes, reading materials and bibliographies will occasionally be provided as handouts and posted to the course web site,

<http://www-ccrma.stanford.edu/courses/424/>

Contents

1. Course Overview	1
Dynamic Range Control	11
2. Definitions; Processing Architectures	11
3. Gain Computation and Level Detection; Analog Detectors	29
4. Applications, Architectures and Improvements	53
Impulse Response Measurement	69
5. LTI Systems, Statistics Review; Impulse Response Measurement	69
6. Golay Code and Allpass Chirp Impulse Response Measurement	83
7. Swept Sinusoid Impulse Response Measurement	101
Reverberation	111
8. Specular Reflections and the Image Method	113
9. The Sabine Theory of Late-Field Reverberation	133
10. Reverberation Statistics and Impulse Response Model	151
11. Reverberation Impulse Response Analysis	155
12. Reverberation Psychoacoustics	167
13. Impulse Response Synthesis and Convolutional Reverberation	169
14. Mechanical and Acoustic Reverberation	179
15. Feedback Delay Network Reverberators	185
Equalization and Filtering	201
16. LaPlace and Fourier Transforms; Bode Plots	201
17. Transform Mechanics	217
18. Bode Plots; Peak and Shelf Filters	227
19. Parametric Sections and Shelf Filters	241
20. Digital Peaking and Shelving Filters, Graphic Equalizer	249
21. Critical Band Smoothing	273
22. Filter Phase, Linear and Minimum Phase	281
23. Frequency Warping; Warped FIR Filter Design	295
24. IIR Filter Design and Prony's Method	303
25. Optimal Filter Design	325
Delay and Filtering	337
26. Parametrized Filters and Discretization	337
27. Time-Varying Filtering	351
28. Distortion Processing	365

1. Course Overview

Introduction

There are typically four steps in producing a CD or movie soundtrack, as shown in Figure 1. In *tracking* sounds are recorded or synthesized and arranged in tracks. The tracks are then processed in the *mixing* stage to form a stereo or multichannel mix. The idea is to arrange the sounds spatially and spectrally, to manipulate their character for artistic purposes, and also to fix problems in the tracks. In *mastering*, subtle adjustments and fixes are made to the mix, and often its dynamic range is limited in preparation for *encoding* and printing on the target medium.

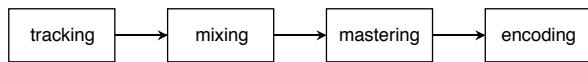


Figure 1: *Audio Production Process*

This class is about audio effects processing, and, in particular, how to build digital versions of the mainline audio effects used in the mixing and mastering stages of audio production. There are four categories of processing commonly employed in mixing and mastering:

- Dynamic Range Control,
- Reverberation,
- Equalization and
- Spatialization.

Other specialized processing, such as delay effects (including echo, chorus, flanging and phasing), distortion, pitch and time stretching and noise removal, is also used. In this class we will explore each of the workhorse processor categories above, and in homework and laboratory exercises you will build examples of each. We will also touch on some of the specialized processors; you may wish to choose one to study as your project.

Dynamic Range Control

Dynamic range control refers to the manipulation of the level of a signal. This may be desirable, for instance, in the case of a singer who moves closer to and further from the microphone while

singing, causing the level to unintentionally rise and fall. It is commonly used to make current pop CDs quite loud, despite their 16-bit integer samples.

As illustrated in Figure 2, dynamic range control may be accomplished using a feed forward architecture in which the signal level is detected and used to compute a gain, which is applied to the input. The gain is computed based on a desired output level as a function of input level.

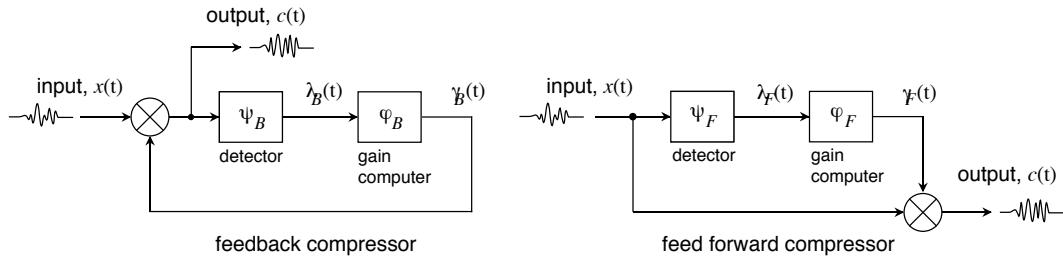


Figure 2: *Feed Forward and Feedback Compressor Architectures.*

Different applications are developed by choice of gain computer and detector. A gain computer which reduces the dynamic range of louder signals results in a *compressor*, and may be used to make a bass or drum track more even. By designing the gain computer to impose a predefined maximum output level, the input is *limited*. A *noise gate*, which eliminates any low-level background noise appearing between notes in the input track, can be implemented by using a gain computer which takes the signal level to zero when the input is small.

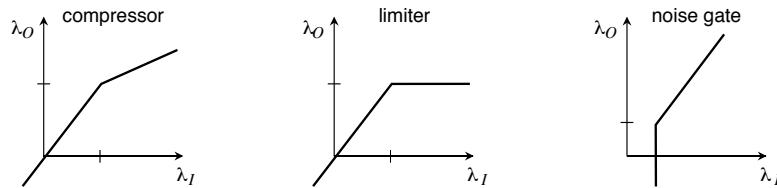


Figure 3: *Compressor, Limiter and Noise Gate Input Level-Output Level Relationships.*

If the detector determines level by examining the input signal over a long time period, the estimated level will change only slowly, and the output will be “transparent”—that is, it will sound much like the input. If the detector only considers a short period of time in evaluating signal level, it can change the envelope of individual notes and the character of the track. Finally, the detector (or the entire architecture, for that matter) can be applied to selected frequency bands, for instance, to reduce loud sibilence.

The dynamic range control unit begins by exploring the notion of signal level and introducing common processing architectures. Signal level detection and gain computation are covered, followed by an overview of applications such as limiting, gating and de-essing, and a presentation of improvements including efficient FIR detection and aliasing elimination. The problem sets will focus on details of detection and gain computation, with the laboratory exercise being to modify a simple compressor.

Reverberation

The acoustics of a space can significantly contribute to the feel of a piece: imagine the “sound” of a jazz hall, a dungeon, outdoors. For this reason it is desired to be able to artificially add environmental cues or *reverberation* to tracks or a mix.

In addition, the tracks of a mix are not often recorded under the same acoustic conditions; consider vocal booths and drum rooms. As a result, artificial reverberation is commonly added to a mix so as to make different tracks feel as if they belong together.

There are two approaches to implementing artificial reverberation in common use today. In one approach, delayed copies of the signal are filtered, mixed and fed back to delay line inputs. The idea is that the process in some ways mimics what happens in actual acoustic spaces, with reflecting surfaces and objects filtering impinging wavefronts and redirecting them to other reflecting surfaces and objects. Loosely speaking, such artificial reverberators are called *feedback delay networks*.

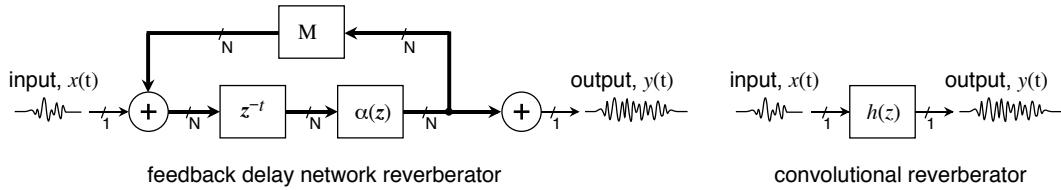


Figure 4: Feedback Delay Network and Convolutional Reverberators.

In the other approach, it is assumed that the acoustic space is approximately linear and time invariant, and can be characterized by its impulse response. Artificial reverberation may then be applied to an input signal simply by convolving it with the desired impulse response. In this approach, impulse responses can be measured and manipulated or synthesized to achieve a desired artistic effect. As an example, Figure 5 shows the time evolution of the impulse response measured between a speaker and microphone in the CCRMA Lobby.

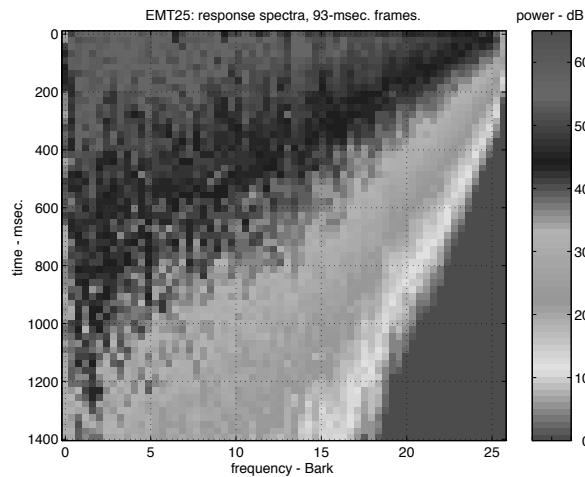


Figure 5: CCRMA Lobby Impulse Response STFT.

In the reverberation unit, the focus is on convolutional reverberation. The unit starts with room impulse response measurement, presenting Golay code and allpass chirp techniques. Reverberation acoustics is then studied, including the image method and Sabine theory. Impulse response analysis and synthesis techniques are covered, and reverberation psychoacoustics is discussed. Problem sets center around reverberation impulse response analysis and synthesis, and the laboratory exercise will be to measure, analyze and manipulate the impulse response of an acoustic space. A convolutional reverberator will be provided to listen to manipulated and synthesized impulse responses. A feedback delay network reverberator will also be provided, and you will learn how to modify it so as to match to sound of a given space or achieve a particular psychoacoustic effect.

Delay Effects and Distortion

There are a number of situations in which a track is purposefully distorted so as to lend a certain character to its sound. It's not unusual, for instance, to send a track in digital form out to an analog tape deck and back to give it a bit of "warmth." In a much less subtle example, guitar amps are often driven to saturation with pleasing results, and guitars tracks are many times recorded clean and distorted appropriately in mixing.

A common architecture for a distortion processor is the cascade of filtering and nonlinear elements, as shown in Figure 6. The nonlinearity has the effect of increasing the bandwidth of its input, and to accommodate the wideband result, it is processed at a high sampling rate. In this portion of the class you will learn how nonlinearities alter the spectral content of a signal, and about antialiasing filter design.

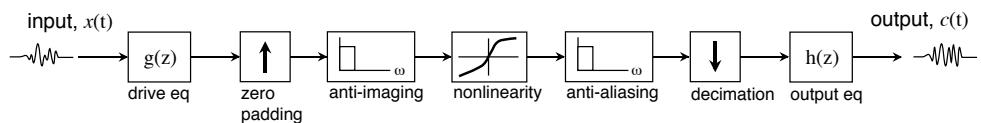


Figure 6: *Distortion Processor.*

Delay effects, including echo, chorus, flanging and phasing, are widely used on keyboards and guitar. They more or less add a series of time-delayed copies of the track to form their output, as shown in Figure 7. If the time delay is sufficiently large, the signal copies will be heard as distinct echoes or a chorus-like sound. When the delay between successive echoes is small enough that it is comparable to the period of a signal in the audio band, the process will be perceived spectrally. In this case a changing time delay results in a changing equalization, as seen in the example in Figure 7. In this portion of the class you will learn how to design fractional and time-varying delays and allpass filters, and will be given a flanger/phaser to modify.

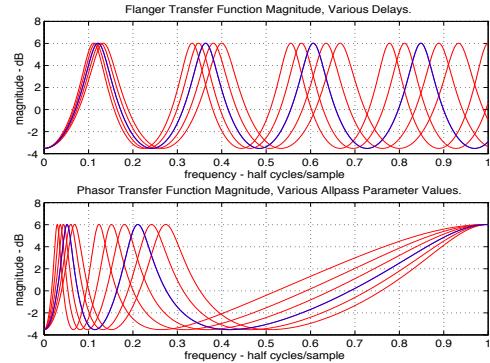


Figure 7: Flanger/Phasor Architecture, Example Transfer Functions.

Equalization and Filtering

Equalization and *filtering* are each the manipulation of the frequency content of a track, with filtering commonly referring to the removal of a band of frequencies and equalization often referring to a more subtle adjustment as a function of frequency.

In many cases equalization and filtering can be used to fix problems with a track. Tracks will sometimes have unwanted frequency components—say, a 60 Hz hum or a rumble from road noise—and filters may be used to remove them. There are other cases where certain frequency components need to be enhanced; for example, a singer's lisp can be corrected somewhat by enhancing high frequencies.

One of the primary uses for equalization on tracks is to help separate different mix elements by having them occupy somewhat different frequency bands.

Equalization may also be used as an effect. Filtering the waveform to a band between 200 Hz and 3200 Hz (in combination with some other processing) makes the track sound as if it is being played through a telephone. In an architecture similar to that of the feed forward compressor above, the signal level may be detected and used to control a filter. In this way, for instance, the onset of a note can be made to have a very different timbre than its decay.

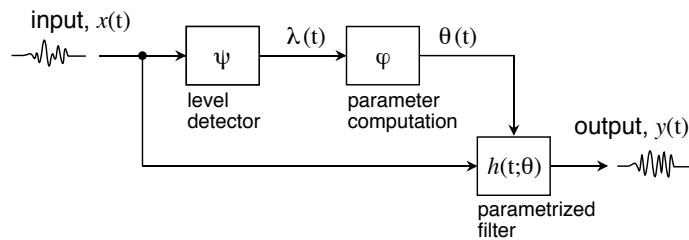


Figure 8: Envelope Following Filter.

It turns out that across a given genre, there is surprisingly little variation in overall spectral content. As a result, the mastering engineer will often (and probably subconsciously) use a program

equalization which brings the power spectrum of the song in line with the standard.

Audio engineers have a preference for analog equalizers, for their transfer function characteristics, their controls and their signal path. The approach to equalization we present here is based on modeling analog equalizers to fix the transfer function and controls, and uses numerically robust filter structures so as to maintain signal integrity.

In the equalization and filtering unit, we first concentrate on peaking and shelving filters which are widely used in mixing and mastering, and which provide excellent building blocks for forming more complicated equalizers. Equalization psychoacoustics is then discussed, including presentation of the Bark and ERB frequency scales, and critical band smoothing. Linear phase and minimum phase filtering are covered, as are techniques for IIR filter design. The problem sets will cover IIR filter design techniques, and the laboratory exercise will include implementing a parametric section and maybe an envelope-following filter.

Panning and Spatialization

In the presence of a stereo or surround output channel configuration, it is possible to position tracks spatially. This provides an important dimension along which tracks may be separated.

The primary technique used to position sounds for stereo or multichannel playback is called *panning*, and places different portions of the signal in the different output channels. In this class we will briefly discuss techniques for determining the panning weights, and might also look at HRTF techniques popular for video games and other interactive environments.

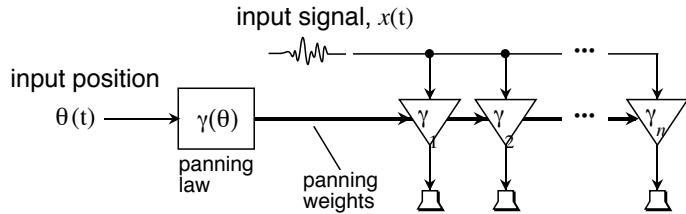


Figure 9: *Multichannel Panning*.

Course Outline

We plan to cover the following topics.

1. Dynamic Range Control (Compression)
 - Definitions; Processing Architectures
 - Detection and Gain Computation; Analog Detectors
 - Applications, Architectures and Improvements
2. Reverberation and Impulse Response Measurement
 - LTI Systems, Statistics Review

- Golay Code, Allpass Chirp Impulse Response Measurement
- Specular Reflections and the Image Method
- Reverberation Analysis and Psychoacoustics
- Reverberation Acoustics and Impulse Response Synthesis
- Low-Latency Convolution

3. Delay and Distortion Processing

- Fractional Sample and Time Varying Delay
- Echo, Chorus, Flanging; Phasing
- Sampling Rate Conversion and Antialiasing Filter Design
- Distortion Processing

4. Equalization and Filtering

- z-Plane, s-Plane and Fourier, Laplace relationships
- Parametric Sections and Shelf Filters; Optimal Filters
- Filter Phase, Linear and Minimum Phase
- Critical-Band Smoothing, Bark and ERB Frequency Scales
- Frequency Warping
- IIR Filter Design and Prony's Method
- Time-Varying and Envelope Filters

5. Panning and Spatialization

- Stereo and Multichannel Panning
- Spatial Hearing and 3D Audio

Dynamic Range Control

2. Definitions; Processing Architectures

Signal Properties

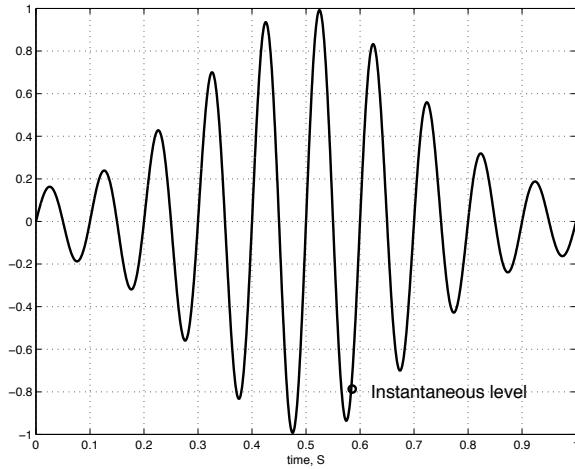


Figure 10: *Instantaneous Signal Level*.

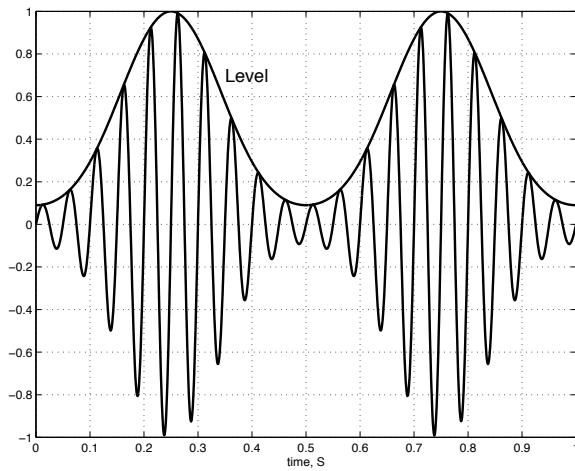


Figure 11: *Signal Envelope Level*.

- instantaneous level $\in \mathcal{R}$

- $\text{level} \geq 0$
- bandwidth of level lower than signal bandwidth

Dynamic Range: difference in dB between highest and lowest levels

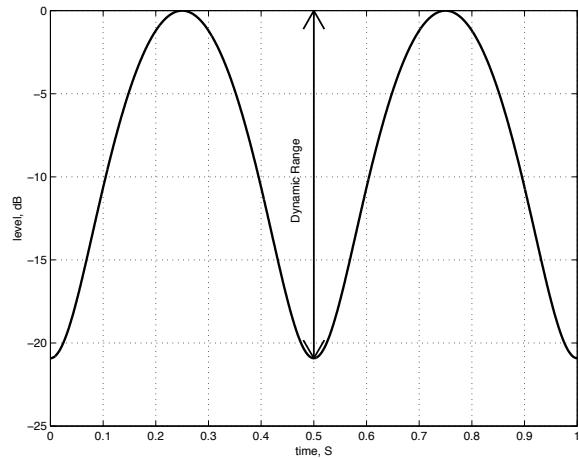


Figure 12: *Dynamic Range of Signal.*

- dB level defined as $20 \cdot \log 10(\cdot)$

What does Compression do to a signal?

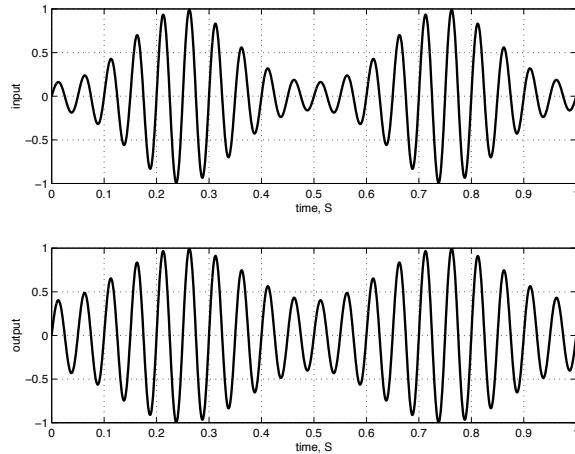


Figure 13: *Signal and Compressed Signal.*

- change dynamic range (compression shown - reduces dynamic range)

Dynamic Range Control Applications

Compression:

- channel/environment has small dynamic range
 - audio in cars
 - recording to tape (hiss)
 - radio broadcast (reduce transmitted bandwidth)
- increase perceptual loudness
- reduce unwanted changes in recording level
- alter envelopes of individual notes
- change timbre of sounds

Expansion:

- suppress noise floor (noise gate or *downwards expansion*)
- 'undo' previous overzealous use of compression on recording
- restore dynamics to signal coming out of channel with reduced dynamic range
 - Dolby, dbx noise reduction for tape
- add dynamic range for perceptual effects

Compressor followed by Expander = Compander

- tape
- digital fx with low bit depth
- analog delays

Signal Detection

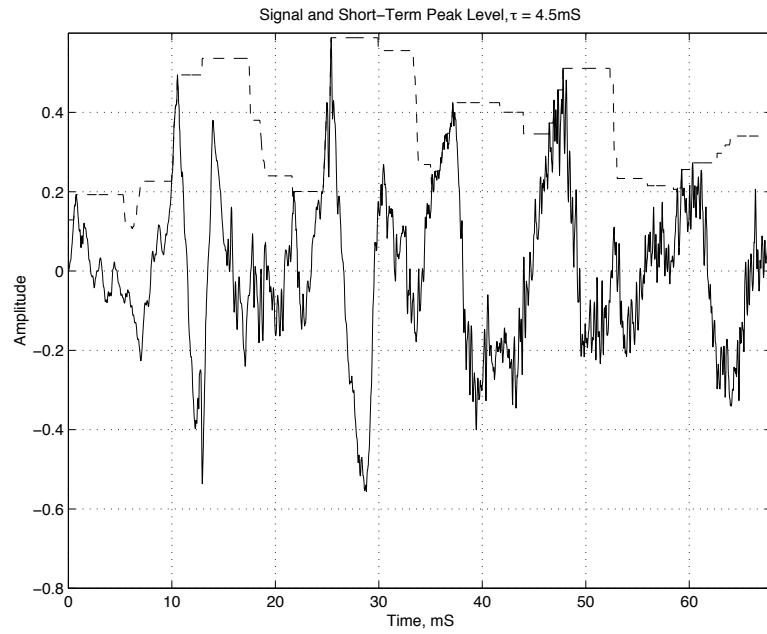


Figure 14: *Signal (solid) and Short-Term Peak Value (dashed), $\tau = 4.5\text{mS}$.*

$$\lambda_\infty(x(t); \tau) = \max_{t-\tau < \xi < t} |x(\xi)|$$

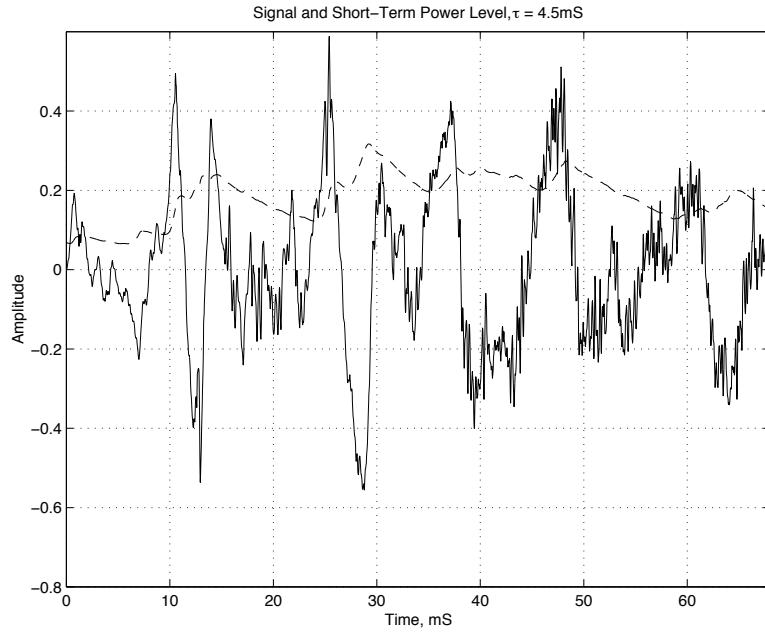


Figure 15: *Signal (solid) and Short-Term Average Root-Power Value (dashed), $\tau = 4.5\text{mS}$.*

$$\lambda_2(x(t); \tau) = \sqrt{x^2(t) * \left[\frac{e^{(-t/\tau)}}{\tau} \cdot u(t) \right]}$$

as $\tau \rightarrow \infty$

$$\lambda_2^* = \left[\frac{1}{\xi} \int_{-\xi/2}^{\xi/2} x^2(t) dt \right]^{1/2}, \quad \xi \rightarrow \infty.$$

For periodic signal with period T ,

$$\lambda_2^* = \left[\frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt \right]^{1/2}.$$

Example:

Triangle wave, 0.25 Hz

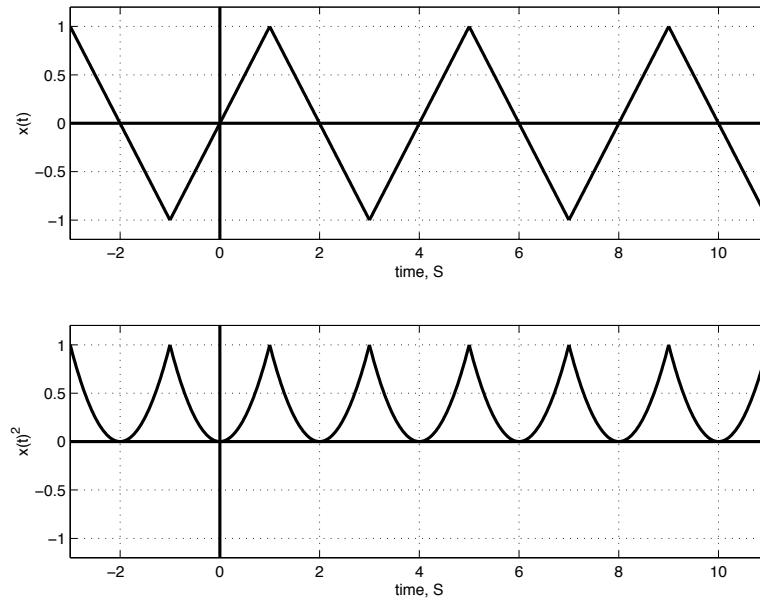


Figure 16: *Triangle wave and squared waveform.*

by symmetry,

$$\lambda_2^* = \left[\int_0^1 t^2 dt \right]^{1/2} = \frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{3}.$$

Short-Term RMS

- convolution with “leaky integrator”

$$\lambda_2(x(t); \tau) = \sqrt{x^2(t) * \left[\frac{e^{(-t/\tau)}}{\tau} \cdot u(t) \right]}$$

$$x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau$$

for causal $h(t)$, $h(t) = 0 \quad \forall \quad t < 0$.

$$x(t) * h(t) = \int_{-\infty}^t x(\tau)h(t - \tau)d\tau$$

- leaky integrator “forgets” old data

contrast with true integration: $\tau \rightarrow \infty$, $h(t) \rightarrow u(t)$. then,

$$x(t) * h(t) = \int_{-\infty}^t x(\tau)d\tau$$

Discrete-Time Implementation of Leaky Integrator
true integration of $x(n)$:

$$y(n) = \sum_{-\infty}^{n_1} x(n)$$

Implement as filter:

$$y(n_1) = y(n - 1) + x(n)$$

$$Y(z) = z^{-1}Y(z) + X(z)$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - z^{-1}}$$

- pole at DC, unstable filter

Now, make “leaky” discrete-time filter:

$$y(n) = a_0y(n - 1) + b_0x(n)$$

$$Y(z) = a_0z^{-1}Y(z) + b_0X(z)$$

$$H(z) = \frac{b_0}{1 - a_1z^{-1}}$$

which is stable for $|a_1| < 1$.

normalization:

$$b_0 = 1 - a_1$$

- This choice makes integrated impulse response have unity magnitude:

we want $\sum h(n) = 1$, where $h(n) = b_0 \cdot a_1^n$.

We can use

$$\sum_{n=0}^{\infty} a_1^n = \frac{1}{1 - a_1}$$

or, evaluate $H(z)$ at DC:

$$\frac{b_0}{1 - a_1} = 1$$

- filtering for an RMS detector is a linear process!

Pole placement for leaky integrator

- can use impulse-invariance:

$$h(n) = a_1^n$$

For sampling rate f_s , samples are spaced at $T = 1/f_s$. Thus, we set $a_1 = e^{-1/(\tau f_s)}$.

Discrete-Time Peak Detector Implementation

- Difficulties with Implementing Peak Sample-Hold
 - expensive max function: $25\text{mS} = 1100 \text{ samples at } 44.1\text{kHz}$ sampling rate
 - output is discontinuous
- implement “leaky” peak detector $\hat{\lambda}_\infty$
 - precedent with analog compressors

During Attack:

$$\frac{d}{dt} \hat{\lambda}_\infty = \frac{1}{\tau_a} \cdot (|x| - \hat{\lambda}_\infty), \quad |x| > \hat{\lambda}_\infty$$

During Release:

$$\frac{d}{dt} \hat{\lambda}_\infty = \frac{1}{\tau_r} \cdot (|x| - \hat{\lambda}_\infty), \quad |x| < \hat{\lambda}_\infty,$$

- reacts quickly if signal is above $\hat{\lambda}_\infty$
- reacts slowly if signal is below $\hat{\lambda}_\infty$

Motivation for asymmetry

- keep $\hat{\lambda}$ near peaks
- keep $\hat{\lambda}$ “smooth” while still being able to catch peaks

Update Equations for Leaky Peak Detector

- similar to RMS detector one-pole model, can use first-order differencing
- because of asymmetry in attack and release behavior, this is not a linear system, and cannot be implemented as a single signal-independent filter!

update:

if $x(n) > \hat{\lambda}_\infty$:

$$\hat{\lambda}_\infty = \hat{\lambda}_\infty + (1 - e^{(-1/\tau_a f_s)})(|x(n)| - \hat{\lambda}_\infty)$$

else

$$\hat{\lambda}_\infty = \hat{\lambda}_\infty + (1 - e^{(-1/\tau_r f_s)})(|x(n)| - \hat{\lambda}_\infty)$$

Variations:

- always release, release to zero:

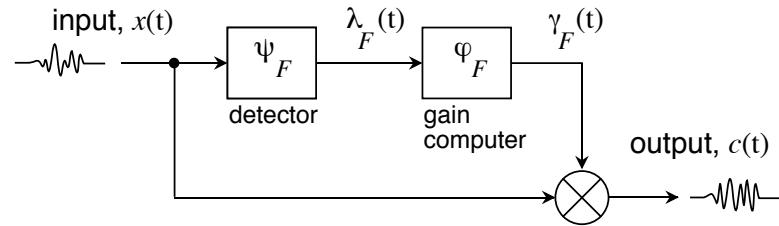
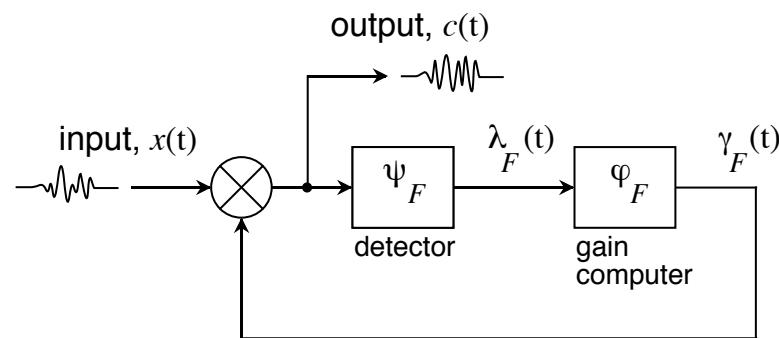
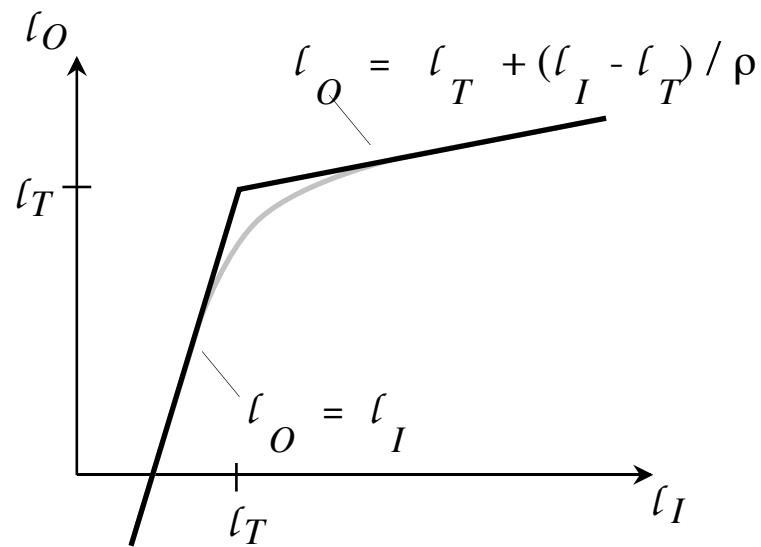
$$\hat{\lambda}_\infty = \hat{\lambda}_\infty - (1 - e^{(-1/\tau_r f_s)})\hat{\lambda}_\infty$$

if $x(n) > \hat{\lambda}_\infty$:

$$\hat{\lambda}_\infty = \hat{\lambda}_\infty + (1 - e^{(-1/\tau_a f_s)})(|x(n)| - \hat{\lambda}_\infty)$$

- combine with sample/hold peak detector

Gain Computation

Figure 17: *Feedforward architecture.*Figure 18: *Feedback architecture.*Figure 19: *Static Compression Curve.*

- static compression curve relates steady-state input and output levels
- sometimes called “transfer function” of compressor
- in linear space, gain = output level / input level.
- in log space, $g = \ell_O - \ell_I$, where $g = 20 \log 10(\text{gain})$, $\ell_O = 20 \log 10(\text{outputlevel})$, and $\ell_I = 20 \log 10(\text{inputlevel})$.
- below threshold (ℓ_T), gain is constant (=1)
- above threshold: 1dB increase in input produces a $1/\rho$ dB increase in output, where ρ is the compression ratio.

Feedforward gain function required for compression ratio ρ (true for $\ell_I > \ell_T$):

$$\begin{aligned}\ell_O &= \ell_T + (\ell_I - \ell_T)/\rho \\ g(\ell_I) &= \ell_O - \ell_I = (\ell_I - \ell_T)/\rho + \ell_T - \ell_I \\ g(\ell_I) &= (\ell_I - \ell_T)(1/\rho - 1) \\ \Phi_F &= \left(\frac{\lambda_I}{\lambda_T} \right)^{(1/\rho-1)}\end{aligned}$$

where $\Phi_F = 10^{(g/20)}$, and λ_I , λ_T are the linear input and threshold levels.

Feedback gain function required for compression ratio ρ (true for $\ell_O > \ell_T$):

$$\begin{aligned}\ell_I &= \ell_T + (\ell_O - \ell_T)\rho \\ g(\ell_O) &= \ell_O - \ell_I = \ell_O - \ell_T - (\ell_O - \ell_T)\rho \\ g(\ell_O) &= (\ell_O - \ell_T)(1 - \rho)\end{aligned}$$

$$\Phi_B = \left(\frac{\lambda_O}{\lambda_T} \right)^{(1-\rho)}$$

where $\Phi_B = 10^{(g/20)}$, and λ_O, λ_T are the linear input and threshold levels.

Example: Compressor Attack / Release

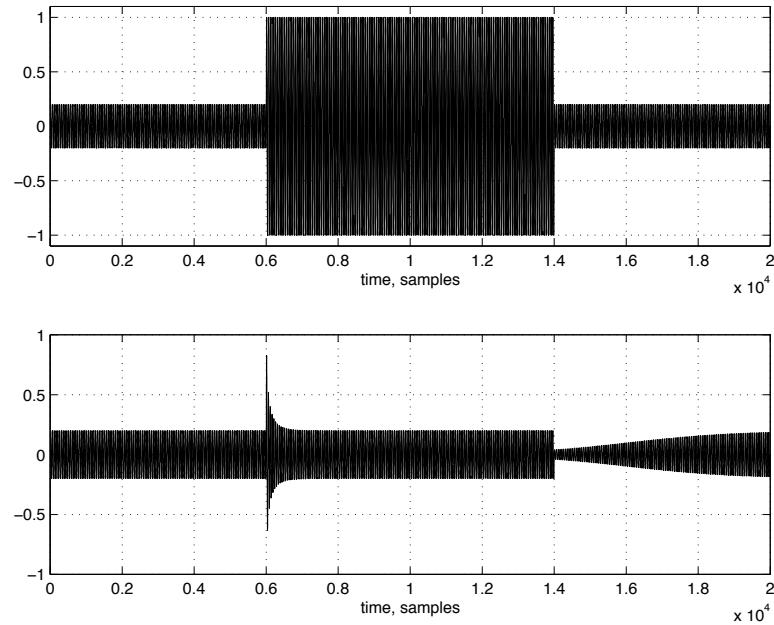


Figure 20: *Compressor Input and Output.*

- this is a limiter, $\rho = \infty$.
- attack is faster than release: $\tau \approx 10mS, \tau_r \approx 100mS$.

Perceptual Ranges for Attack and Release Times

- attack, release times affect perceptual quality of compression:

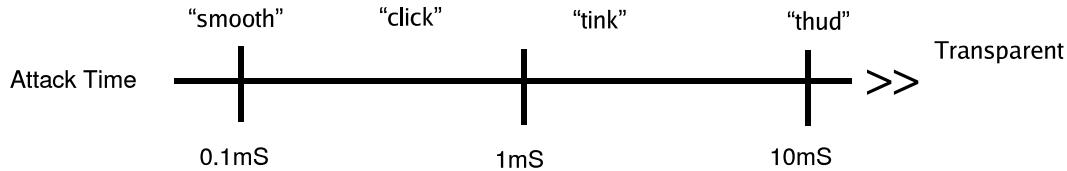


Figure 21: *Perceptual qualities of attack times.*

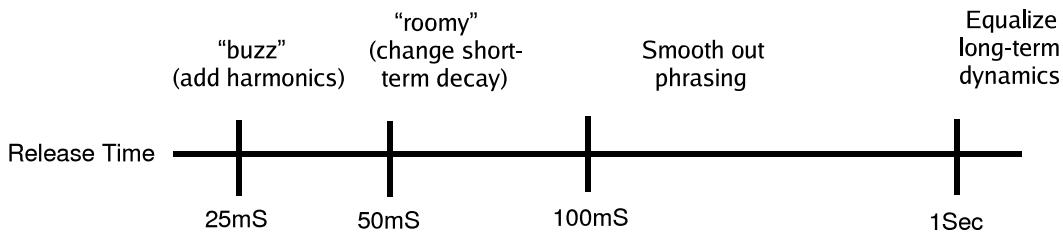


Figure 22: *Perceptual qualities of release times.*

3. Gain Computation and Level Detection; Analog Detectors

Static Compression Curves

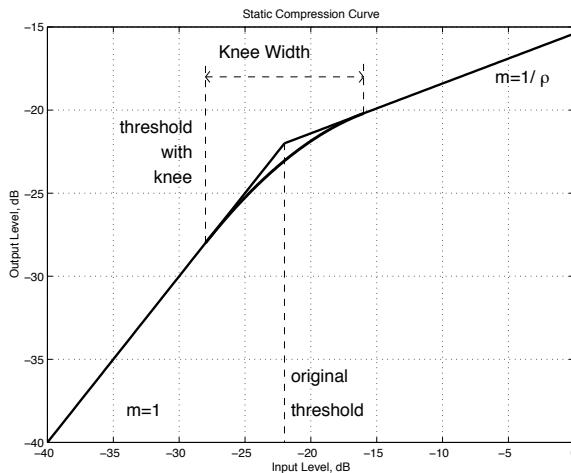


Figure 23: *Static Compression Knee*.

For 'hard' knee:

$$\begin{aligned} \ell_I < t &\rightarrow m = 1 \\ \ell_I > t &\rightarrow m = 1/\rho \end{aligned}$$

one possible strategy:

- make static curve continuous with several continuous derivatives

idea:

- traversing curve should be “smooth”.

Example:

- Use parabolic segment for region around threshold

- parabolic form gives three degrees of freedom:

$$\ell_O = a\ell_I^2 + b\ell_I + c$$

- at left edge of curve $m = 1$.
- at right edge of curve $m = 1/\rho$.
- at left edge of curve $\ell_I = \ell_O$

$$\ell_O = \ell_I - a(\ell_I - T)^2$$

for knee 'width' w ,

$$\begin{aligned} \frac{d}{d\ell_I} \ell_O &= 1 - 2a \cdot (\ell_I - T) \\ 1 - 2a \cdot w &= 1/\rho \\ a &= \frac{1 - 1/\rho}{2w} \end{aligned}$$

Why smooth the knee?

- Reducing BW of the gain function $\Phi(t)$ minimizes amount of harmonic and inharmonic distortion introduced by the dynamic range control.

Example:

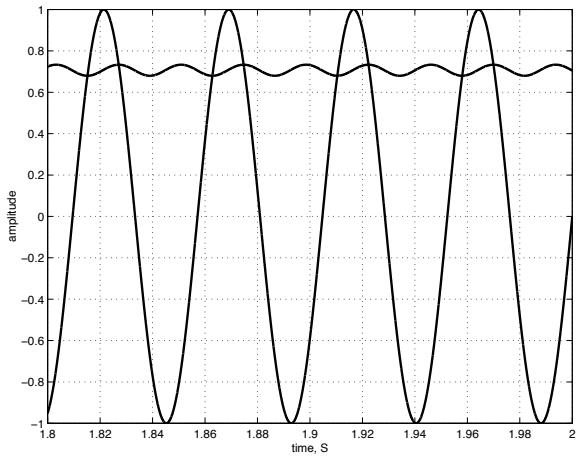


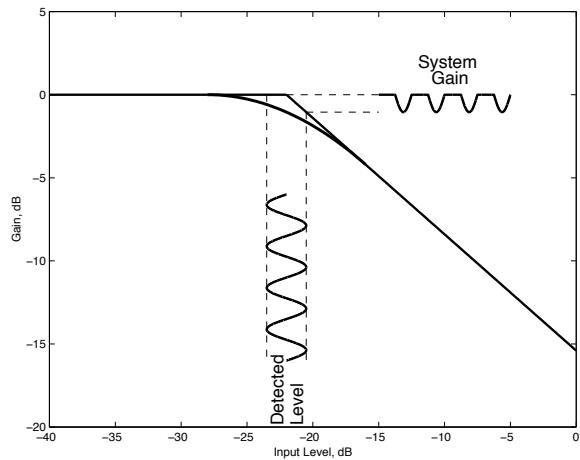
Figure 24: *Input signal and RMS detector output.*

for RMS detector:

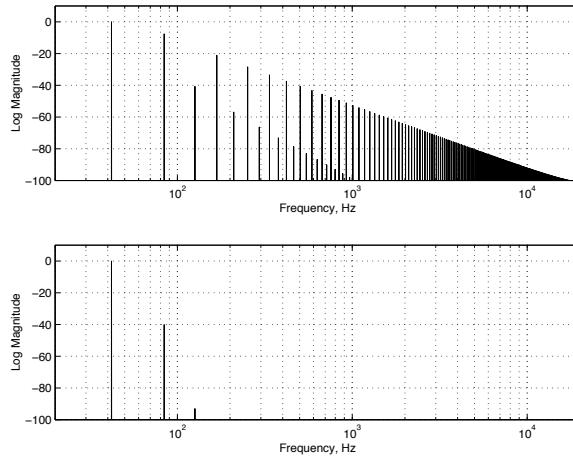
$$\lambda(t) = \sqrt{l + a \cdot \sin(2\omega t + \phi)}$$

- squared detector output is sinusoidal because detector is a linear system acting on the squared input.
- frequency doubles because detector senses squared input:

$$\cos^2(\omega t) = \frac{1}{2} + \frac{1}{2} \cos(2t).$$

Figure 25: *Detected level and resulting output gain.*

- with same level detector, output gain depends on shape of static knee.

Figure 26: *Spectra for hard- and soft- knee, with same detection.*

- soft-knee produces $\Phi(t)$ with lower bandwidth

Time-varying gain adds BW to signal

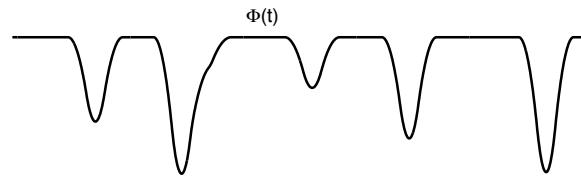


Figure 27: *Gain function $\Phi(t)$ and corresponding spectrum $\Gamma(\omega)$.*

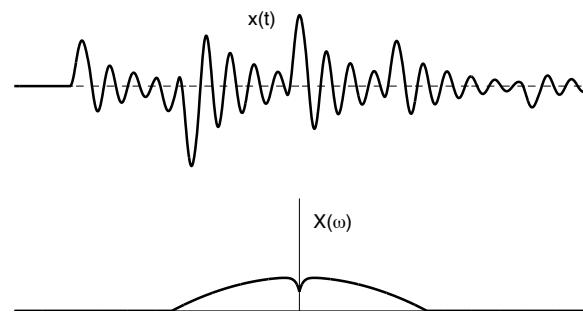


Figure 28: *Input signal $x(t)$ and corresponding spectrum $X(\omega)$.*

$$x(t) \cdot \Phi(t) \iff X(\omega) * \Gamma(\omega)$$

- spectral signal components are 'smeared' by $\Gamma(\omega)$.

Special Case: Sinusoidal input $x(t) = A\cos(\omega_0 t)$

- gain $\Phi(t)$ will have energy at *even* harmonics of ω_0 .
- → harmonic distortion at *odd* harmonics of ω_0 .

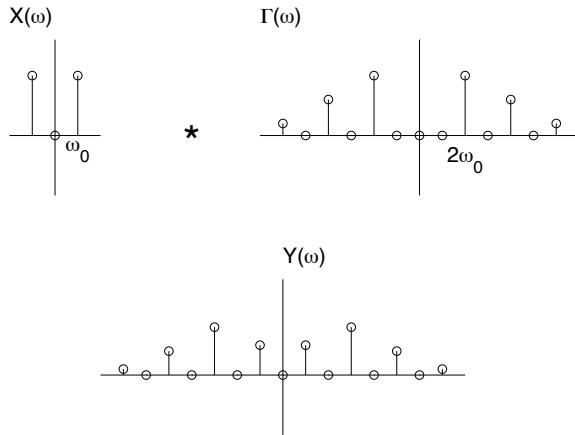


Figure 29: Spectra of input, detector output for sinusoidal input.

If input $x(t)$ has a continuous spectrum:

- $\Phi(t)$ will likely have a continuous spectrum
- input spectrum $X(\omega)$ will be 'smeared' by $\Gamma(\omega)$
- result: harmonic and *inharmonic* distortion

If input $x(t)$ is dominated by one frequency ω_1 :

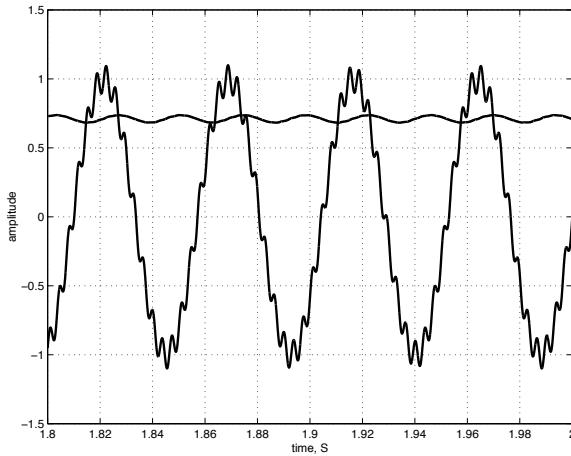


Figure 30: *Input, detector signals for input with one dominant frequency.*

- $\Phi(t)$ will be dominated by even harmonics of ω_1
- harmonic and inharmonic distortion
- sum, difference tones

Discrete-time implementation

- aliasing can occur
- upsampling can minimize aliasing
- the lower the BW of $\Gamma(\omega)$, the less aliasing will happen

How fast does the detector have to react?

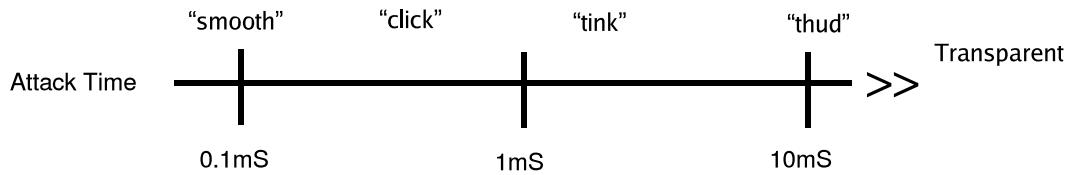


Figure 31: *Perceptual qualities of attack times.*

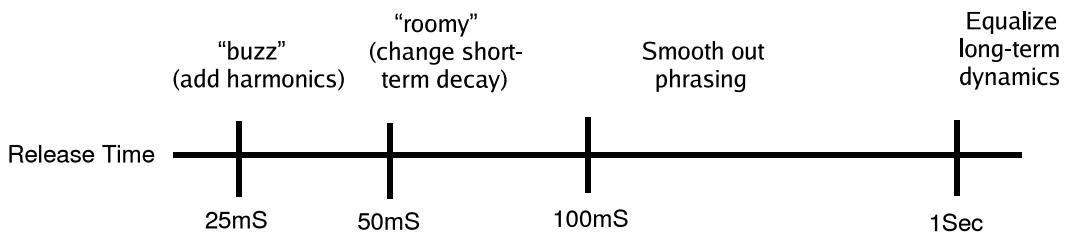


Figure 32: *Perceptual qualities of release times.*

Attack:

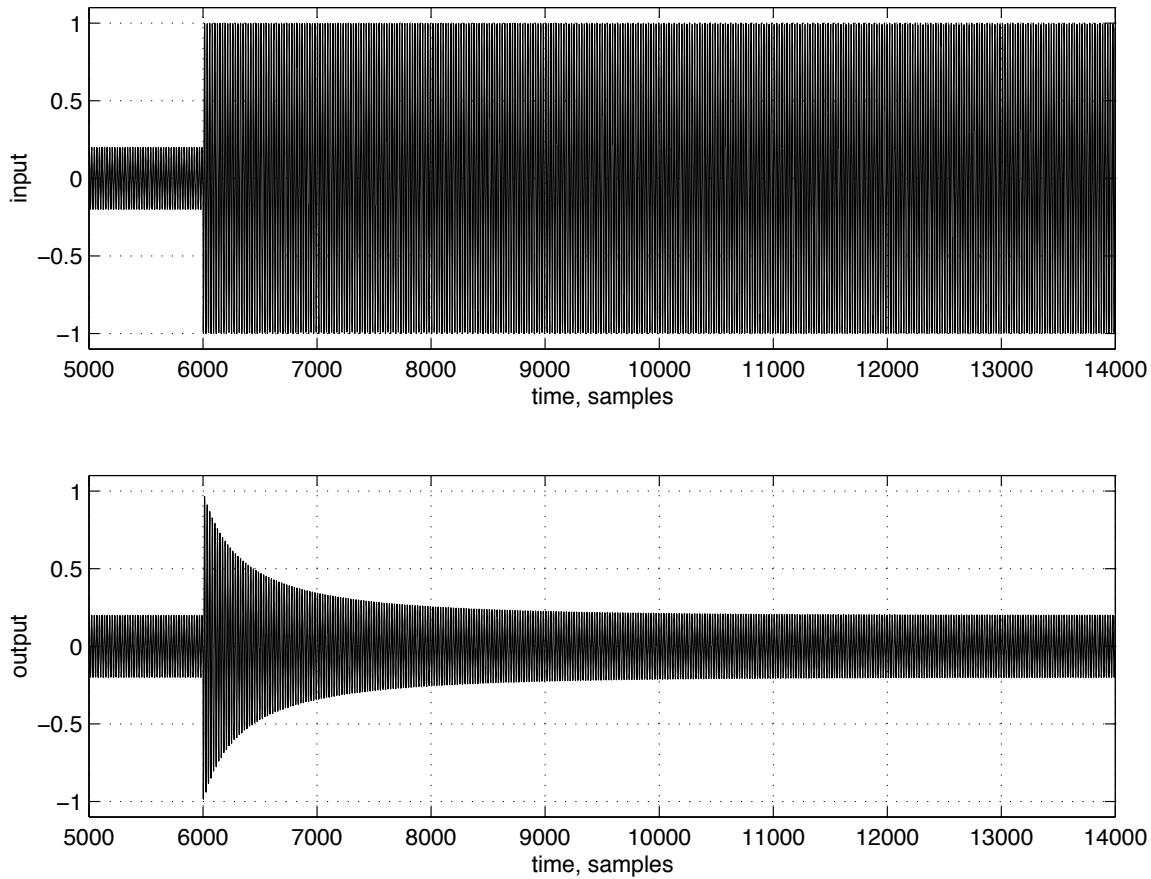


Figure 33: Envelope of compressed signal during attack.

- compression attack imposes artificial 'envelope' on input signal
 - BW of transient $\propto 1/\tau_A$

Release:

- $\tau_R \gg \tau_A \rightarrow \tau_r$ controls how fast modulation can happen due to gain changes.
- $\tau_R < 50mS$: $1/50mS = 20Hz$
 - modulation at audio rates!
 - timbral changes, hear 'distortion'
- $\tau_R > 50mS$:
 - think 'time domain', we are "out of" audio band.
 - decay of note envelopes
 - phrasing
 - long-term dynamics
- roughly true for any process, that:
 - $\tau < 50mS$ – think spectral
 - $\tau > 50mS$ – think time

Example: chorus/flanger

- same algorithm can be used for both
 - $\tau < 50mS$: flanger, hear filtering
 - $\tau > 50mS$: chorus/doubler, hear echoes/pitch modulation

Analog Detector Circuits

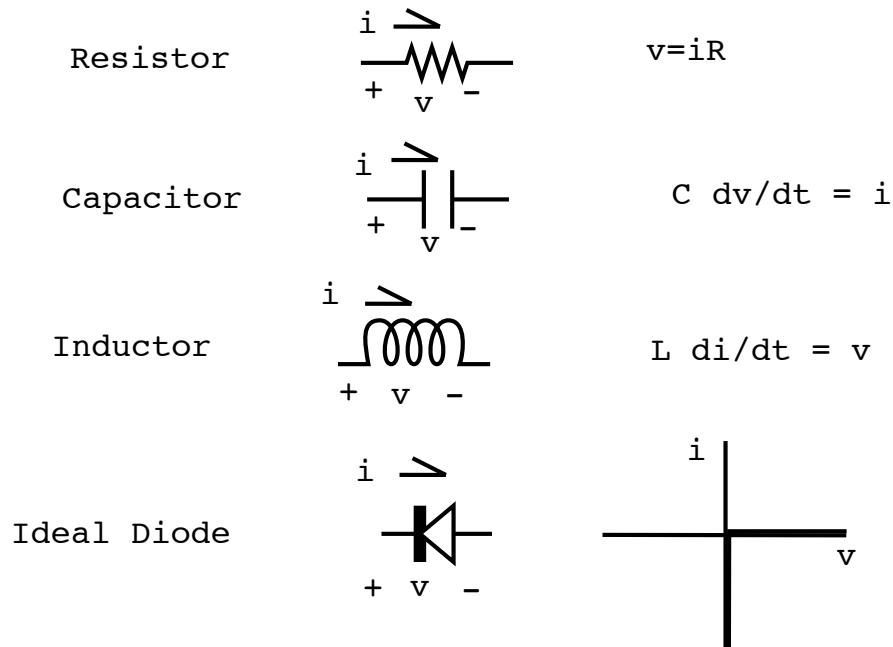


Figure 34: Lumped element definitions.

- R, L, C are *linear* elements.

Full-wave rectifier (Diode bridge)

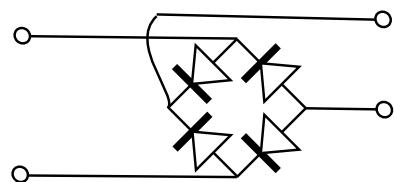


Figure 35: Diode Bridge. Inputs on left, outputs on right

Rectification

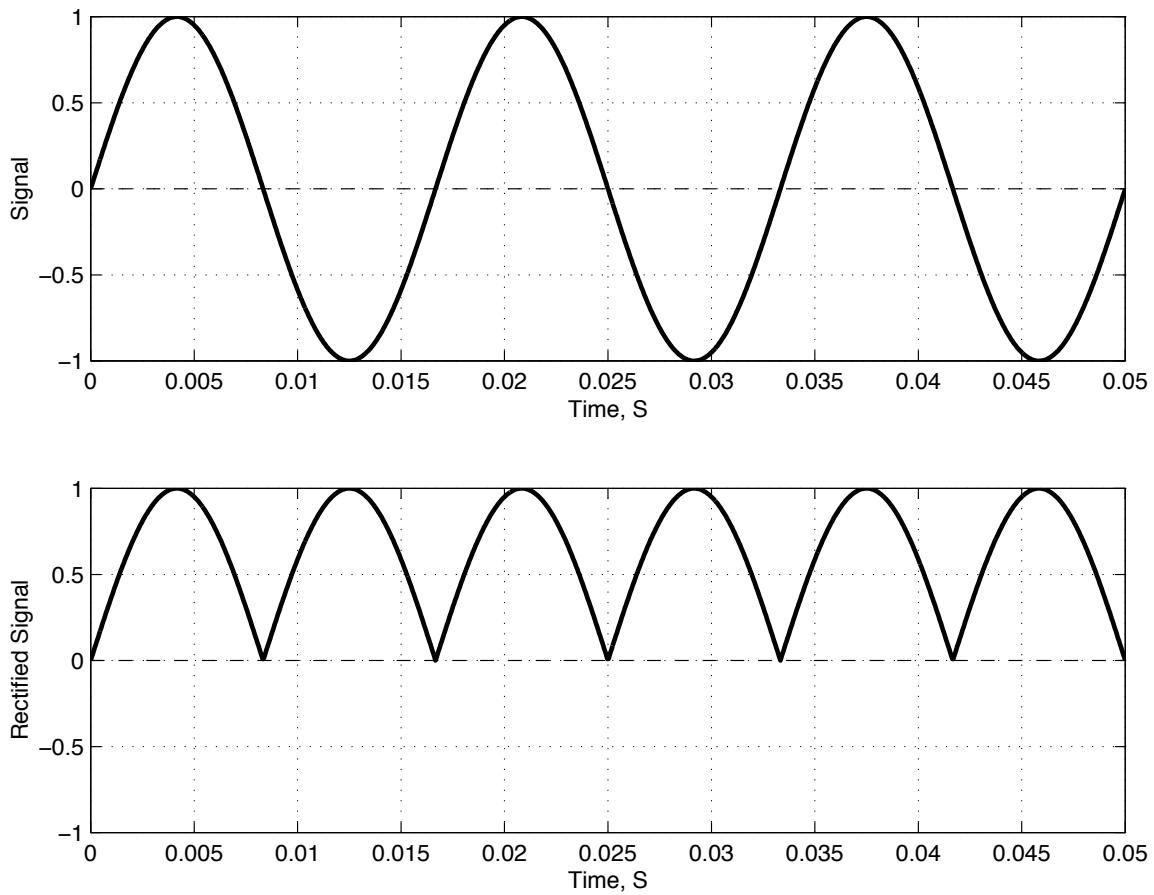


Figure 36: *Signal rectification.*

Analog Peak Detection

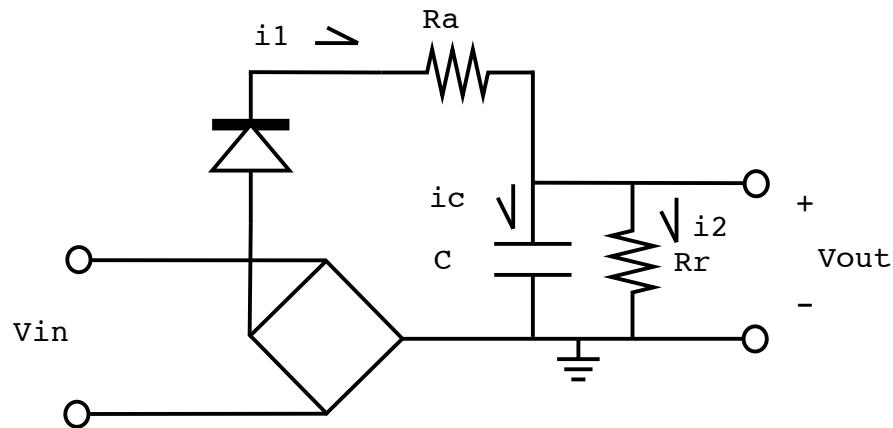


Figure 37: Peak Detection Circuit

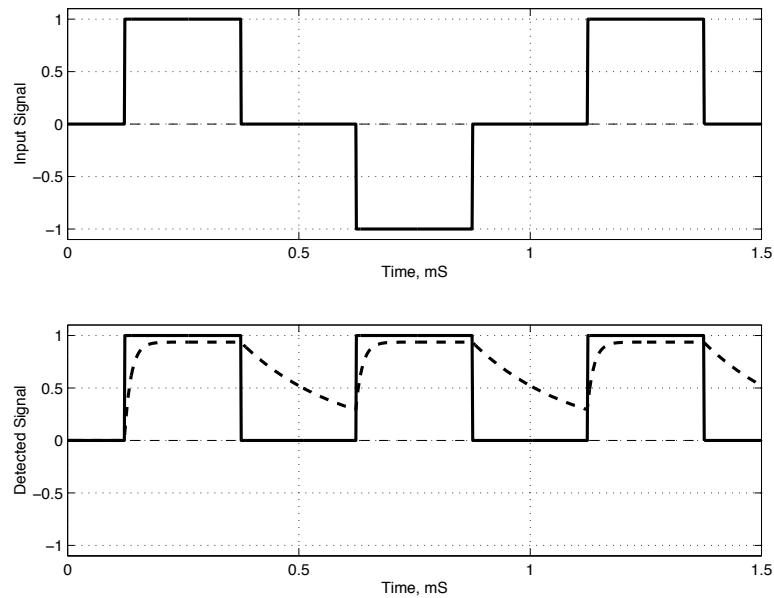


Figure 38: Peak detector input, output

Peak Detector, cont'd

- detector tracks to value of signal exponentially
- rate of tracking is asymmetric, different for attack and release
- detector output does not converge exactly to signal level

State equations for peak detector, with diode conducting:

$$\begin{aligned} v_{\text{out}} &= v_C \\ i_1 &= \frac{(v_{\text{in}} - v_{\text{out}})}{R_A} \\ i_2 &= \frac{v_C}{R_r} \\ i_C &= i_1 - i_2 \\ C \cdot \frac{dv_C}{dt} &= i_C \end{aligned}$$

Step response for filter: for $v_{\text{in}} = u(t)$ and $v_C|_{t=0} = 0$,

$$v_{\text{out}} = \frac{R_R}{R_R + R_A} \left[1 - e^{-t/(R_R \| R_A)C} \right]$$

- can be implemented as a filter

State equations for peak detector, with diode 'off'

$$v_{\text{out}} = v_C$$

$$i_1 = 0$$

$$i_2 = \frac{v_C}{R_r}$$

$$i_C = i_2$$

$$C \cdot \frac{dv_C}{dt} = i_C$$

Response for filter, diode 'off': for $v_C|_{t=0} = A$,

$$v_{\text{out}} = Ae^{-t/(R_R C)}$$

- can also be implemented as a filter

Two time-constants describe the behavior:

- release: $\frac{dx}{dt} = -\frac{1}{\tau_R}x$, where $\tau_R = R_R C$
- attack: $\frac{dx}{dt} = (\xi - x) \cdot \frac{1}{\tau_A}$, where $\tau_A = (R_A \parallel R_R) C$

We need to switch between two separate filters for attack and release. **This becomes a nonlinear system.**

- need to determine when to use which filter
 - use attack filter when $v_{\text{in}} > v_{\text{out}}$ (diode 'on')
 - use release filter when $v_{\text{in}} < v_{\text{out}}$ (diode 'off')

Discrete-time implementation of peak detector

Problem: How to preserve filter states when switching between filters

Solution: Each filter has one pole, no zeros:

- each filter has exactly one state
- $v_C \stackrel{\text{def}}{=} v_{\text{out}}$ is continuous in time.

v_C remains constant during transitions between attack and release.
 → if the filter is implemented so that its state represents v_C , no state modification is necessary when switching filter coefficients.
 For sampling rate f_s , input $x(n)$, and detector state D:

$$D = D - \left[1 - e^{-1/\tau_R f_s} \right] \cdot D$$

if($|x(n)| > D$)

$$D = D + \left[1 - e^{-1/\tau_A f_s} \right] \cdot (|x(n)| - D)$$

RMS Detection

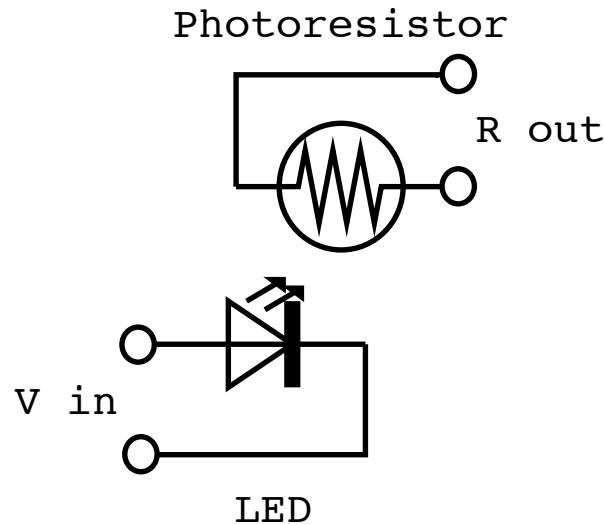


Figure 39: *RMS detector*.

This is approximate.

- want to have light produced $\propto x^2(t)$
- power absorbed by LED's $\sim |x(t)|$

But, this is referred to as RMS-type detector, mainly because it is NOT a peak detector.

colloquially: 'RMS' refers to detectors which have some of the following properties

- relatively slow reaction time
- tracks 'average' level or power rather than peak level
- no separate controls for attack and release times

Program Dependence

Problem:

slow release causes dropouts after transients; fast release → too much harmonic distortion

Goal:

- recover from transients quickly (no pumping)
- slow release down at higher average levels to reduce pumping and distortion

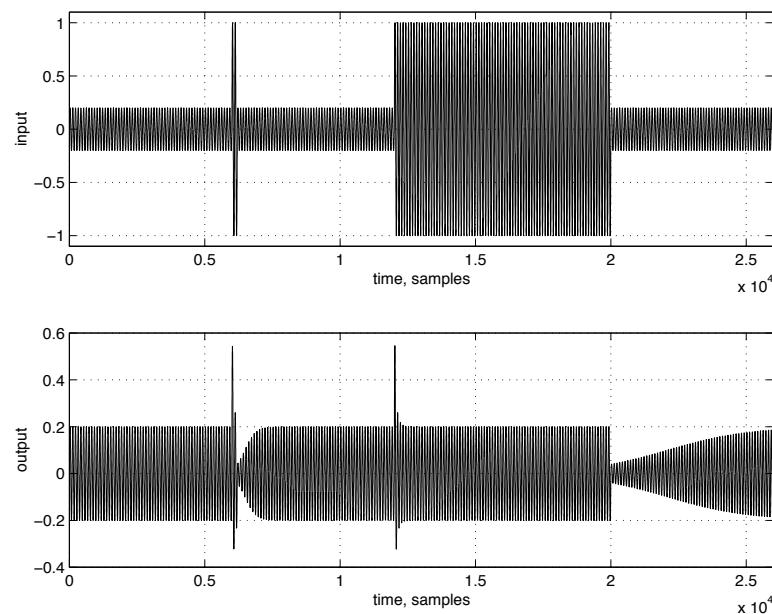


Figure 40: Program-dependent release.

Common methods for implementing program dependence:

1. Higher order filter for detector

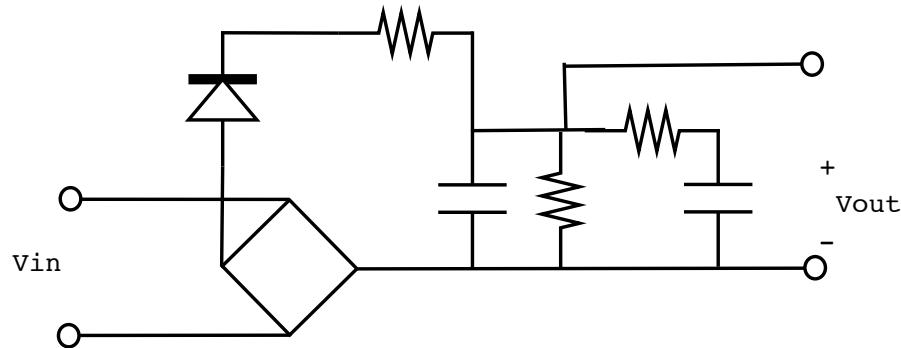


Figure 41: Program-dependent detector.

2. Run two separate detectors, one fast and one slow: **either**

- have fast detector decay towards slow detector

or

- adjust τ_R for fast detector depending on the ratio $D_{\text{fast}}/D_{\text{slow}}$

then

- use fast detector output for detected level

Affine Detectors

- for analog compressors, gain computer determined by choice of components
 - FET
 - variable-mu tubes
 - VCA
- we would like to control ratio

From last time: for compression ratio ρ , gain Φ_B

$$\Phi_B = \lambda_O^{(1-\rho)}$$

for a feedback compressor.

→ ρ seemingly determined by Φ_B .

Can we change the ratio?

- Apply affine transformation to λ_O :

$$\hat{\lambda}_O = a \cdot \lambda_O + b$$

Example:

$$\rho = 2 \rightarrow \Phi = \lambda_O^{-1}$$

try to generate warped estimate to obtain compression ratio R:

$$\begin{aligned}\hat{\lambda}_O &= (a\lambda_O + b) \\ \Phi &= (a\lambda_O + b)^{-1}\end{aligned}$$

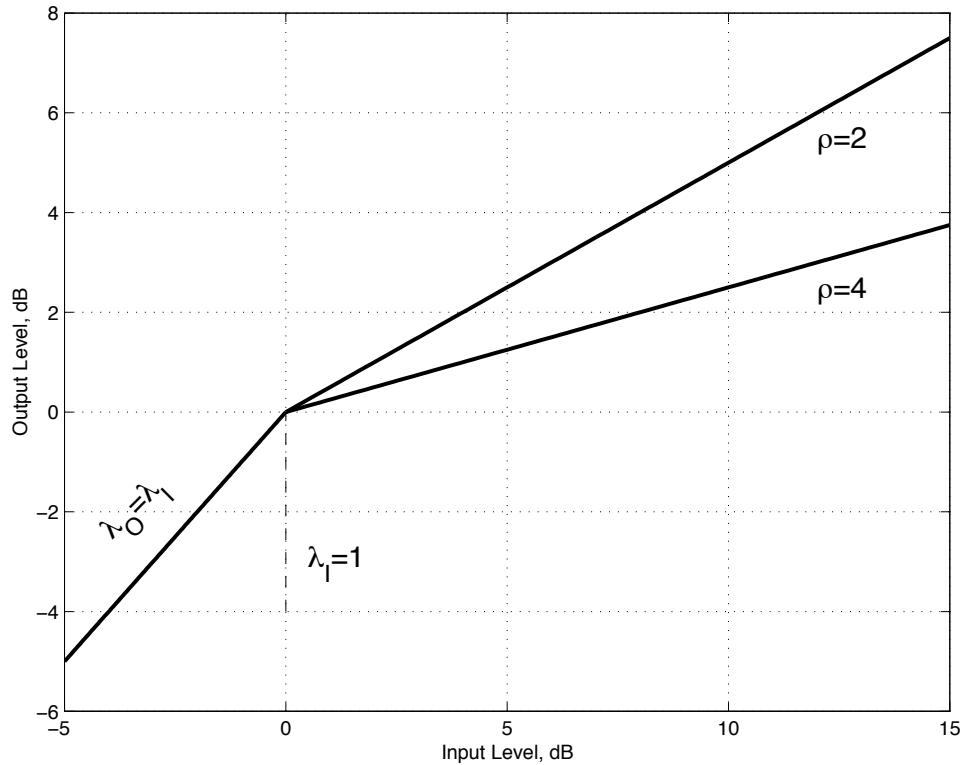


Figure 42: Multiple compression ratios.

we want $(a\lambda_O + b)|_{\lambda_O=1} = 1$ so $a + b = 1$, giving

$$\hat{\Phi}_R = (a\lambda_O + (1 - a))^{-1}$$

which results in

$$\lambda_I = \frac{\lambda_O}{\hat{\Phi}_R} = a\lambda_O^2 + (1 - a)\lambda_O$$

for true compression at ratio $\rho = R$,

$$\Phi_R = \lambda_O^{(1-R)} \rightarrow \lambda_I = \lambda_O^R$$

choose a s.t. $\hat{\Phi}_R \approx \Phi_R$.

we want to minimize over a the quantity

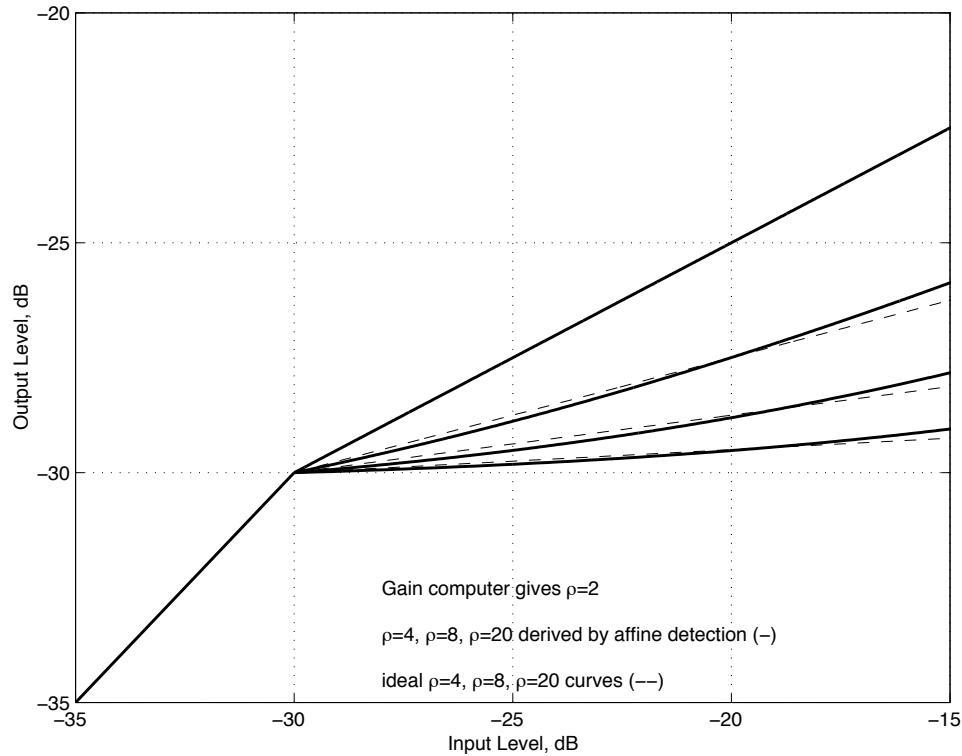
$$\int [\log(a\lambda_O^2 + (1-a)\lambda_O) - \log(\lambda_O^R)]^2 d\lambda_O$$

try

$$\int \log(a\lambda_O^2 + (1-a)\lambda_O) - \log(\lambda_O^R) d\lambda_O = 0$$

which gives

$$\begin{aligned} & (\lambda_O + \frac{1-a}{2a}) \log(a\lambda_O^2 + (1-a)\lambda_O) \\ & - 2\lambda_O + \frac{1-a}{a} \tanh^{-1} \left(\frac{2a\lambda_O + 1-a}{1-a} \right) \\ & - R(\lambda_O \log \lambda_O - \lambda_O) \Big|_{\lambda_O \text{min}}^{\lambda_O \text{max}} = 0 \end{aligned}$$

Figure 43: *Multiple compression ratios.*

Alternatively, we can pick a s.t. we get δ amount of compression when λ_I is ϵ above the threshold T ($\epsilon \cdot T$ since λ is a linear variable):

$$\hat{\lambda}_{IR} = a\lambda_O^2 + (1-a)\lambda_O$$

if the original ratio was $\rho = 2$,

$$\epsilon = a\epsilon^{2/R} + (1-a)\epsilon^{1/R}.$$

Or,

$$a = \frac{\epsilon - \epsilon^{1/R}}{\epsilon^{2/R} - \epsilon^{1/R}}.$$

4. Applications, Architectures and Improvements

Compression Applications

Stereo Compression: Apply compression to stereo signal with two channels, $x_L(t)$ and $x_R(t)$.

- preserve imaging: to prevent image shift, gain applied to the left and right channels should be the same at any instant in time.

$$g_L(t) = g_R(t) \quad \forall t$$

- must make single level estimate from multiple inputs
 - peak: find max peak over all channels

$$\lambda = \max(|x_L(t)|, |x_R(t)|)$$

- RMS: find average power across all channels

$$\lambda = (0.5(x_L^2(t) + x_R^2(t)) * e^{-t/\tau})^{1/2}$$

- for both peak and RMS, we are doing the same thing across multiple channels that we do with respect to time.

Multi-channel compression:

- usually n-generalization of stereo compression
 - peak: $\lambda = \max_n |x_n(t)|$
 - RMS: $\lambda = (\frac{1}{n} \sum_n x_n^2(t) * e^{-t/\tau})^{1/2}$
- for 5.1 surround, often center channel contains all movie dialog
→ can compress separately.

Frequency-selective compression:

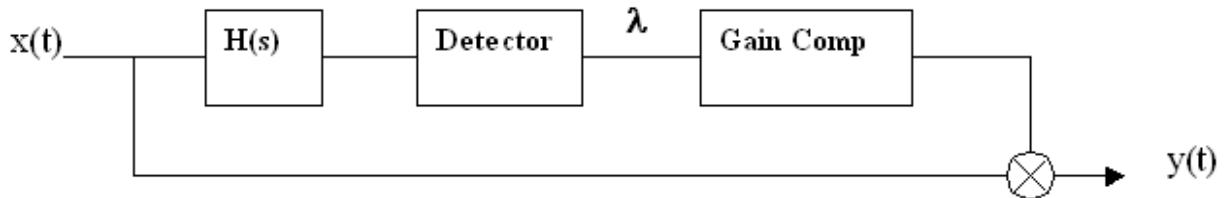


Figure 44: Block diagram for frequency-selective compression.

- make compressor sensitive to certain frequency ranges
 - de-essing (sibilance)
 - gating (rumble)
- $H(s)$ moves the threshold to different places for different frequencies

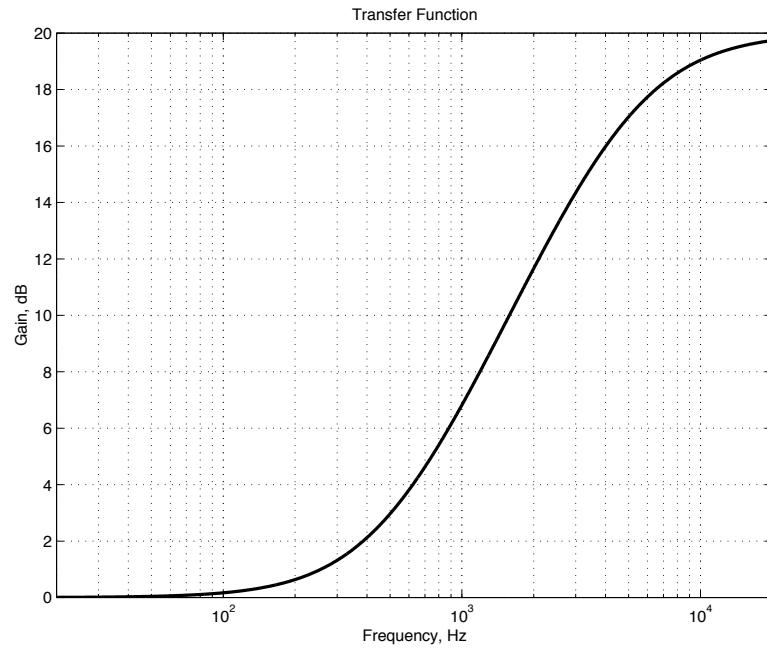


Figure 45: *Transfer function for signal detection filter.*

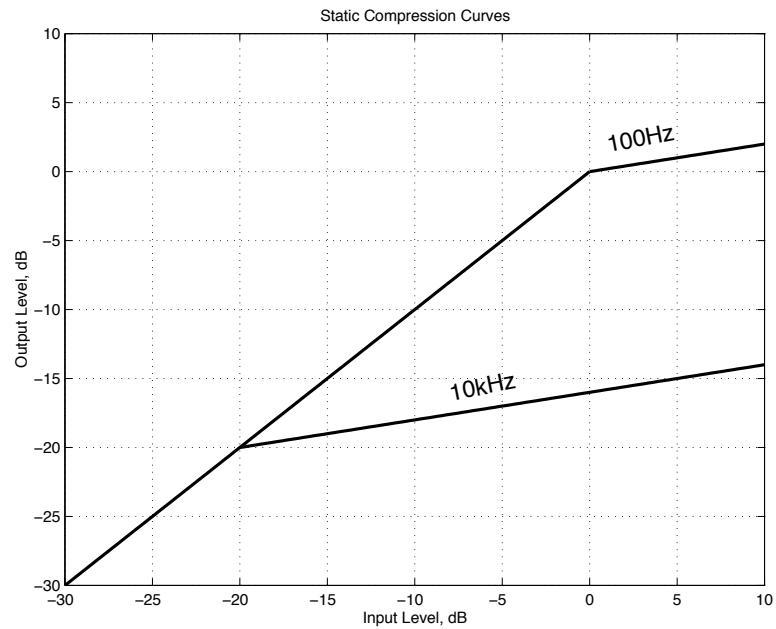


Figure 46: *Static compression curves for frequency-selective compressor.*

Frequency-selective compression, cont'd

- good for single-source tracks
 - de-ess vocal track
 - remove rumble from isolated track
- changes gain across entire audio band, not ideal for mastering or working with mixed material

Why not use EQ?

- wanted/unwanted sources may overlap in frequency
- may want to remove isolated problems but be transparent elsewhere

Sidechain input:

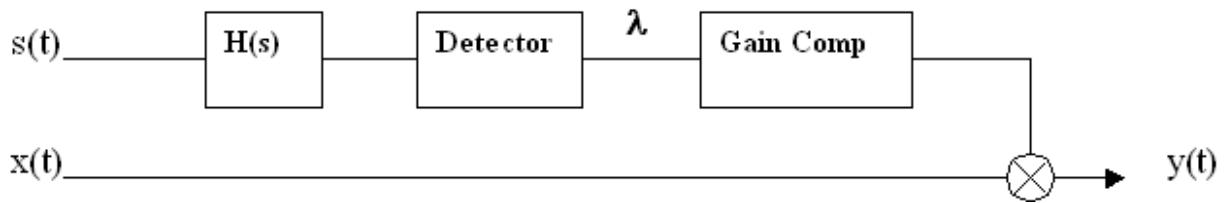


Figure 47: Block diagram for compressor with sidechain input.

- common for gating
 - example: gate noisy good-sounding signal with other more clean signal

Multiband Compression:

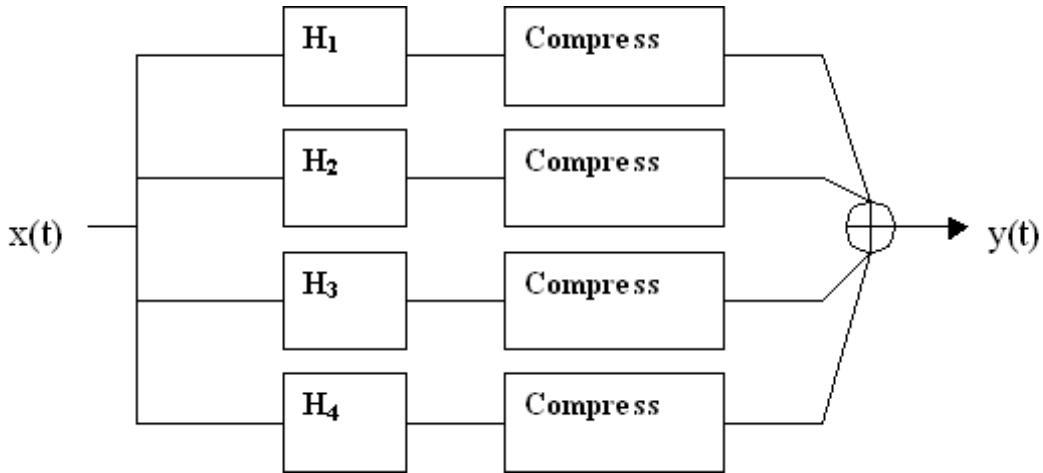


Figure 48: *Block diagram for multiband compressor.*

- filterbank: use either perfect reconstruction or perfect magnitude reconstruction filterbank.
- can be used for compressing mixes ; minimizes “pumping”
- sometimes called “dynamic EQ”
- can use different attack/release times for different bands
- not good for peak limiting, since transients are wideband

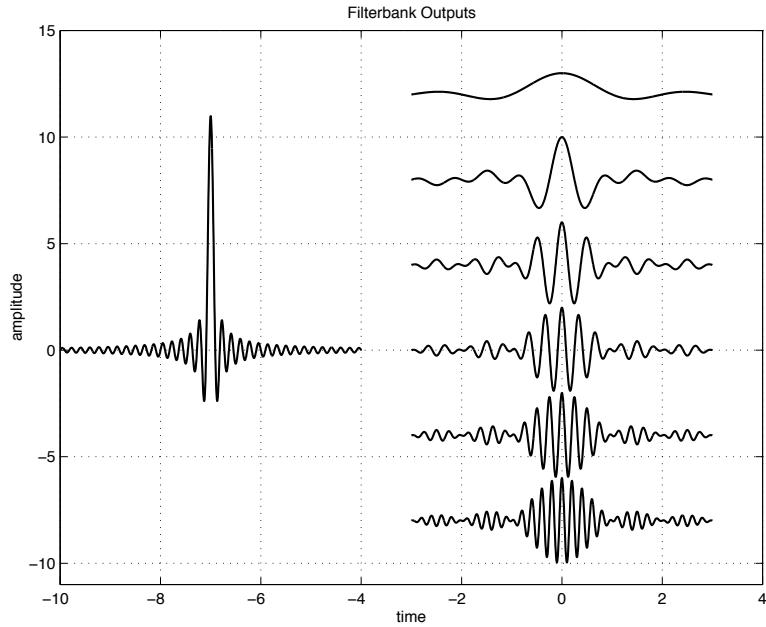


Figure 49: *Input transient and decomposition by filterbank.*

- energy from transient spread evenly among bands ; peak input level not obvious by looking at separate bands
- peak location smeared because of limited bandwidth available to each band
- each band has characteristic “time response” which is $\propto 1/\text{BW}$. Attack and release times need not be faster than this.
- can minimize harmonic distortion by choosing longer release times for low-frequency bands.

Aliasing / Upsampling:

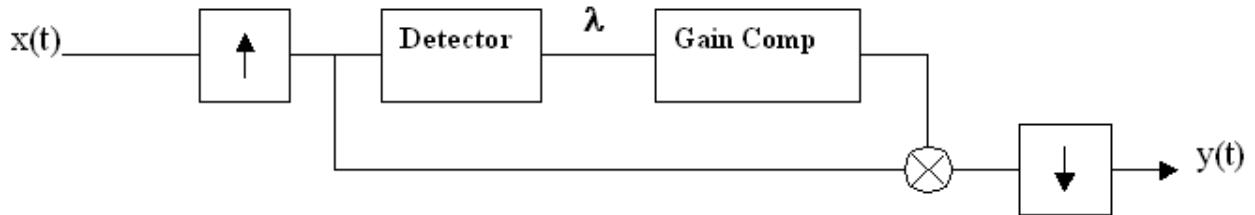


Figure 50: Block diagram for compression with upsampling.

- upsample → compress → downsample
- gives guard band for added BW from modulation

Question: What is the bandwidth of $g(t)$?

Answer: Usually $g(t)$ is lowpassed because of smoothing in the signal detector, but the bandwidth of $g(t)$ can go up to the Nyquist limit at whatever upsampled rate is used.

Alternative idea:

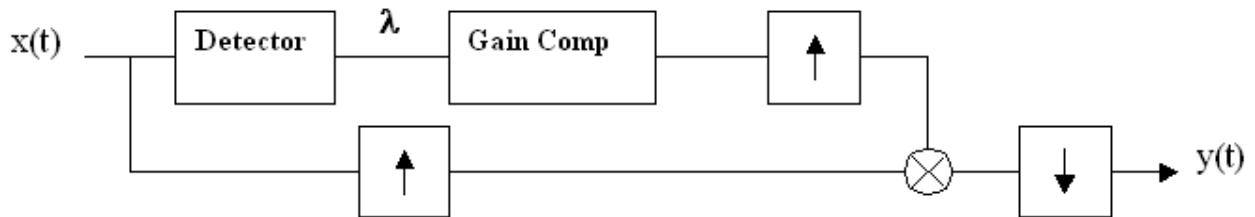


Figure 51: Alternate block diagram for compression with upsampling.

What upsampling factor is needed for this scheme?

- with 2x upsampling, we are guaranteed NO ALIASING.
 - should use same latency upsampling for sidechain and audio path to preserve alignment

- usually, gain is low BW, so a lower-cost (shorter) upsampler can be used for the detector, along with time-alignment delay.
- two upsamplers are required

Peak limiting

- want to keep peak levels below a certain value AT ALL TIMES
 - must “attack” in a finite amount of time
- common strategy: lookahead along with FIR smoothing for finite attack behavior
 - lookahead – delay audio relative to signal detection path

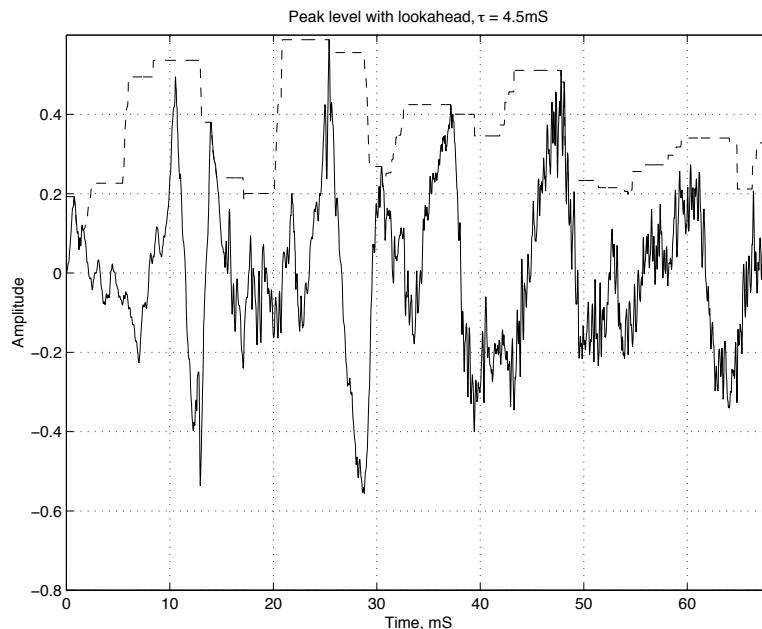


Figure 52: Peak detection with lookahead.

$$\lambda_p(t_0) = \max_{t_0 < t < t_0 + \tau} (|x(t)|)$$

- usually holdoff is set equal to the lookahead time

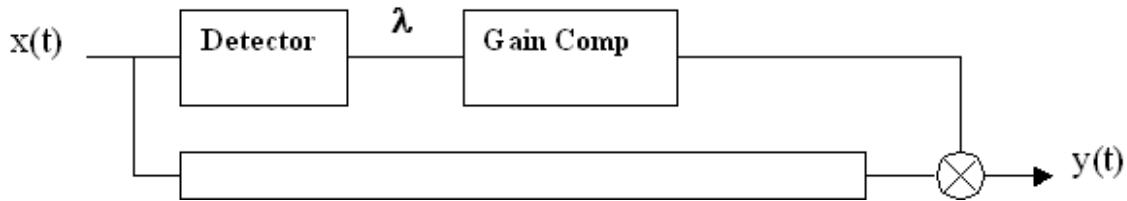


Figure 53: *Block diagram for compressor with lookahead.*

- Detection smoothing: cannot use IIR (exponential) filtering for attack – must reach peak (not asymptotically).
- can use exponential release
- need smoothing to bandlimit detector output

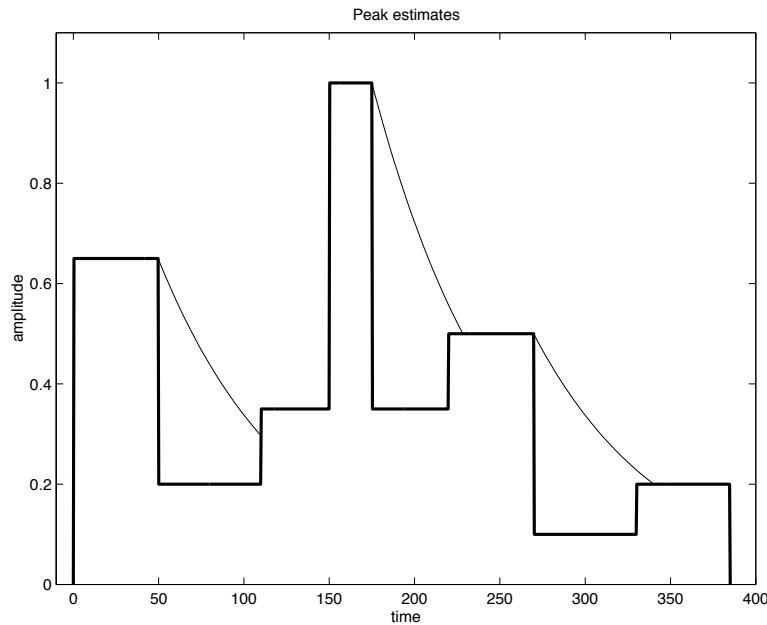


Figure 54: *Smoothed peak detection with lookahead.*

- FIR smoothing: $\text{BW} \propto 1/\tau$

- $\sum h(n) = 1$ for signal tracking
- can use “window” method to design smoothing filter (normalize window)

Spline Filter

- step response is a “spline”

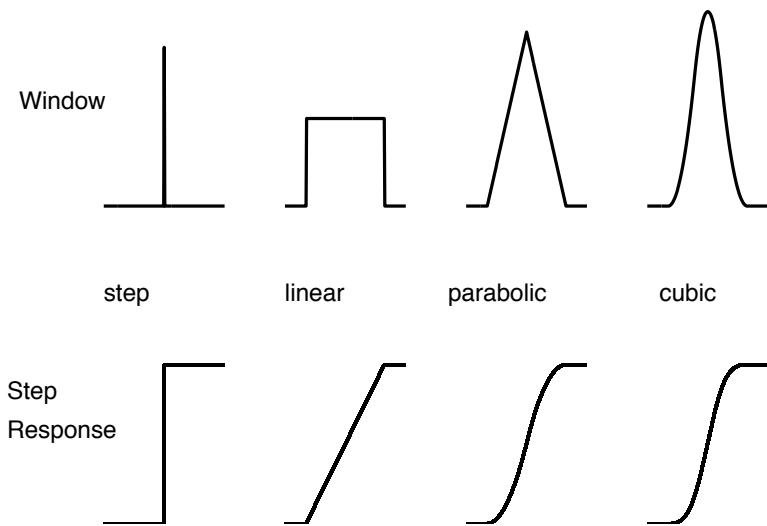


Figure 55: *Spline filters and associated step responses.*

- n rectangular windows cascaded gives n -spline window.
- window has $n - 1$ continuous derivatives
- window transform will roll off at $n \cdot 20\text{dB/decade}$
- central limit theorem: as $n \rightarrow \infty$, we approach a gaussian window

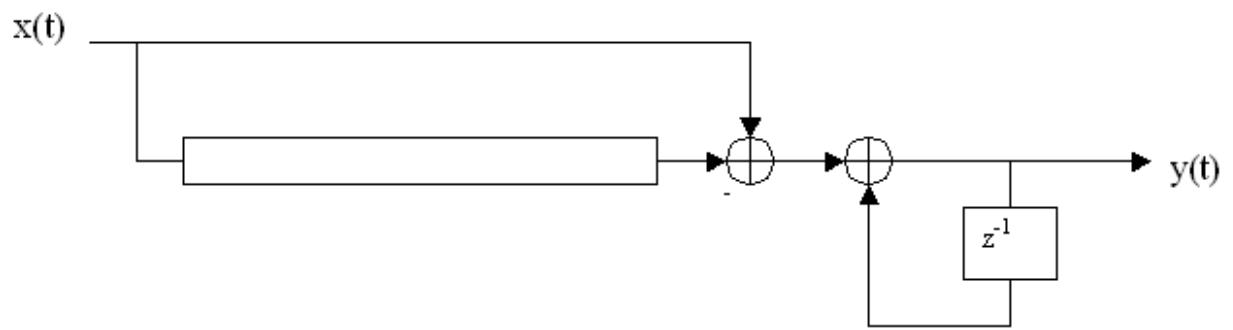


Figure 56: *Low-cost implementation for rectangular window filter.*

How does smoothing window transform appear at output?

$$\begin{aligned}\gamma(t) &= x(t) * h(t) \\ \Gamma(s) &= X(s)H(s) \\ Y(s) &= \Gamma(s)X(s) \\ &= X(s)H(s) * X(s)\end{aligned}$$

- assume $X(s)$ is wideband:

so the bandwidth of $Y(s) \approx X(s) * H(s)$ for window $h(t)$.

Offset spline design:

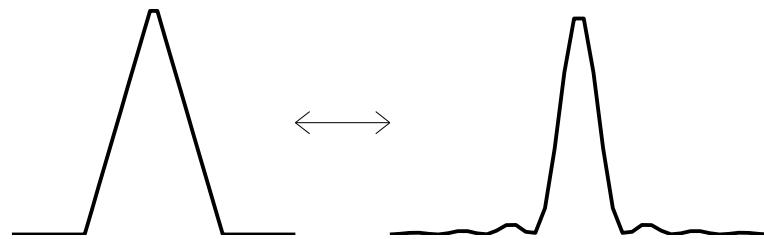
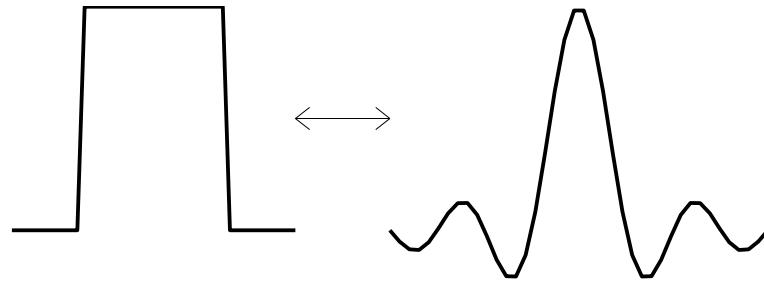


Figure 57: Transforms for rectangular and Bartlett windows.

- for Bartlett (triangular) window, $H(s) = \text{sinc}^2(s)$.
 - all zeros of sinc^2 are double-zeros

- cascade two rectangular windows of different lengths:

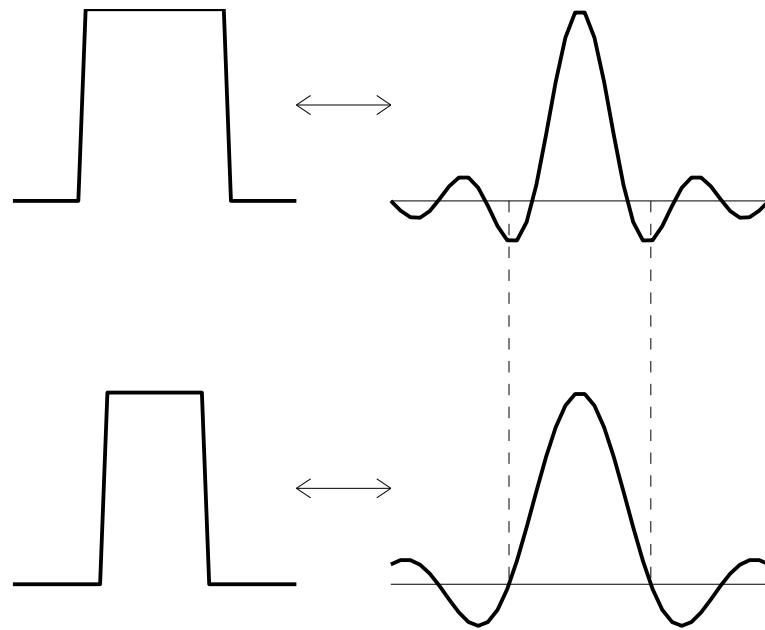


Figure 58: *Transforms for rectangular windows.*

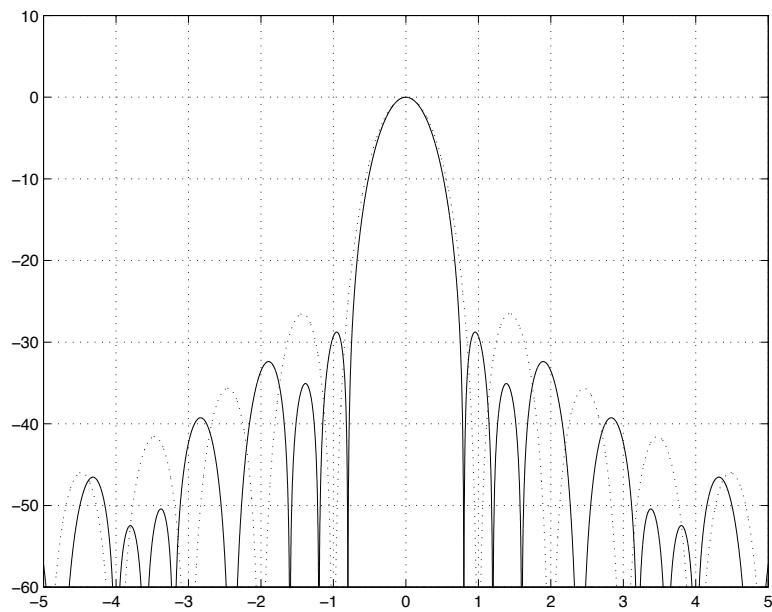


Figure 59: *Transforms for Offset Spline (solid) and Bartlett (dashed) windows.*

- two window transforms have “interleaved” zeros
- time-domain window function is trapezoidal
- can choose ratio of lengths to minimize height of first sidelobe
- similar in philosophy to Hamming window
- retain sidelobe rolloff envelope because we have same number of continuous derivatives as Bartlett window
- cascading three stretched rectangular windows can give transform similar to a Blackman window
- offset spline windows not suitable for OLA applications because windows do not sum to constant (for aliasing cancellation) with large hop sizes

Lookahead Summary:

- use lookahead, holdoff to anticipate peaks
- delay signal path by holdoff time
- use infinitely fast attack followed by FIR smoothing
- spline filters good for limiting bandwidth
- for 1.5mS lookahead \rightarrow 3kHz BW for $g(t)$ so may not need to upsample

Impulse Response Measurement

5. LTI Systems, Statistics Review; Impulse Response Measurement

Reverberation and Linear Time-Invariant Systems

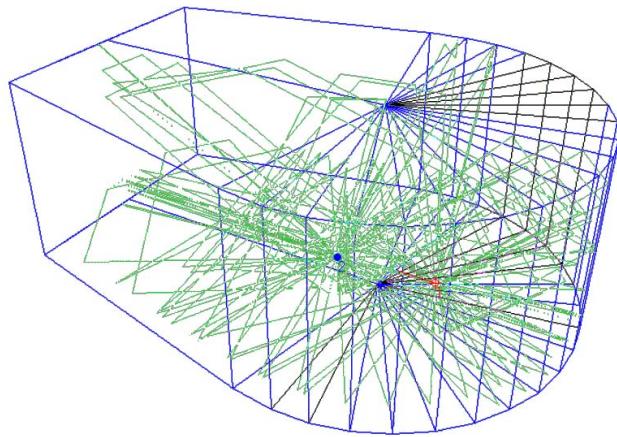


Figure 60: *Reflective Environment*

Reverberation is the arrival of energy reflected by the environment; it creates the sense that sounds are prolonged by the environment.

- Everyday objects and construction materials are very reflective of acoustic energy at audio frequencies.
- Reverberation is experienced in virtually every enclosed space, with source signals arriving along thousands of distinct paths.
- Arriving reflected energy carries with it information about the geometry and composition of the space, lending the space its “feel.”

Reverberation and Linear Time-Invariant Systems

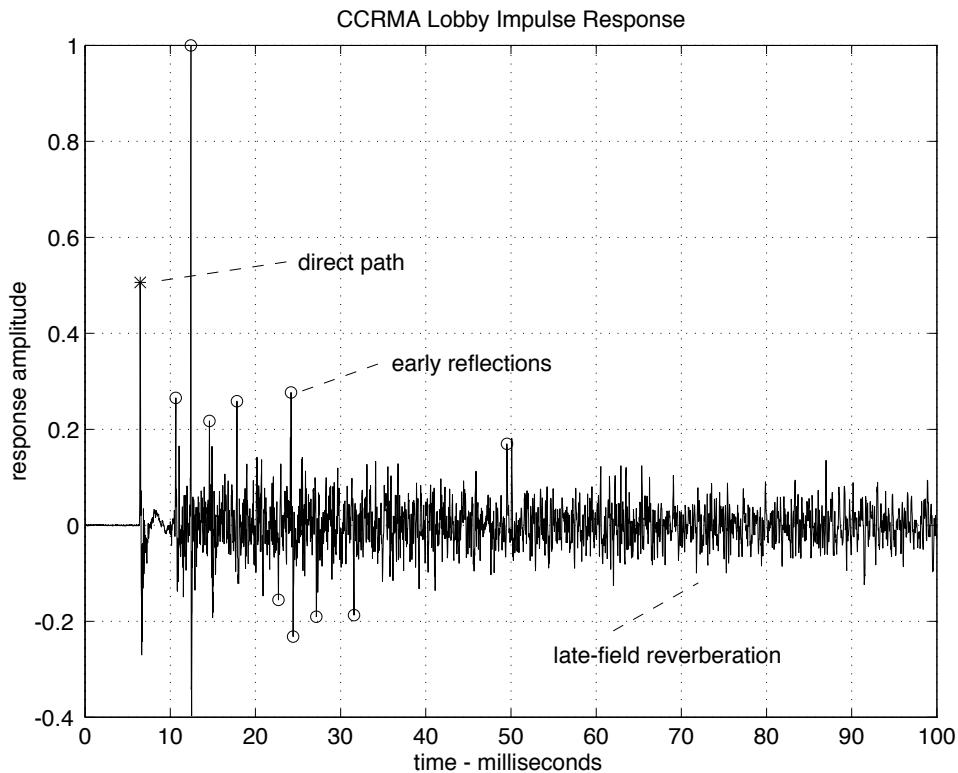
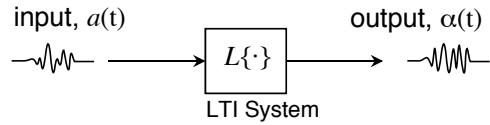


Figure 61: CCRMA lobby response to transient signal

- The arrival times and nature of reflected source signals are sensitive to the details of the environment geometry and materials.
- As a result, reverberation may seem unmanageably complex.
- Fortunately, reverberation has two properties which allow its analysis and synthesis without having to know the details:
 - *linearity*, and
 - *time invariance*.
- Here, we explore reverberation by studying impulse responses of enclosed spaces.

LTI Systems

Figure 62: *Linear Time-Invariant System.*

Definition:

Consider a system $\mathcal{L}\{\cdot\}$ operating on signals $a(t)$ and $b(t)$,

$$\begin{aligned} \alpha(t) &= \mathcal{L}\{a(t)\} \\ \beta(t) &= \mathcal{L}\{b(t)\} \end{aligned}$$

$\mathcal{L}\{\cdot\}$ is said to be *linear time-invariant* if it satisfies the following two properties.

- *Superposition or Linearity.*

$$\mathcal{L}\{a(t) + b(t)\} = \alpha(t) + \beta(t).$$

– Note: superposition implies scaling,

$$\mathcal{L}\{\gamma a(t)\} = \gamma \alpha(t).$$

- *Time Invariance.*

$$\mathcal{L}\{a(t - \tau)\} = \alpha(t - \tau).$$

- The *impulse response* of an LTI system $\mathcal{L}\{\cdot\}$ is its response to a unit pulse $\delta(t)$.

$$h(t) = \mathcal{L}\{\delta(t)\}$$

- How does the system respond to a delayed, scaled impulse, $s \cdot \delta(t - \tau)$?

$$s \cdot h(t - \tau) = \mathcal{L}\{s \cdot \delta(t - \tau)\}$$

- What about the response to an arbitrary signal $s(t)$?

- Any signal $s(t)$ can be decomposed as the sum of delayed, scaled impulses,

$$s(t) = \sum_{\tau} s(\tau) \cdot \delta(t - \tau).$$

- Using superposition, the system output $\mathcal{L}\{s(t)\}$ is

$$\mathcal{L}\left\{\sum_{\tau} s(\tau) \cdot \delta(t - \tau)\right\} = \sum_{\tau} s(\tau) \cdot \mathcal{L}\{\delta(t - \tau)\}.$$

- Substituting $h(t - \tau) = \mathcal{L}\{\delta(t - \tau)\}$, the system output, called the *convolution* of $s(t)$ and $h(t)$, is

$$\mathcal{L}\{s(t)\} = s(t) * h(t) \stackrel{\text{def}}{=} \sum_{\tau} s(\tau) \cdot h(t - \tau).$$

→ Note that any LTI system is completely characterized by its impulse response, $h(t)$.

Convolution and Correlation

Definitions:

- Convolution

$$h(t) * s(t) \stackrel{\text{def}}{=} \sum_{\tau} h(t - \tau) \cdot s(\tau).$$

- Correlation

$$h(t) \star s(t) \stackrel{\text{def}}{=} \sum_{\tau} h(\tau - t) \cdot s(\tau).$$

- Relationship: $h(t) \star s(t) = h(-t) * s(t)$.

Properties:

- Associativity. (Filters may be applied in succession, or their impulse responses convolved and applied.)

$$\begin{aligned} s(t) * (h(t) * g(t)) &= (s(t) * h(t)) * g(t), \\ s(t) \star (h(t) \star g(t)) &= (s(t) \star h(t)) \star g(t). \end{aligned}$$

- Commutivity. (Filters cascade order irrelevant.)

$$\begin{aligned} s(t) * h(t) &= h(t) * s(t), \\ s(t) \star h(t) &= \text{time_flip}\{h(t) \star s(t)\}. \end{aligned}$$

- Distribution. (Parallel filters add.)

$$s(t) * (h(t) + q(t)) = s(t) * h(t) + s(t) * q(t)$$

Impulse Response Measurement

Measurement Model:

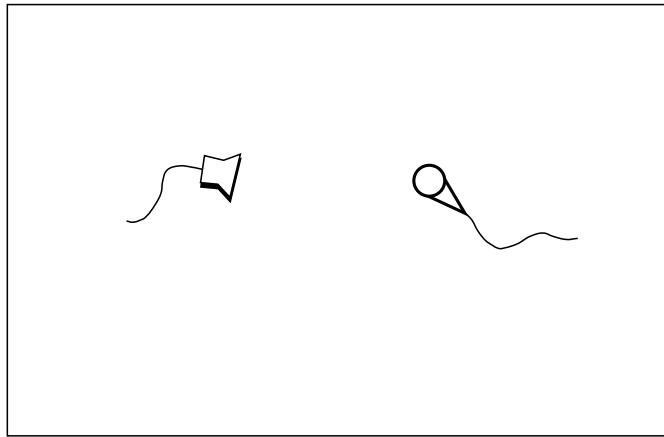


Figure 63: *Measurement Configuration*

- Room presumed LTI.
- Why not just crank an impulse out the speaker, record the result and declare victory?

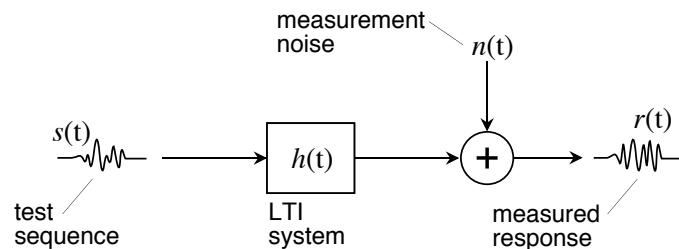


Figure 64: *Measurement Model*

- Noise assumed additive, unrelated to the test signal. (How will it be related to the system?)

Measurement Approach:

- Given the measurement model

$$r(t) = s(t) * h(t) + n(t),$$

the idea is to find a test signal which will reveal the impulse response $h(t)$.

- Put an impulse into the room,

$$s(t) = \delta(t),$$

and estimate $h(t)$ as the system response,

$$\hat{h}(t) = r(t).$$

- How good is this estimate $\hat{h}(t)$?

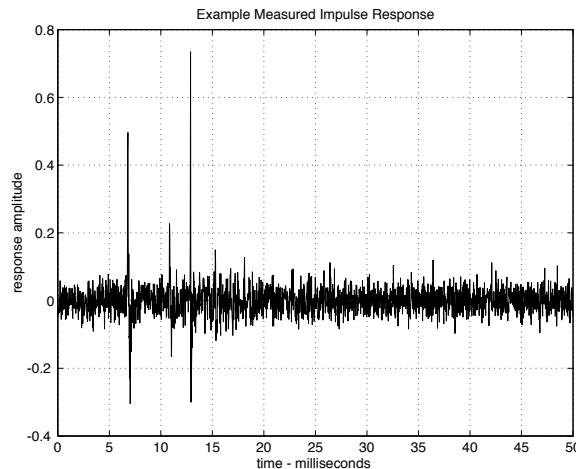


Figure 65: *Measured Impulse Response*

Probability and Statistics:

- The probability that a *random variable* x will take on a value in the interval $x \in [a, b]$ is the integral of its *probability density* $\varphi(x)$ over the interval,

$$\Pr\{x \in [a, b]\} = \int_a^b \varphi(x) dx.$$

- The *expectation* of a *statistic* $T(x)$ with respect to the random variable x is the integral over x of the value of the statistic $T(x)$ weighted by the probability of x occurring, $\varphi(x)$,

$$\mathbb{E}\{T(x)\} = \int_{x \in \Omega} T(x) \varphi(x) dx.$$

- Expectation is linear,

$$\begin{aligned} \mathbb{E}\{a(x) + b(x)\} &= \int_{x \in \Omega} [a(x) + b(x)] \varphi(x) dx \\ &= \mathbb{E}\{a(x)\} + \mathbb{E}\{b(x)\}. \end{aligned}$$

- Expectations of products of independent quantities multiply: If $x \perp y$, i.e., if

$$\varphi(x, y) = \varphi(x) \cdot \varphi(y),$$

then

$$\mathbb{E}\{a(x) \cdot b(y)\} = \mathbb{E}\{a(x)\} \cdot \mathbb{E}\{b(y)\}.$$

- The *variance* of a statistic $T(x)$ is the expectation of the square difference from its mean,

$$\begin{aligned} \text{Var}\{T(x)\} &= \mathbb{E}\{[T(x) - \mathbb{E}\{T(x)\}]^2\}, \\ &= \mathbb{E}\{T(x)^2\} - \mathbb{E}\{T(x)\}^2. \end{aligned}$$

Impulse Response Measurement Quality:

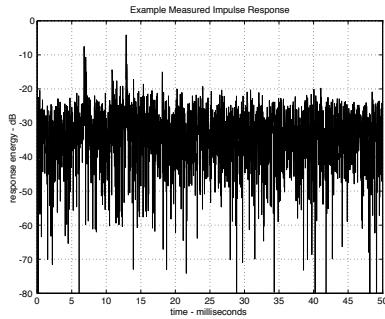


Figure 66: *Measured Impulse Response*

- Recall, $\hat{h}(t) = h(t) + n(t)$.
- Let's assume that $n(t)$ is zero mean, with variance σ^2 ,

$$\mathbb{E}\{n(t)\} = 0, \quad \text{Var}\{n(t)\} = \sigma^2.$$

- The estimated impulse response is *unbiased* (i.e., on average correct),

$$\mathbb{E}\{\hat{h}(t)\} = h(t),$$

and has variance σ^2 .

- Is this good? It depends on how loud the impulse is compared to σ^2 .
- Since impulse responses are transient by nature, the ratio of the square peak impulse response level to the noise variance σ^2 (usually expressed in dB) is often used to judge quality.

Measurement Method:

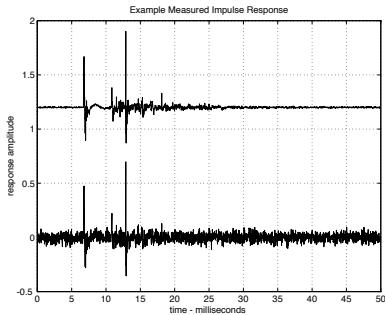


Figure 67: *Loud Impulse Test Signal*

- How about using a LOUD impulse as a test signal?
- We have

$$s(t) = \gamma \delta(t),$$

and estimate the impulse response as

$$\begin{aligned}\hat{h}(t) &= r(t)/\gamma = [\gamma \cdot \delta(t) * h(t) + n(t)]/\gamma, \\ &= h(t) + n(t)/\gamma.\end{aligned}$$

- Again the estimated impulse response is unbiased,

$$E \left\{ \hat{h}(t) \right\} = h(t),$$

but this time with variance σ^2/γ^2 .

- By putting γ^2 more energy into the room, we improved the impulse response estimate signal-to-noise ratio by the same factor of γ^2 .
- This works to a point (people do pop balloons to measure reverberation).

Measurement Method:

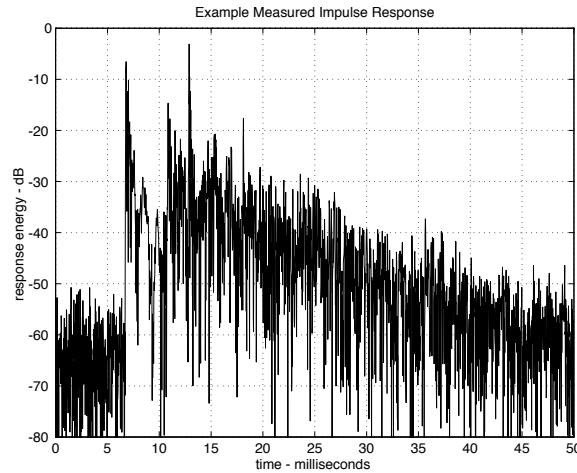


Figure 68: *Repeated Impulse Response Measurement*

- How about repeating the experiment?

- We have

$$s_1(t) = s_2(t) = \delta(t),$$

and

$$r_1(t) = \delta(t) * h(t) + n_1(t),$$

$$r_2(t) = \delta(t) * h(t) + n_2(t).$$

Note that the responses measure the same system, but with different noise realizations.

- The impulse response estimate is found by averaging the measured responses,

$$\hat{h}(t) = \frac{1}{2} [r_1(t) + r_2(t)].$$

Repeated Measurement Performance:

- The estimated impulse response is the average of the responses,

$$\begin{aligned}\hat{h}(t) &= \frac{1}{2} [r_1(t) + r_2(t)], \\ &= h(t) + \frac{1}{2} [n_1(t) + n_2(t)].\end{aligned}$$

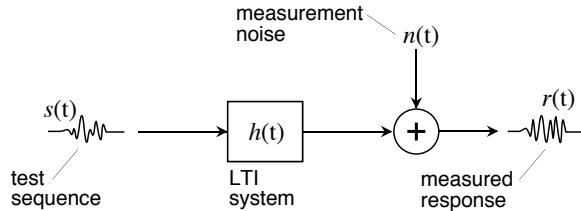
- The estimate is unbiased,

$$\mathbb{E} \left\{ \hat{h}(t) \right\} = \mathbb{E} \{ h(t) \} + \frac{1}{2} \mathbb{E} \{ n_1(t) + n_2(t) \} = h(t).$$

- The variance is halved compared to a single measurement (assuming independent measurement noise sequences, each with variance σ^2),

$$\text{Var} \left\{ \hat{h}(t) \right\} = \mathbb{E} \{ [n_1(t) + n_2(t)]^2 / 4 \} = \sigma^2 / 2.$$

- For a sequence of N independent measurements, the impulse response estimate noise variance is reduced by $1/N$ compared to that of a single measurement. We say that the *SNR gain* of the measurement process is N .
- To estimate a room impulse response with a 60 dB SNR given a 30 dB noise floor, around 1000 measurements are required. If the room has a couple second decay time, it will take about a half hour—and everybody has to sit still (why?).

Figure 69: *Measurement Model*

Measurement Approach (Golay Codes, Allpass Chirps, Sine Sweeps).

- Design a set of test signals $s_k(t)$, $k = 1, 2, \dots, K$, such that the sum of their autocorrelation sequences is an impulse of height γ^2 ,

$$\sum_{k=1}^K s_k(t) \star s_k(t) = \gamma^2 \delta(t).$$

- Measure the system separately with each test sequence $s_k(t)$, recording the responses $r_k(t)$.
- Estimate the impulse response as the sum of test sequence-measured response correlations,

$$\hat{h}(t) = \frac{1}{\gamma^2} \sum_{k=1}^K s_k(t) \star r_k(t).$$

Impulse Response Estimate Statistics.

- If the additive noise $n(t)$ is zero-mean with variance σ^2 , then

$$\mathbb{E} \left\{ \hat{h}(t) \right\} = h(t), \quad \text{Var} \left\{ \hat{h}(t) \right\} = \sigma^2 / \gamma^2.$$

6. Golay Code and Allpass Chirp Impulse Response Measurement

Impulse Response Measurement

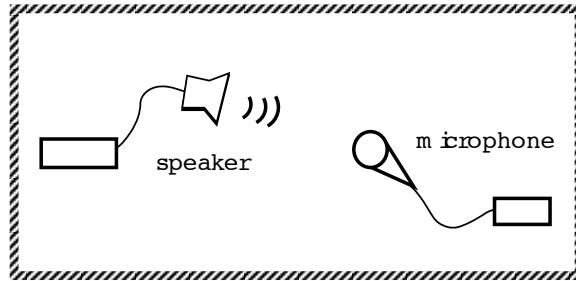


Figure 70: *Measurement Scenario*

Measurement Scenario.

- A test signal $s(t)$ is applied to the system $h(t)$, and the response $r(t)$ recorded.

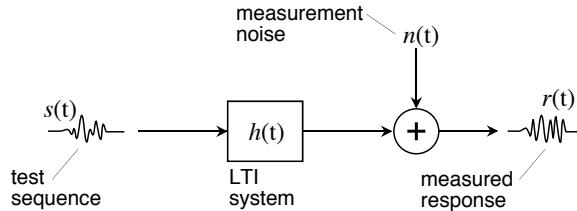
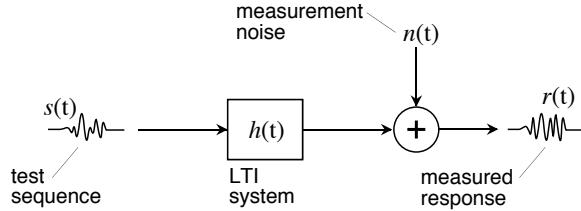


Figure 71: *Measurement Model*

Measurement Model.

- The system $h(t)$ is presumed LTI; the noise $n(t)$, additive and independent of the test sequence $s(t)$.

Figure 72: *Measurement Model*

Measurement Approach.

- Design a set of test signals $s_k(t)$, $k = 1, 2, \dots, K$, such that the sum of their autocorrelation sequences is an impulse of height γ^2 ,

$$\sum_{k=1}^K s_k(t) \star s_k(t) = \gamma^2 \delta(t).$$

- Measure the system separately with each test sequence $s_k(t)$, recording the responses $r_k(t)$.
- Estimate the impulse response as the sum of test sequence-measured response correlations,

$$\hat{h}(t) = \frac{1}{\gamma^2} \sum_{k=1}^K s_k(t) \star r_k(t).$$

Impulse Response Estimate Statistics.

- If the additive noise $n(t)$ is zero-mean with variance σ^2 , then

$$E \left\{ \hat{h}(t) \right\} = h(t), \quad \text{Var} \left\{ \hat{h}(t) \right\} = \sigma^2 / \gamma^2.$$

Golay Code Impulse Response Measurement

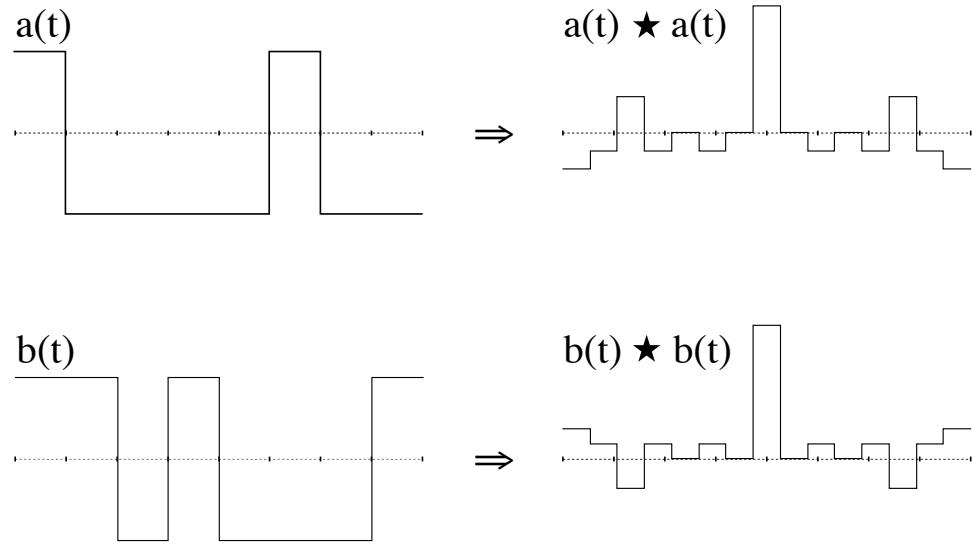


Figure 73: Example Golay Code Pair

Definition. The length L bilevel sequences $a(t)$ and $b(t)$ are said to be *complementary* or *Golay* if their autocorrelations sum to an impulse,

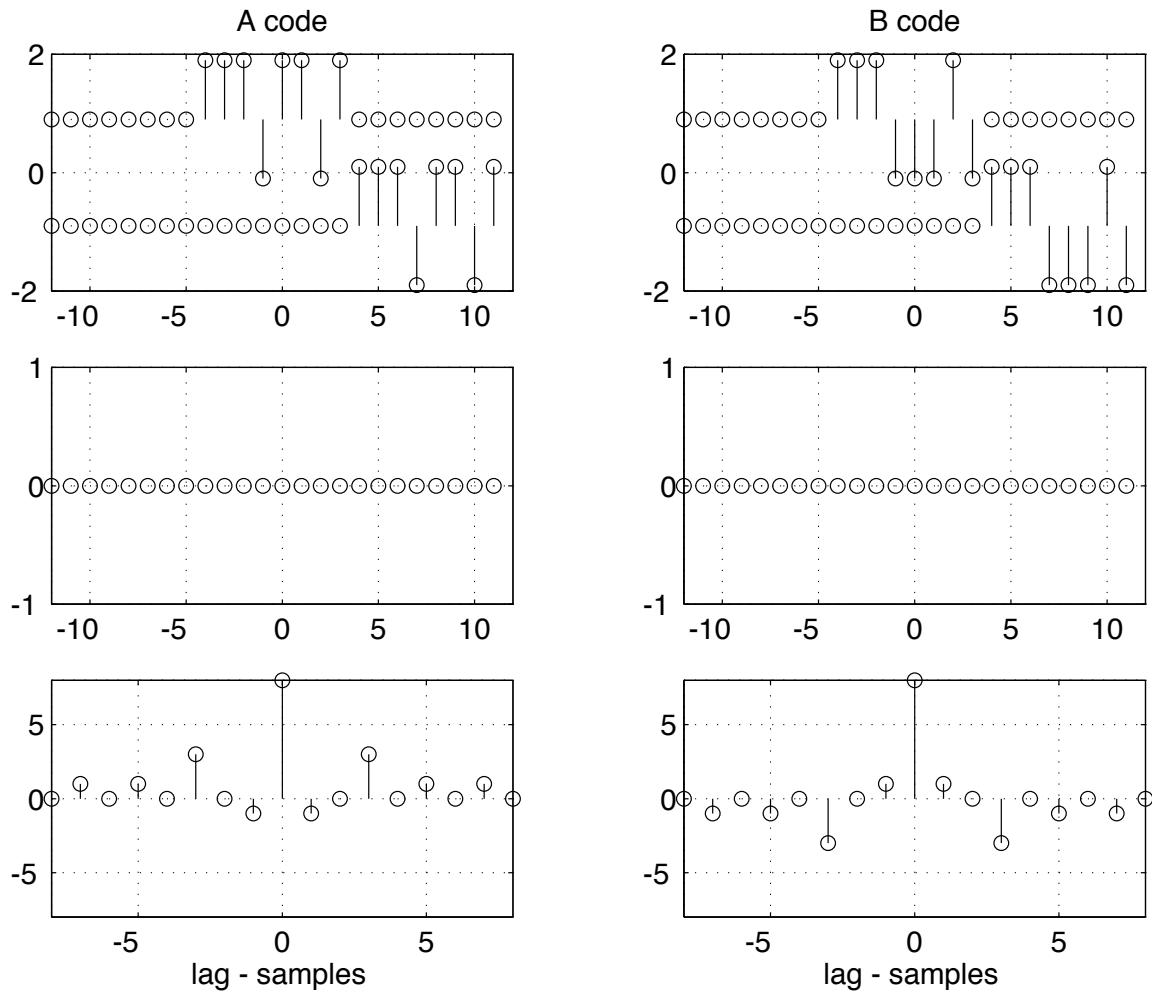
$$a(t) \star a(t) + b(t) \star b(t) = 2L\delta(t).$$

Theorem. Given length L complementary sequences $a_L(t)$ and $b_L(t)$, the sequences

$$a_{2L}(t) = [a_L(t) \ b_L(t)], \quad b_{2L}(t) = [a_L(t) \ -b_L(t)]$$

are length $2L$ complementary sequences, as are

$$a_{2L}(t) = [a_L(t) \ b_L(-t)], \quad b_{2L}(t) = [b_L(t) \ -a_L(-t)].$$

Figure 74: *Animated Golay Code Pair Correlation*

Golay Code SNR Gain.

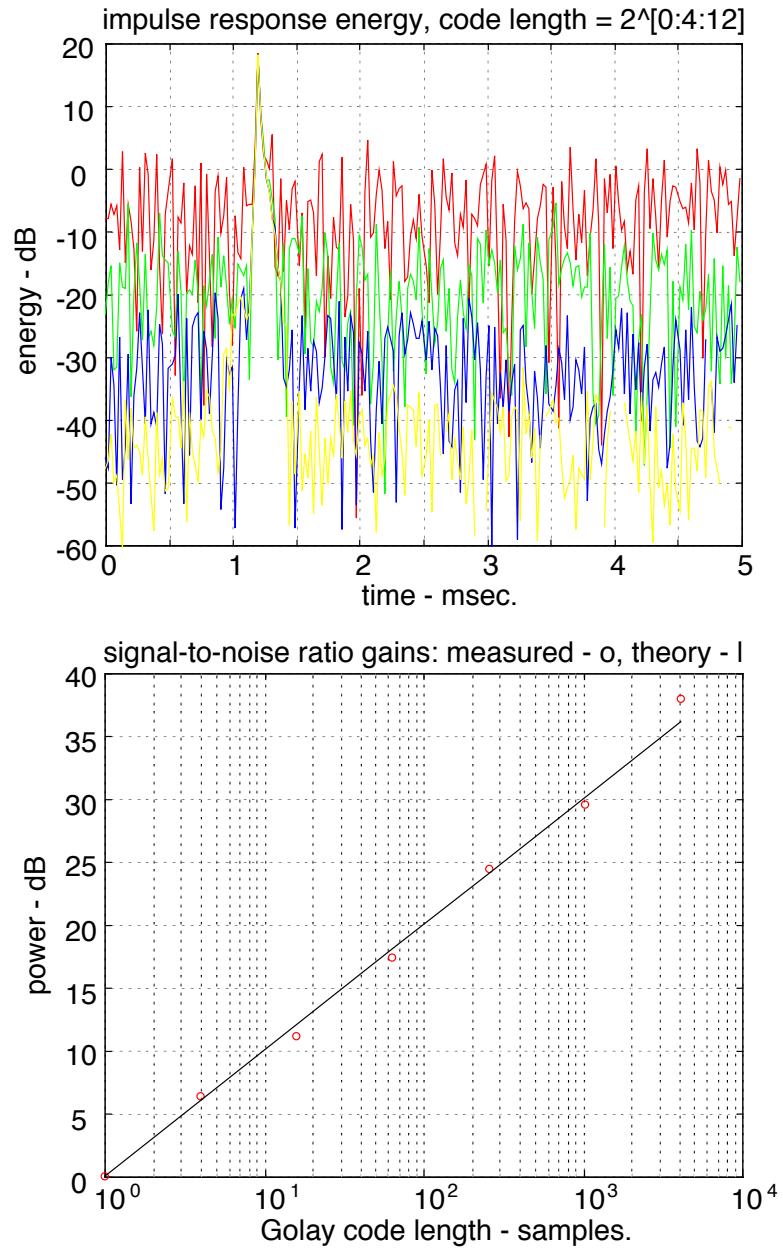


Figure 75: *Golay Code Impulse Response Measurement SNR Gain*

- Compared to using a single impulse, length L Golay codes provide an SNR gain of $10 \log_{10}\{2L\}$.

Allpass Chirp Impulse Response Measurement

First-Order Allpass Filter

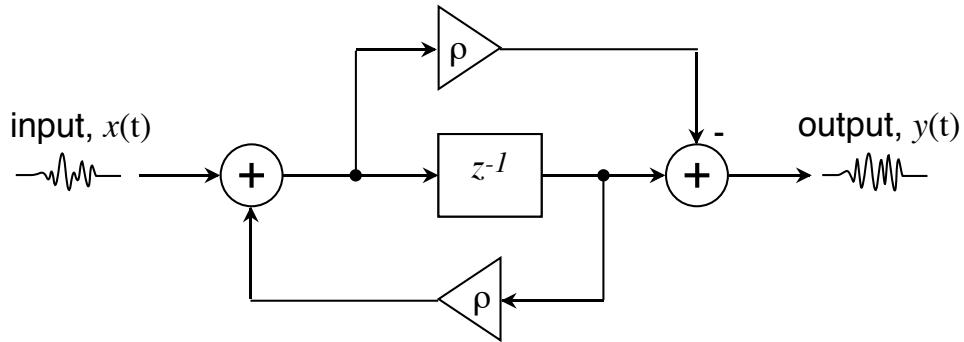


Figure 76: First-Order Allpass Filter Signal Flow

- The first-order *allpass filter* shown above is described by the difference equation

$$y(t) = -\rho x(t) + x(t-1) + \rho y(t-1),$$

and has transfer function

$$G(z) = \frac{-\rho + z^{-1}}{1 - \rho z^{-1}}.$$

Its impulse response is

$$g(t) = \begin{cases} 0, & t < 0, \\ -\rho, & t = 0, \\ (1 - \rho^2)\rho^{t-1}, & t \geq 1. \end{cases}$$

Allpass Filter Impulse Response Autocorrelation

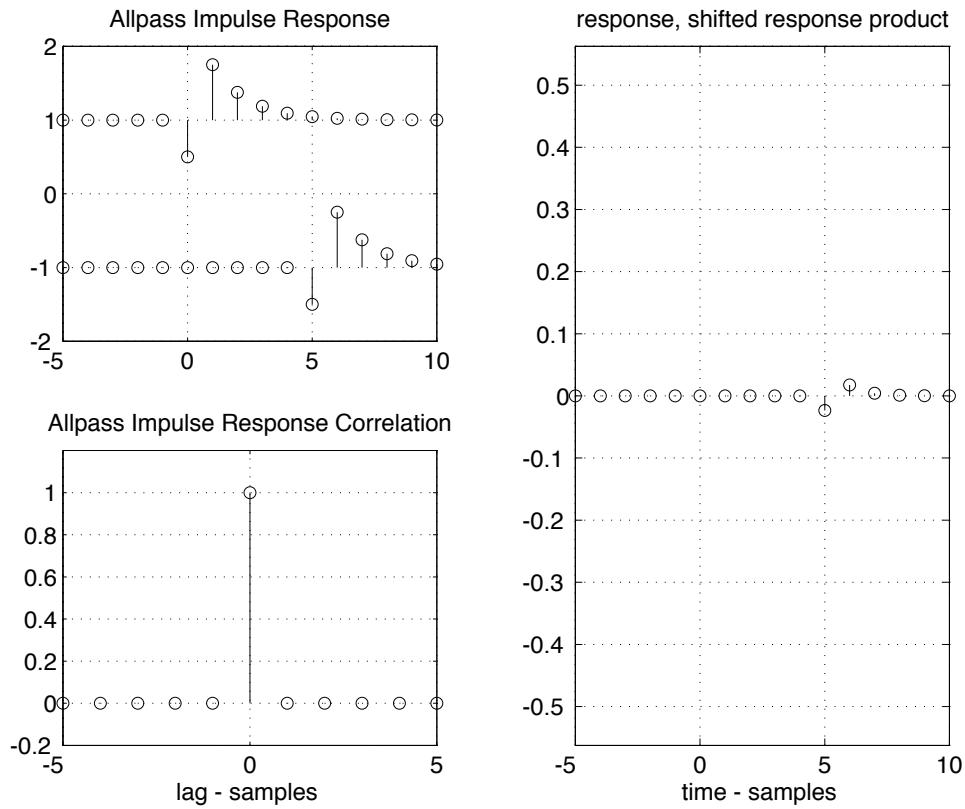


Figure 77: Animated First-Order Allpass Response Correlation

- Note that the allpass filter has an impulse response which is smeared in time, yet its autocorrelation is an impulse.

$$g(t) \star g(t) = g(-t) * g(t) \longleftrightarrow G^*(\omega)G(\omega) = 1.$$

Impulse Response Measurement Using Allpass Sequences

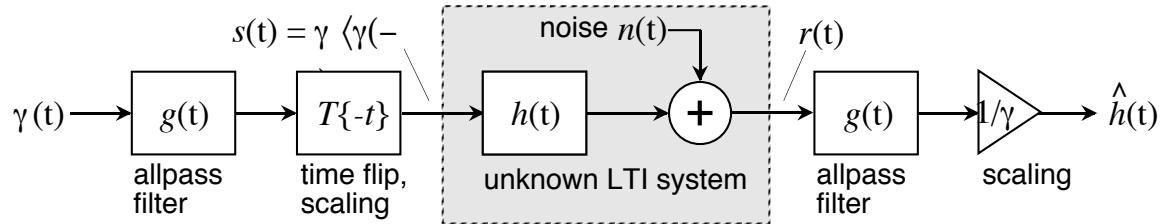


Figure 78: System Measurement Using Time-Flipped Allpass Sequences

- Allpass filter impulse responses autocorrelate to unit pulses, and therefore can be used to measure system impulse responses as follows.

- Form an allpass sequence $g(t)$ and select a scaling factor γ .
- Measure the response of the unknown LTI system to the signal $\gamma \cdot g(t)$,

$$r(t) = \gamma \cdot g(t) * h(t) + n(t).$$

- Finally, correlate the measured response $r(t)$ with the input allpass sequence $g(t)$ and normalize by γ to recover an estimate of the system impulse response,

$$\hat{h}(t) = g(t) \star r(t)/\gamma = h(t) + g(t) \star n(t)/\gamma.$$

- Note that the noise is suppressed by making γ as large as possible while maintaining linearity and other constraints such as subject comfort.

Generating Allpass Sequences

- The idea is to generate an allpass sequence with a small crest factor.
- Allpass sequences may be generated as impulse responses of allpass filters.
 - Allpass filters may be synthesized by generating time-flipped numerator and denominator coefficients,

$$G(z) = \frac{\rho_n + \rho_{n-1}z^{-1} + \cdots + z^{-n}}{1 + \rho_1z^{-1} + \cdots + \rho_Nz^{-n}}.$$

- The cascade of allpass filters is allpass,

$$|G_1(z) \cdot G_2(z)| = |G_1(z)| \cdot |G_2(z)| = 1.$$

- An allpass filter will remain allpass under the substitution of $\zeta^{-1}(z)$ for z^{-1} , if $\zeta^{-1}(z)$ takes on the form of an allpass filter. (Why?)

Measuring Room Acoustics with a Reverberator

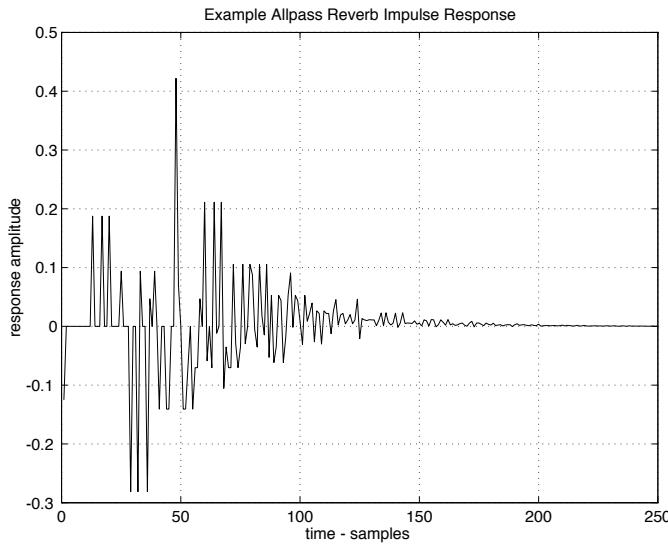


Figure 79: Example Allpass Reverberator Impulse Response

- Setting z^{-1} to z^{-n} creates an allpass filter with an impulse response consisting of a series of echos, spaced n samples apart.
- Cascading a number of these (with different delay line lengths n_i) creates an allpass system with a reverberation-like impulse response which can be used to measure room acoustics.
- A system can then be measured using the time-flipped allpass sequence as the test signal, and applying the system response to the “reverberator” that generated the test sequence.

First-Order Allpass Cascade

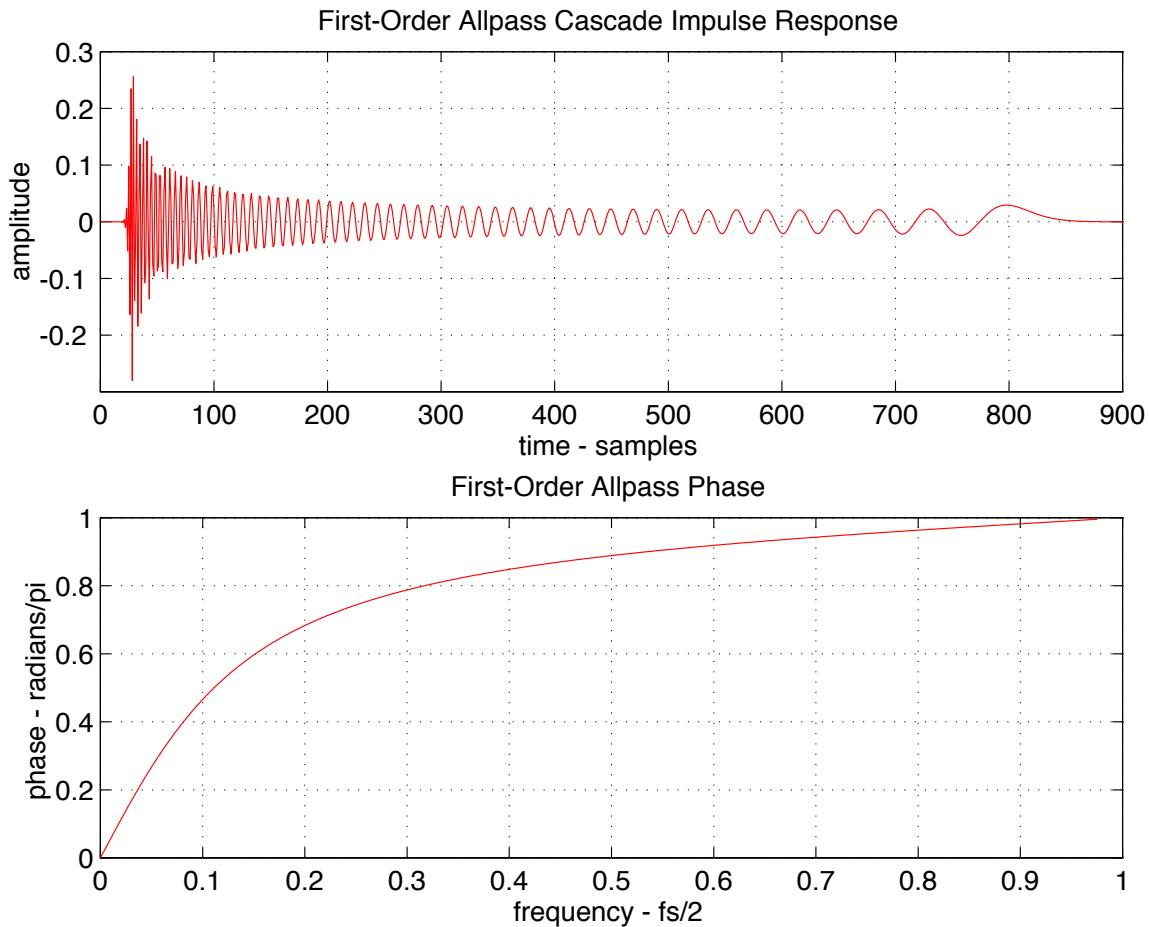


Figure 80: Example Allpass Impulse Response and Phase Response

- Cascading a large number of first-order allpass filters smears an input impulse into a sort of chirp, with low and high frequencies separated in time.
- Note the amplitude scale—the impulse response shown is unit energy.

First-Order Allpass Cascade Autocorrelation

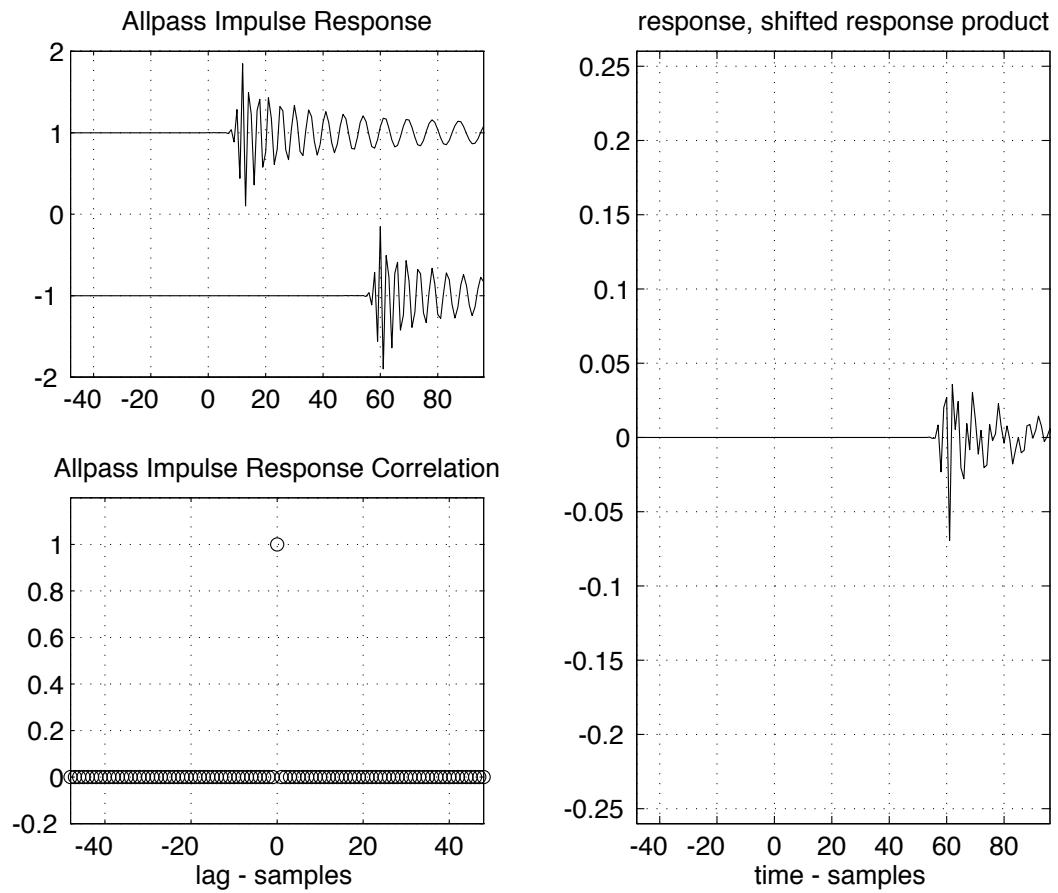
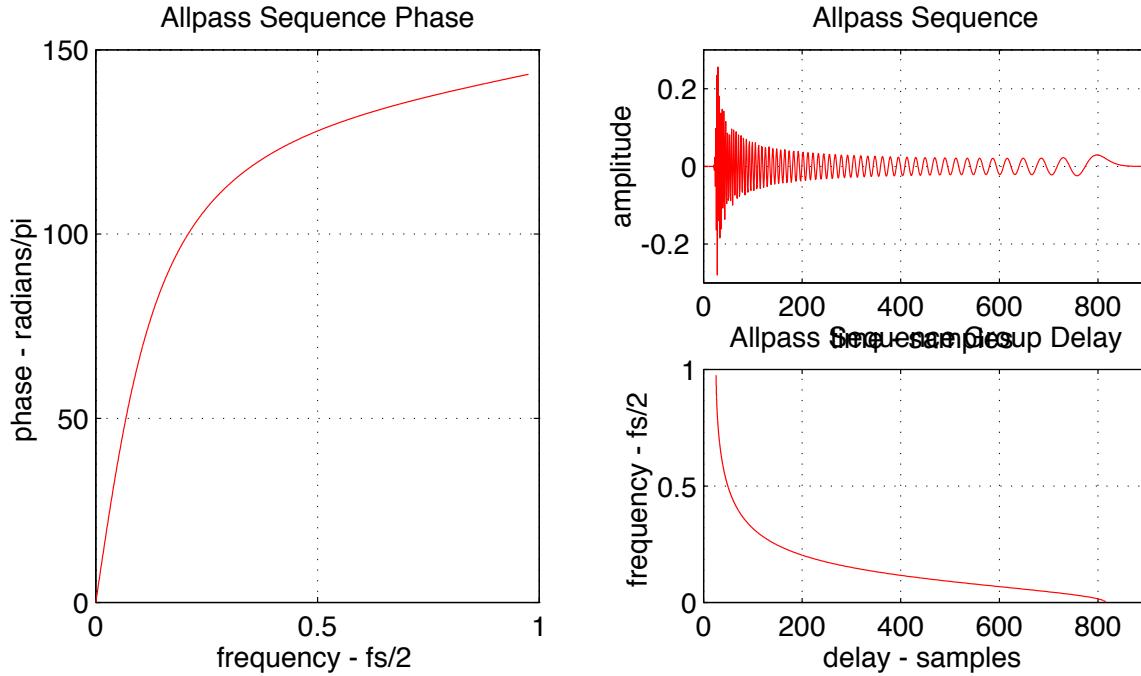


Figure 81: *Animated Allpass Chirp Response Correlation*

Allpass Sequence



- An allpass sequence has a unit-magnitude Fourier transform,

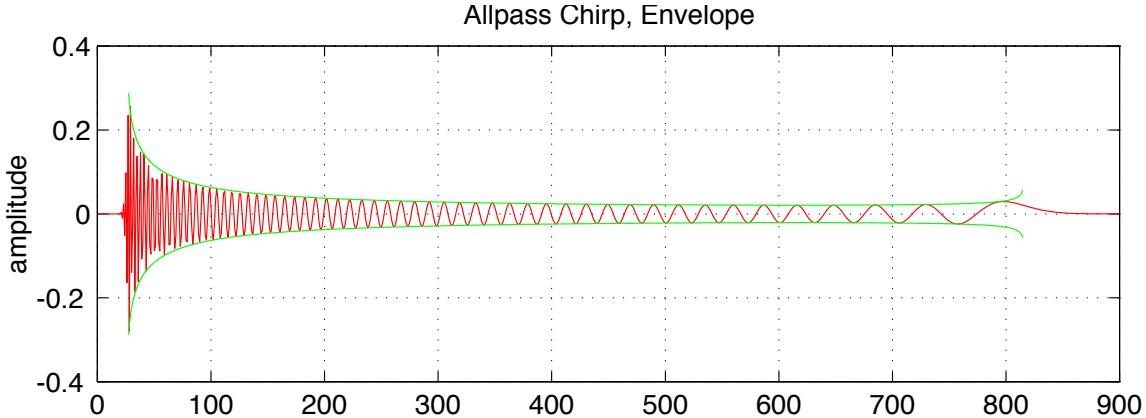
$$g(t) = \mathcal{F}^{-1} \left\{ e^{-j\phi(\omega)} \right\}.$$

- Group delay $\tau(\omega)$ is the derivative of the phase lag $\phi(\omega)$ with respect to frequency,

$$\tau(\omega) = \frac{d\phi(\omega)}{d\omega}.$$

→ If $\tau(\omega)$ is monotonic, and the impulse response sufficiently long lasting, a chirp results.

Allpass Chirp Envelope



What is the envelope $\alpha(\omega)$ of an allpass chirp $g(t)$ with group delay $\tau(\omega)$?

- Denote by ω_{\pm} two nearby frequencies having difference Δ and mean ω ,

$$\Delta = \omega_+ - \omega_-, \quad \omega = (\omega_+ + \omega_-)/2.$$

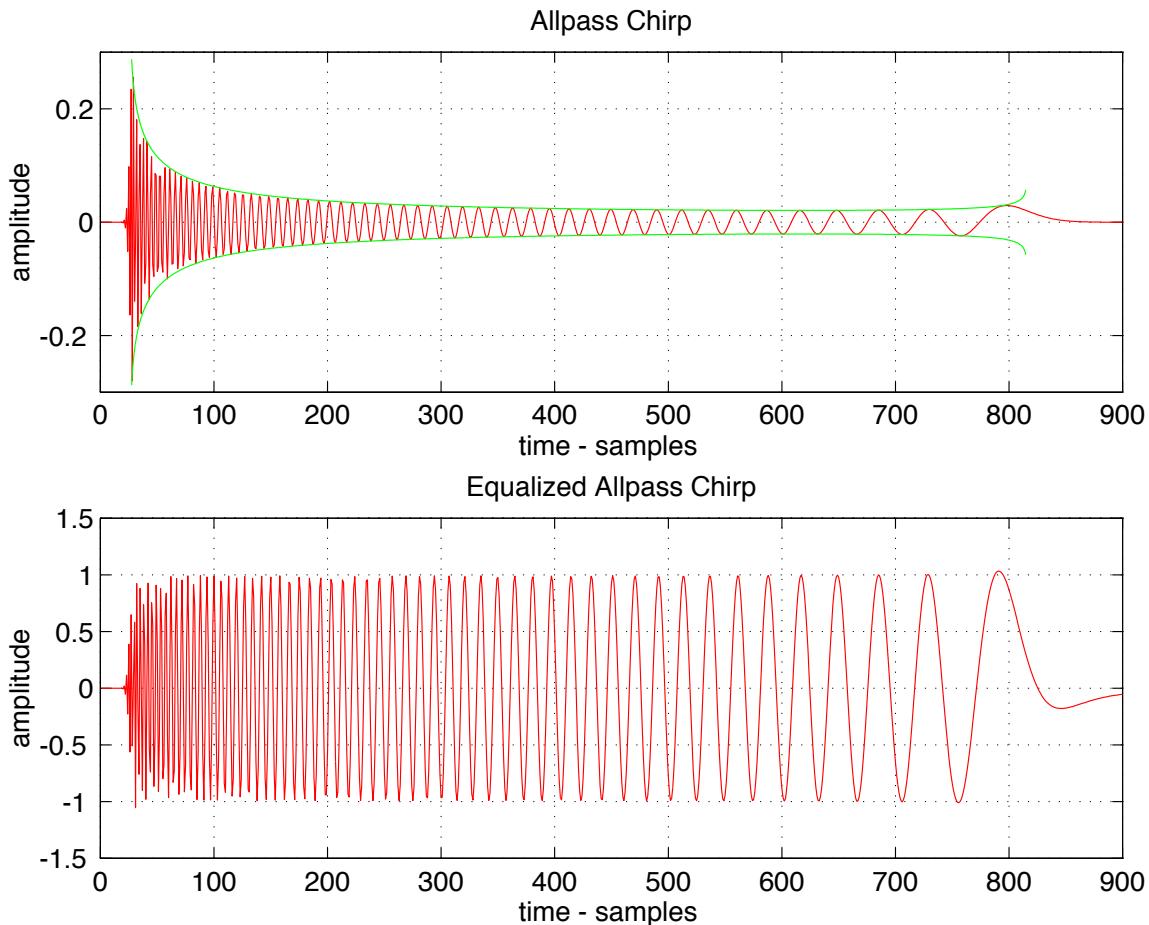
- The allpass chirp has energy Δ/π in the frequency interval $[\omega_-, \omega_+]$.
- That energy is roughly equal to the signal energy in the time interval $[\tau(\omega_-), \tau(\omega_+)]$,

$$\Delta/\pi \approx |\tau(\omega_-) - \tau(\omega_+)| \cdot \alpha^2(\omega)/2.$$

- Taking the limit $\Delta \rightarrow 0$, gives

$$\alpha(\omega) = \left[\frac{\pi}{2} \cdot \left| \frac{d\tau(\omega)}{d\omega} \right| \right]^{-\frac{1}{2}}.$$

Equalized Allpass Chirp

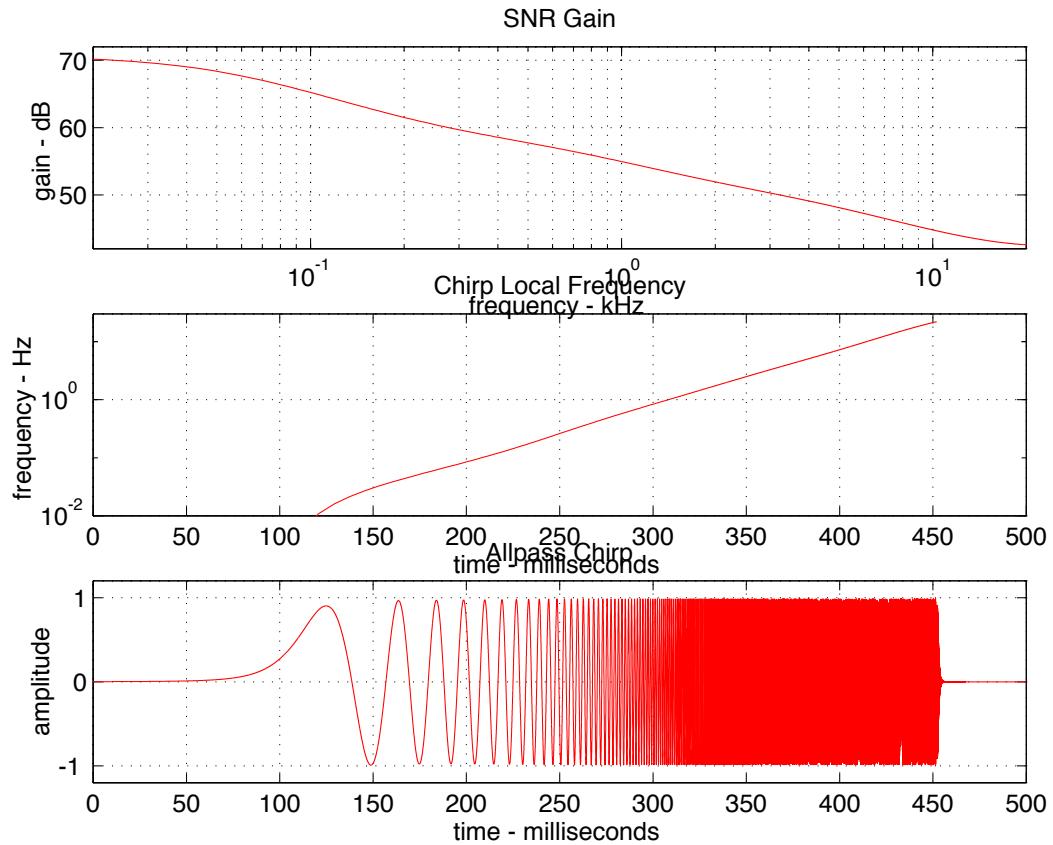


- An allpass chirp may be filtered by an invertible normalization $\nu(\omega)$ with a magnitude chosen to approximate the inverse of its envelope $\alpha(\omega)$,

$$|\nu(\omega)| \approx 1/\alpha(\omega) = \left[\frac{\pi}{2} \cdot \left| \frac{d\tau(\omega)}{d\omega} \right| \right]^{\frac{1}{2}}.$$

- Provided the filter doesn't smear the phase of the signal too much, the resulting sequence will have a nearly constant envelope.

Equalized Allpass Chirp Synthesis

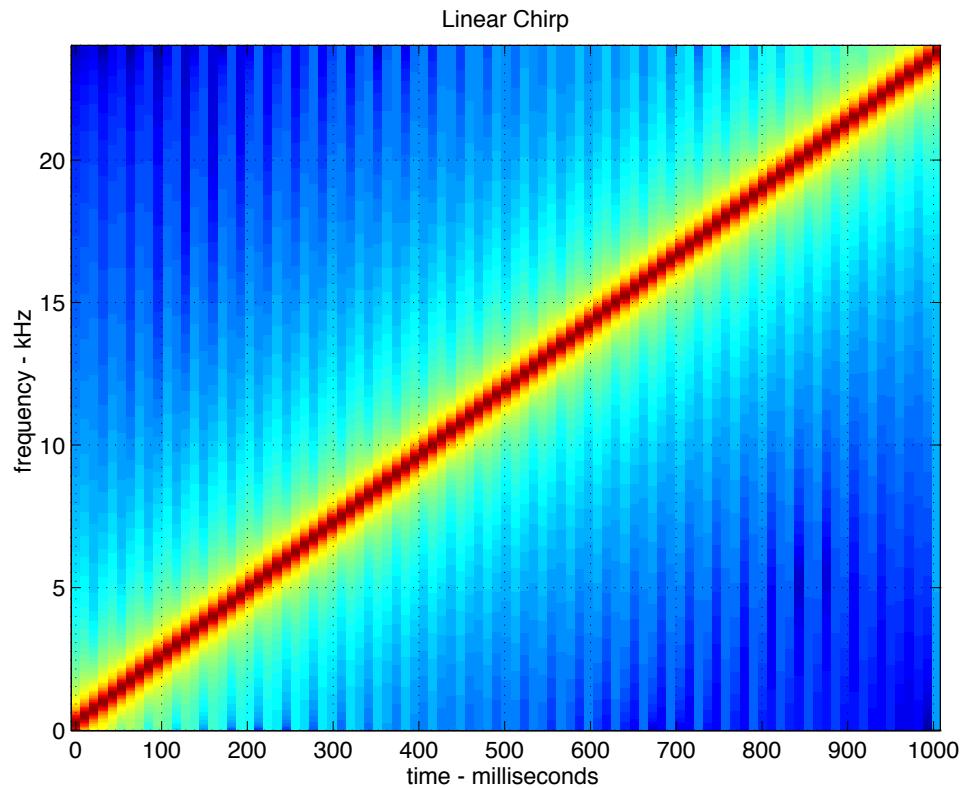


- Select $\nu(\omega)$ to achieve a desired SNR gain as a function of frequency.
- Based on $\nu(\omega)$, compute the allpass sequence group delay $\tau(\omega)$, and, in turn, the phase $\phi(\omega)$ and time series $g(t)$,

$$\tau(\omega) = \int_0^\omega |\nu(\xi)|^2 d\xi, \quad \phi(\omega) = \int_0^\omega \tau(\xi) d\xi.$$

$$g(t) = \mathcal{F}^{-1} \left\{ e^{-j\phi(\omega)} \right\}.$$

Linear Chirps

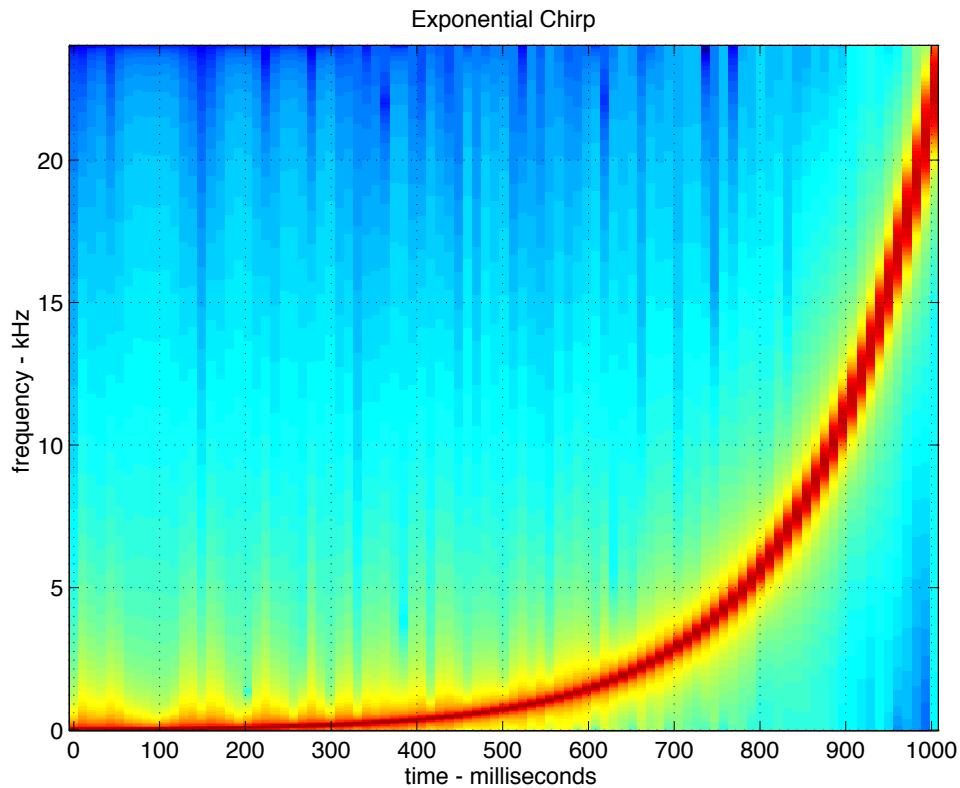


- Linear chirps take the same length of time to traverse any fixed bandwidth,

$$\tau(\omega) = \eta \cdot \omega.$$

- Linear chirps provide a constant SNR gain across frequency.

Exponential Chirps



- Exponential chirps take the same length of time to traverse any given octave,

$$\tau(\omega) = \eta \cdot \ln \omega.$$

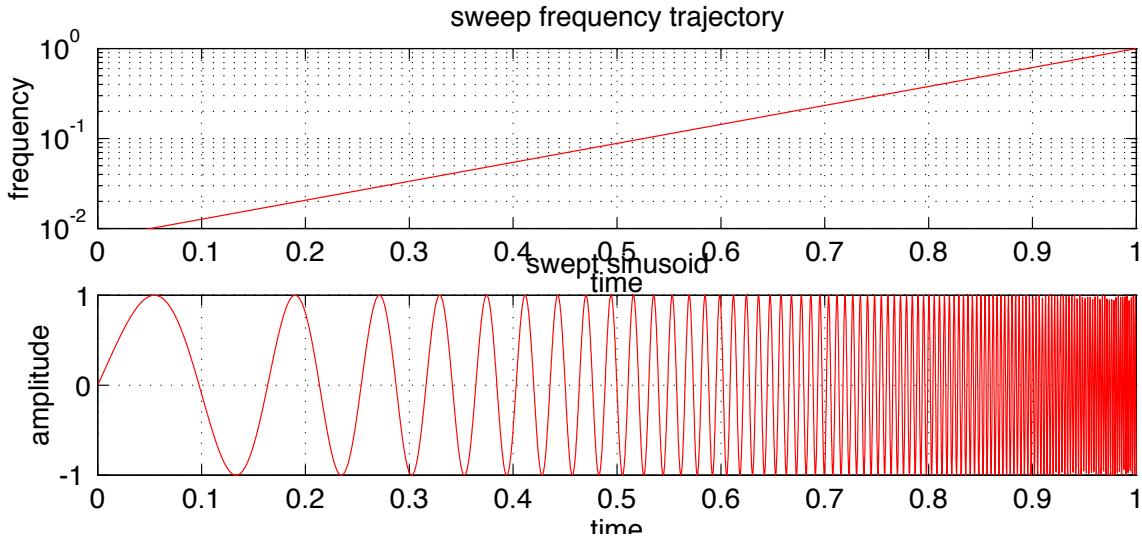
- Exponential chirps give an SNR gain inversely proportional to frequency,

$$|\nu(\omega)|^2 \propto \frac{d\tau(\omega)}{d\omega} = \eta/\omega;$$

they would be appropriate for measuring impulse responses in the presence of pink noise.

7. Swept Sinusoid Impulse Response Measurement

Swept Sinusoids



- The signal

$$s(t) = \sin \phi(t), \quad \phi(t) = \int_0^t \omega(\tau) d\tau$$

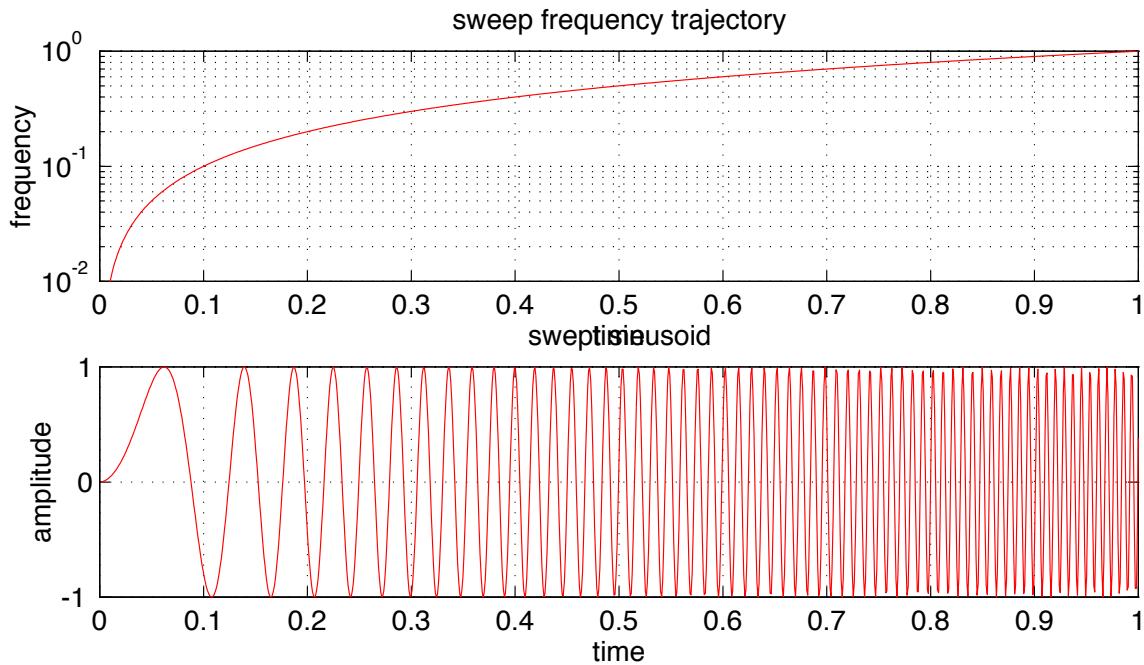
is a *swept sinusoid*, tracing out the frequency trajectory $\omega(t)$, assumed to be continuous and monotonic.

- When $\omega(t)$ is everywhere slowly changing, correlating the signal

$$\xi(t) = \nu(t) \cdot \sin \phi(t), \quad \nu(t) = 2 \left| \frac{d\omega}{dt} \right|,$$

with the sweep $s(t)$ will roughly produce an impulse, bandlimited to the initial and final frequencies of $\omega(t)$.

Linear Sweep



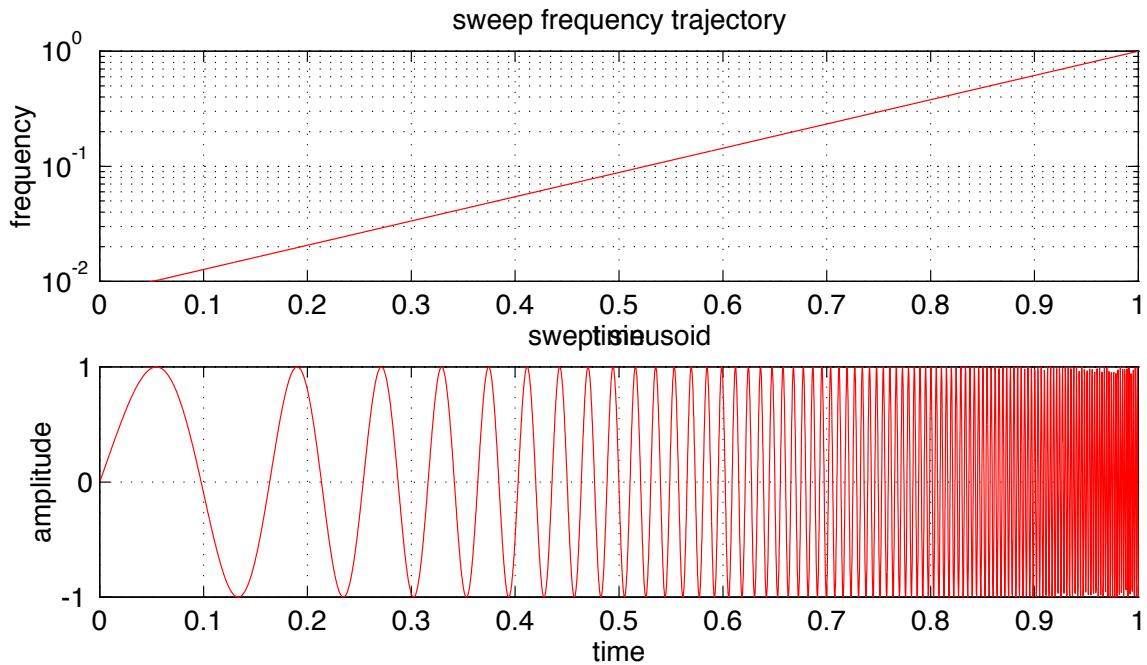
- A *linear sweep* is defined by the trajectory

$$\omega(t) = \eta \cdot t, \quad t \in [0, T].$$

- Its phase and correlating filter normalization are given by

$$\phi(t) = \eta \cdot t^2 / 2, \quad \nu(t) = 2\eta.$$

Exponential Sweep



- An *exponential sweep* is defined by the trajectory

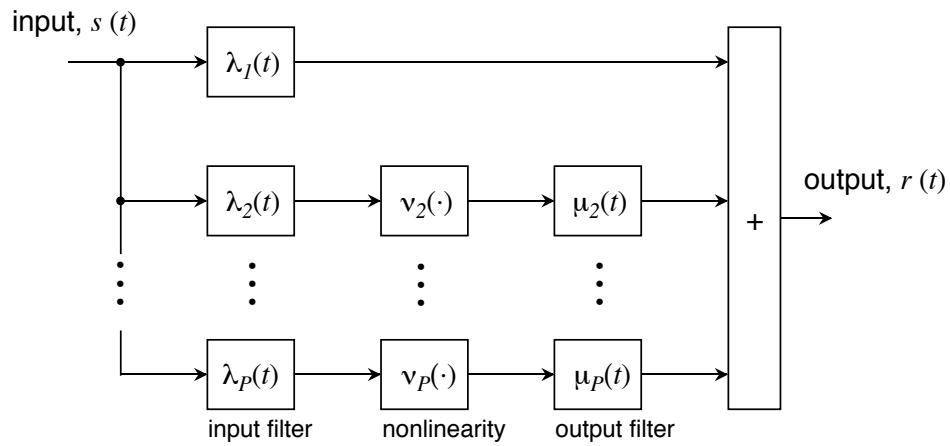
$$\omega(t) = \omega_0 \cdot \exp\{\eta \cdot t\}, \quad \eta = \frac{1}{T} \log(\omega_1/\omega_0), \quad t \in [0, T].$$

- Its phase and correlating filter normalization are given by

$$\phi(t) = \frac{\omega_0}{\eta} [\exp\{\eta \cdot t\} - 1], \quad \nu(t) = 2\omega_0 \eta \exp\{\eta \cdot t\},$$

Measuring Weakly Nonlinear Systems

Volterra Series

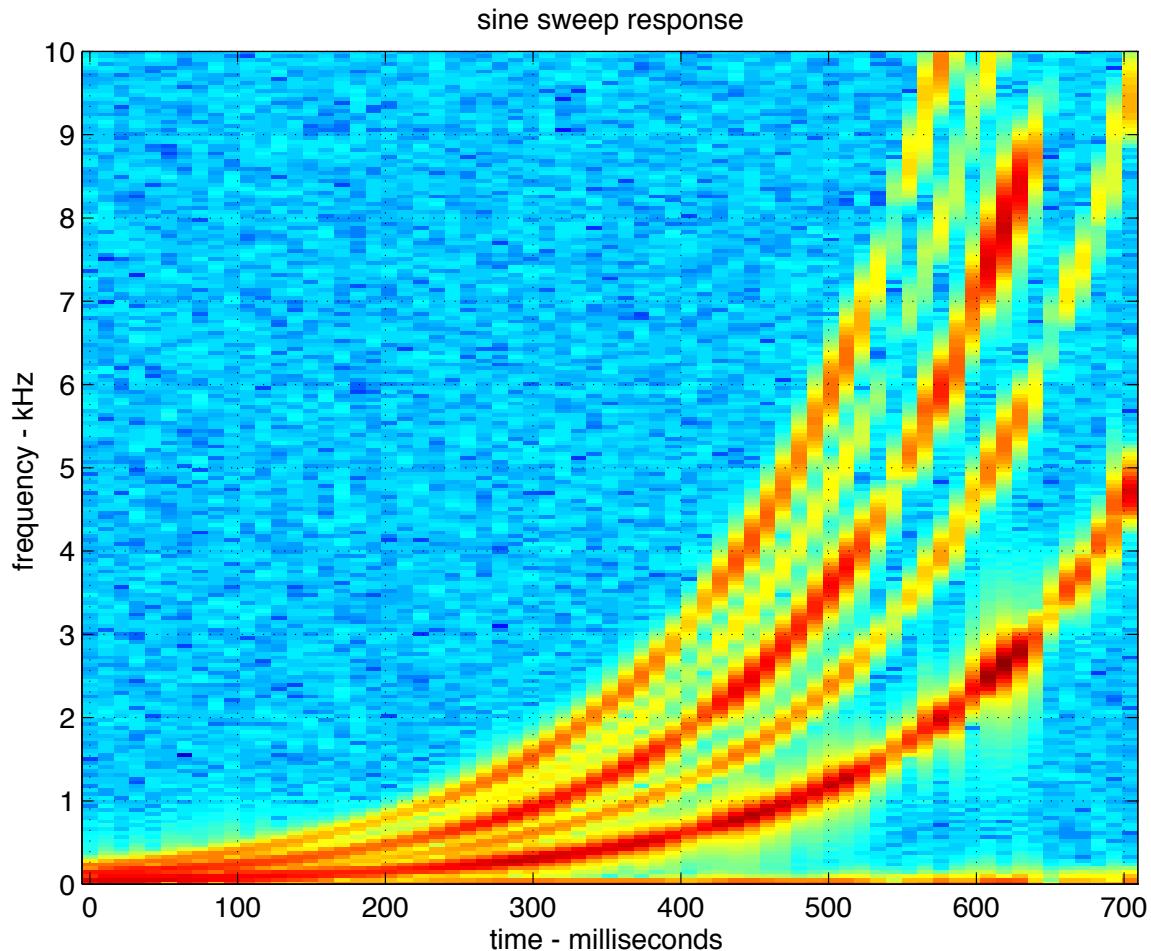


- Subject to regularity conditions, a system with filtering and memoryless nonlinearities connected via summing and splitting junctions can be represented by a Volterra series,

$$y(t) = \sum_{k=1}^K [x(t) * \lambda_k(t)]^k * \mu_k(t).$$

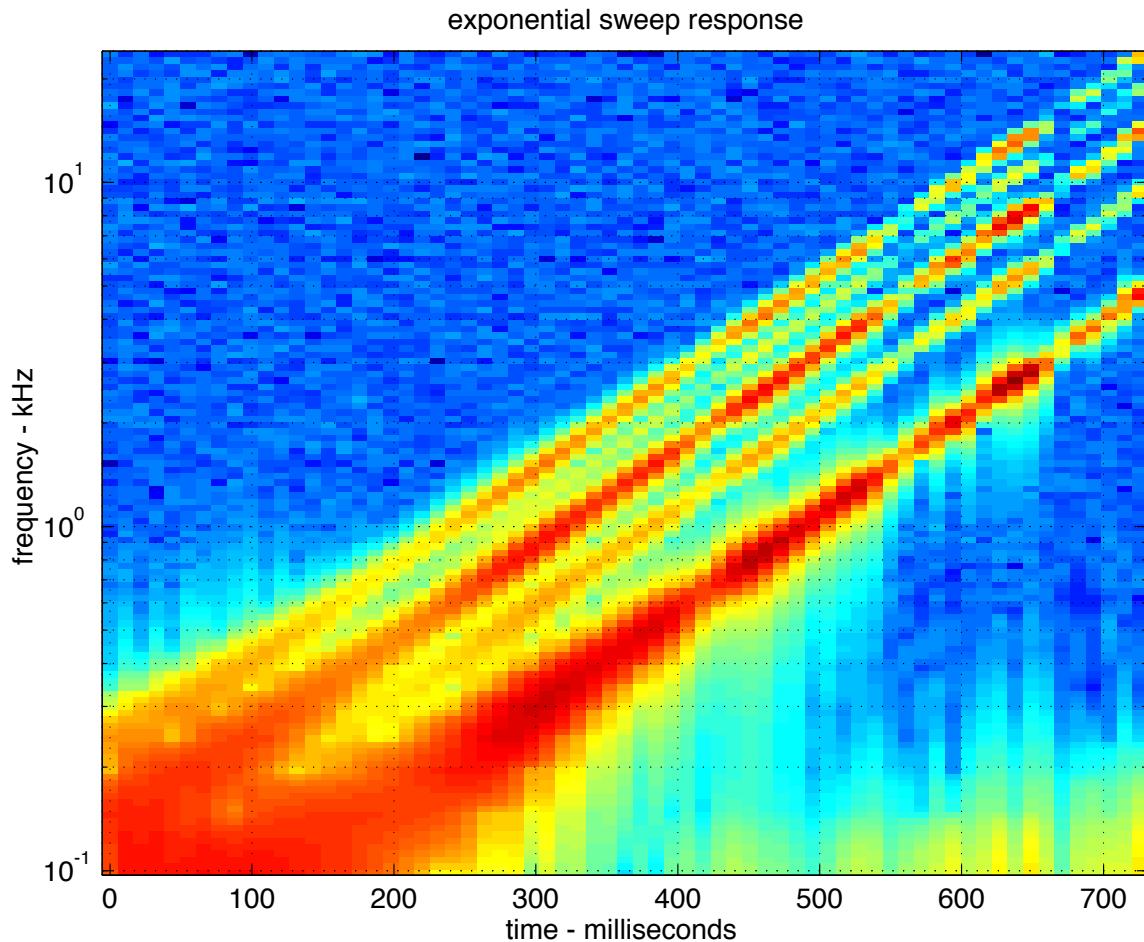
- Example: Guitar distortion is often modeled as the cascade of filtering and nonlinear operations,

$$y(t) = f(x(t) * g(t)) * h(t).$$

Volterra System Exponential Sine Sweep Response

- Weak nonlinearities can cause higher harmonics to appear in response to an input sinusoid.

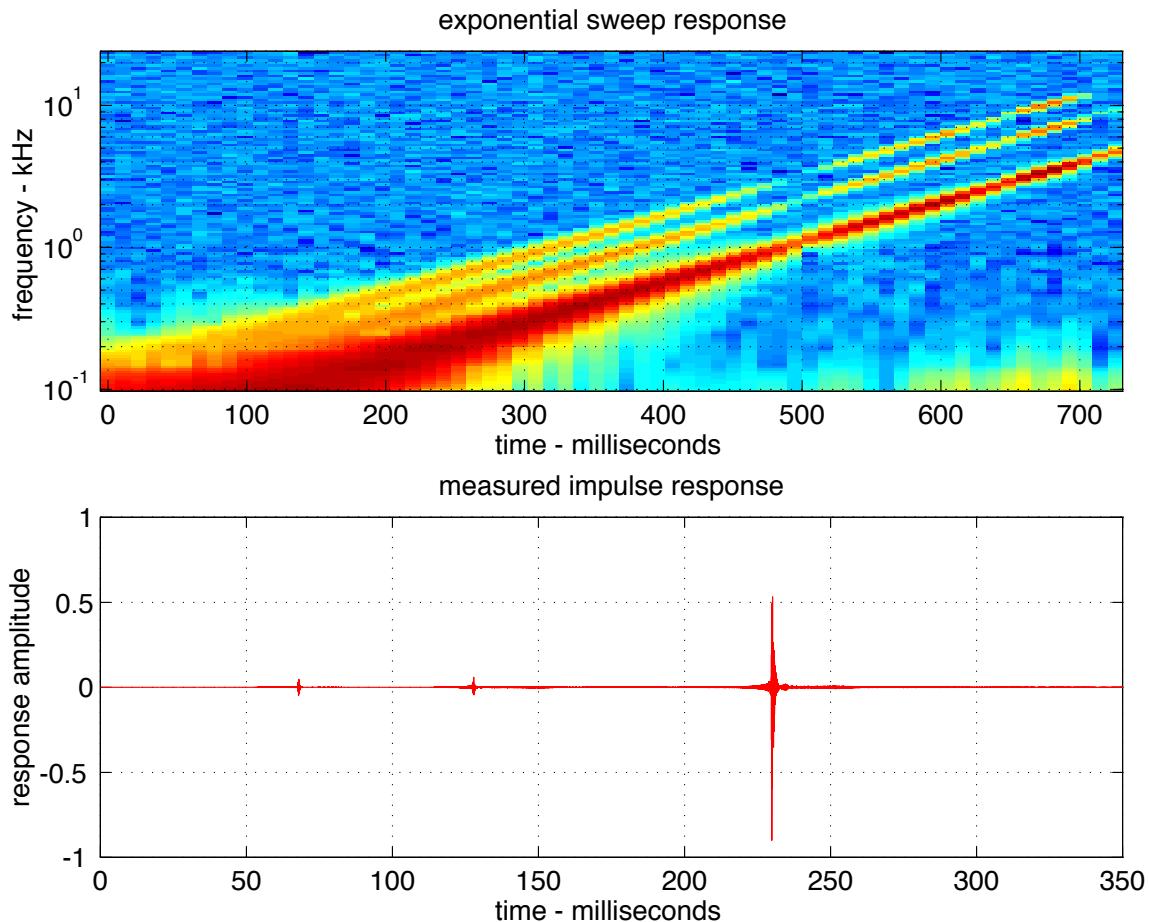
Volterra System Exponential Sine Sweep Response



- Harmonics in the system response trace out multiples of the input sine sweep frequency trajectory.
- When an exponential sweep is used, harmonics in the response may be viewed as typical system responses, but offset in time,

$$\omega_k(t) = \omega_0 \cdot k \exp\{\eta \cdot t\} = \omega_0 \cdot \exp\{\eta \cdot (t - t_k)\}, \quad t_k = (\log k)/\eta.$$

Volterra System Exponential Sine Sweep Response

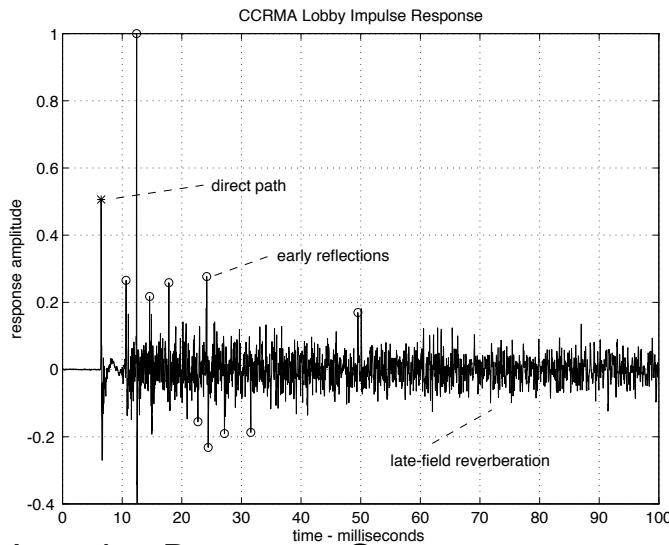


- Denoting by T the duration of the exponential sweep, and by ω_0 and ω_1 the starting and ending sweep frequencies, the k th harmonic will appear in the reconstructed impulse response as an impulse response-like feature, advanced relative to linear system response by

$$\frac{T \cdot \log k}{\log(\omega_1/\omega_0)}.$$

Reverberation

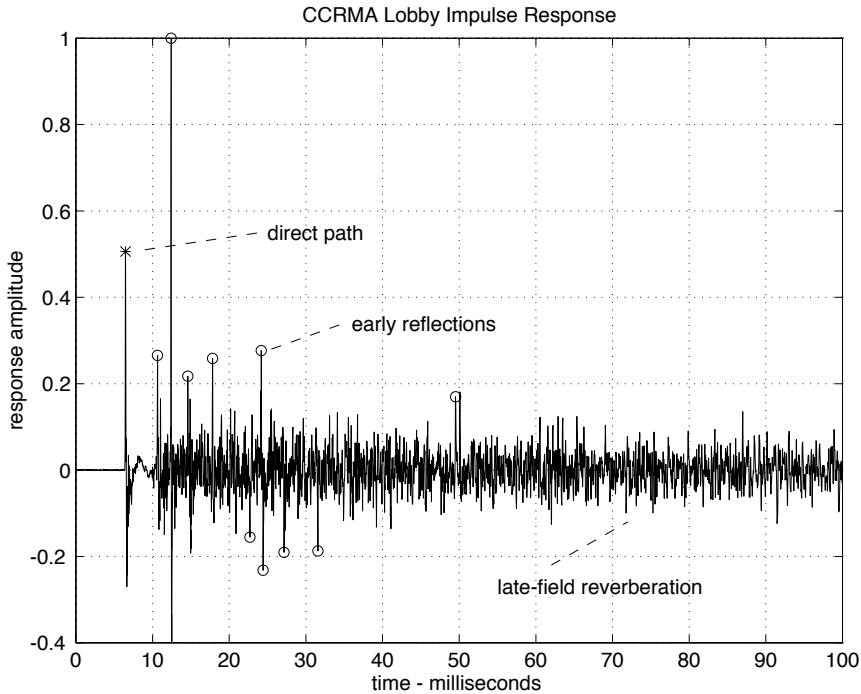
Reverberation Unit Overview



Reverberation Impulse Response Components

- *Direct Path.* The direct path is the arrival of source signals along a straight-line path from the source. The arrival time and amplitude are fixed by the source-listener distance; filtering dependent on the source direction and orientation are also present.
- *Early Reflections.* Typically a small number of specular reflections from environment surfaces will arrive at the listener well separated in time or amplitude from other reflected energy. These so-called early reflections convey a sense of the environment geometry and size.
- *Late-Field Reverberation.* After a period of time, source signals have interacted sufficiently with room objects and surfaces that the density of arrivals is great enough to be indistinguishable from noise. This late-field reverberation is further characterized by a frequency-dependent exponential decay determined by the environment size and materials present.

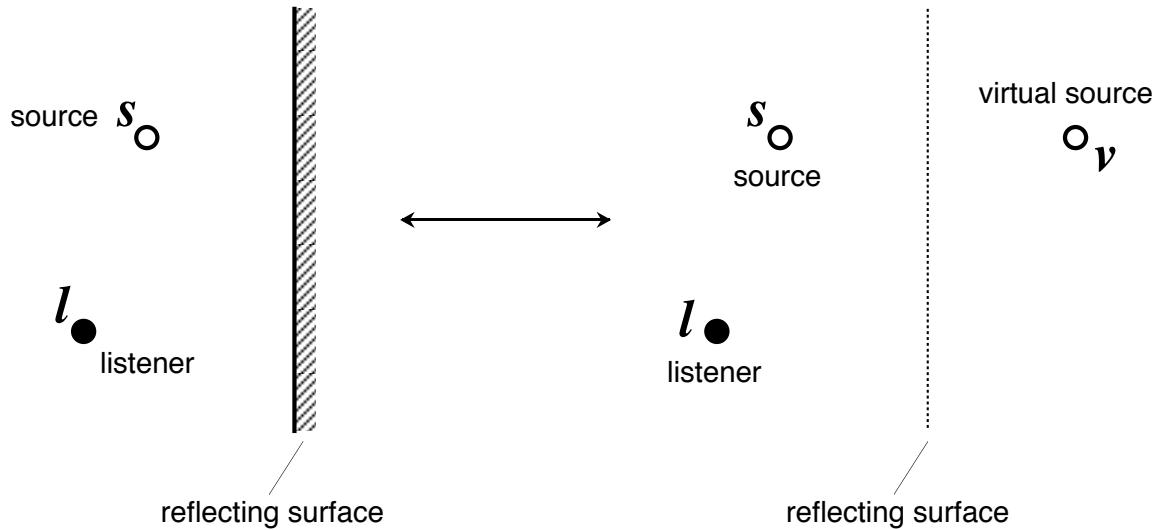
Reverberation Unit Outline



- Physical Acoustics
 - physical origin of impulse response components
- Impulse Response Analysis
 - estimation of psychoacoustically relevant impulse response features
- Psychoacoustics
 - impulse response feature perception
- Synthesis
 - impulse response synthesis and convolutional reverberation methods

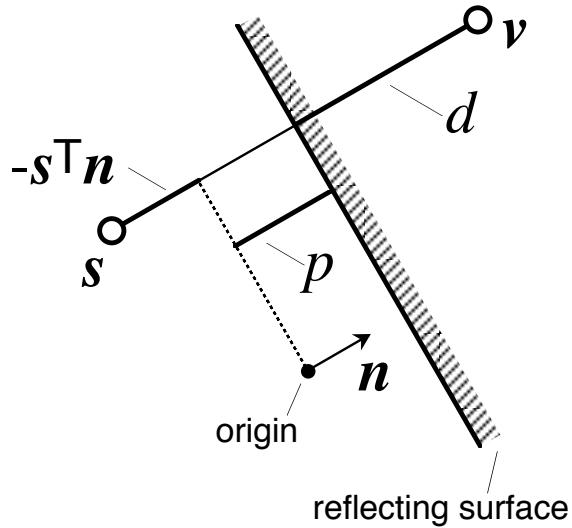
8. Specular Reflections and the Image Method

The Image Method



Single Reflection.

- A reflection is equivalent to two sources, the actual source and a reflected or “virtual” source.



Reflection Geometry.

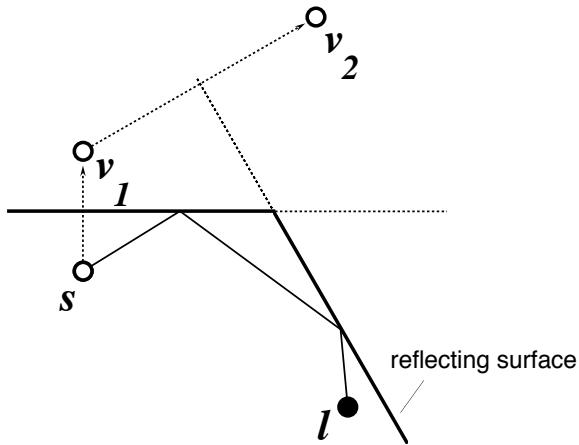
- Consider a reflecting plane located a distance p from the coordinate origin, and having normal vector \mathbf{n} .
- Denoting by \mathbf{s} the column of source position coordinates, the reflected source position \mathbf{v} is the source position plus twice the distance d from the source to the reflecting plane along the direction normal to the plane.

$$\mathbf{v} = \mathbf{s} + 2d\mathbf{n}.$$

- The distance between the source and reflecting plane is the distance from the plane to the origin, less the projection of the source position on the plane normal,

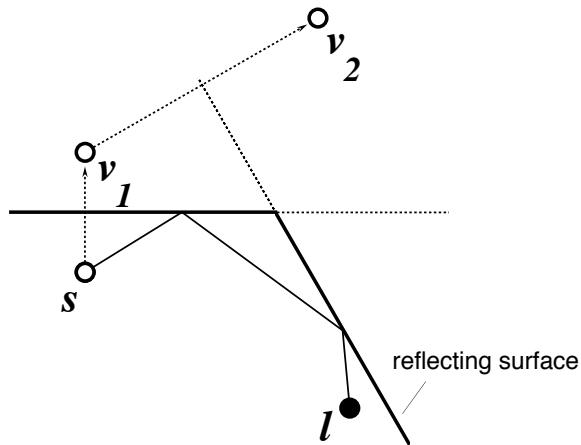
$$d = p - \mathbf{s}^\top \mathbf{n}.$$

The Image Method.



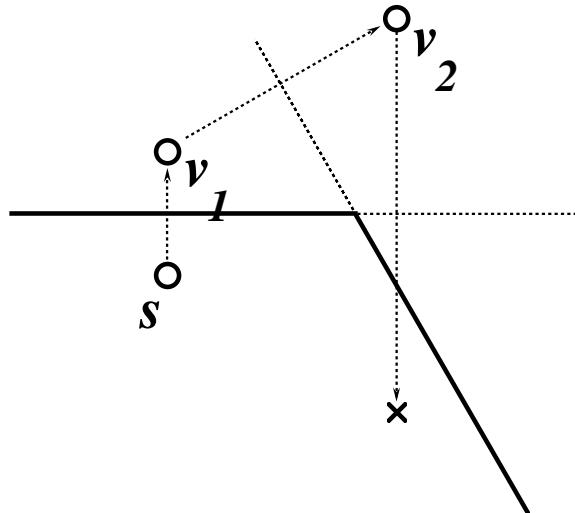
- Consider a pair of reflecting planes.
- A source may be reflected through the first plane, and then the second to create a so-called *second-order* reflection. In general, the number of surfaces a reflection has encountered is the *order* of the reflection.
- In an enclosed space, as signals radiate out from the source, they encounter reflecting surfaces, creating reflections. The reflections, in turn interact with other surfaces creating additional reflections.
- The signal arriving at a listener in an idealized enclosed space is equivalent to the signal arriving at the listener from the source and a set of properly located (and oriented) virtual sources.

Computing the Virtual Source Locations (Convex Spaces).



- Start with the actual source, and reflect it through each of the surfaces in the space to create a set of first-order reflected sources.
- Iteratively generate the set of next-order reflections by taking each reflection of the current order and reflecting it through each surface, checking for *validity* and *visibility*.
 - Only the inside of the room walls are “mirrored,” and reflections through the outside of the surfaces are invalid, and should be discarded.
 - Not all reflections are visible to a given listener, and invisible reflections should be marked as such. Invisible reflections should be maintained, as they might generate higher-order visible reflections.

Reflection Validity.

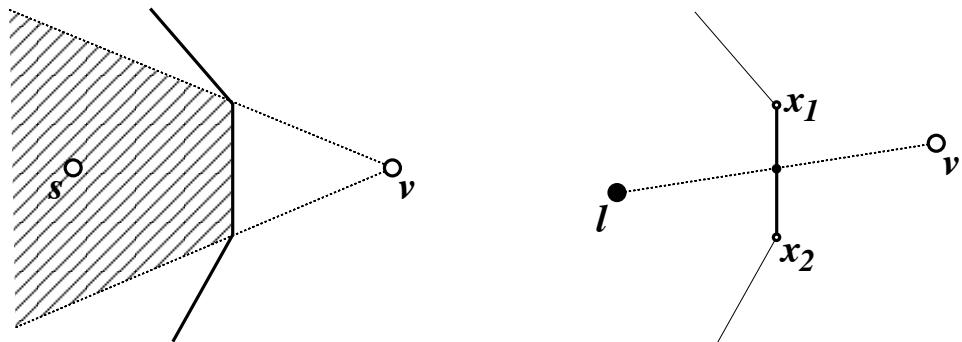


- A reflection is valid if the signed distance d

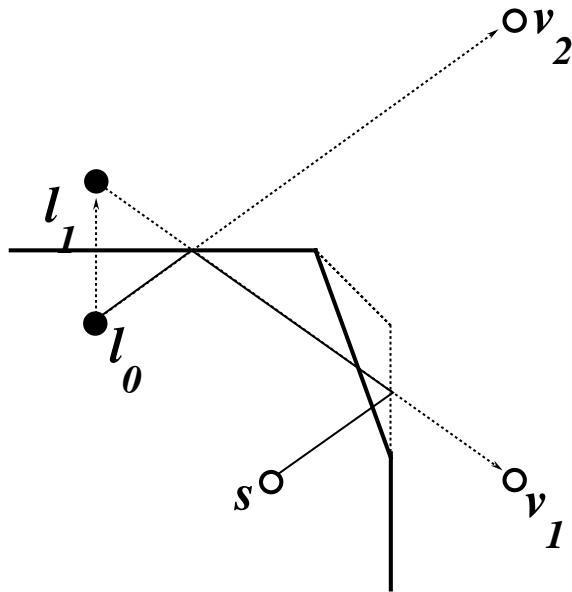
$$d = p - s^\top n$$

along the normal to the plane pointing outward from the room n is positive. Otherwise the reflection is through the back of the surface.

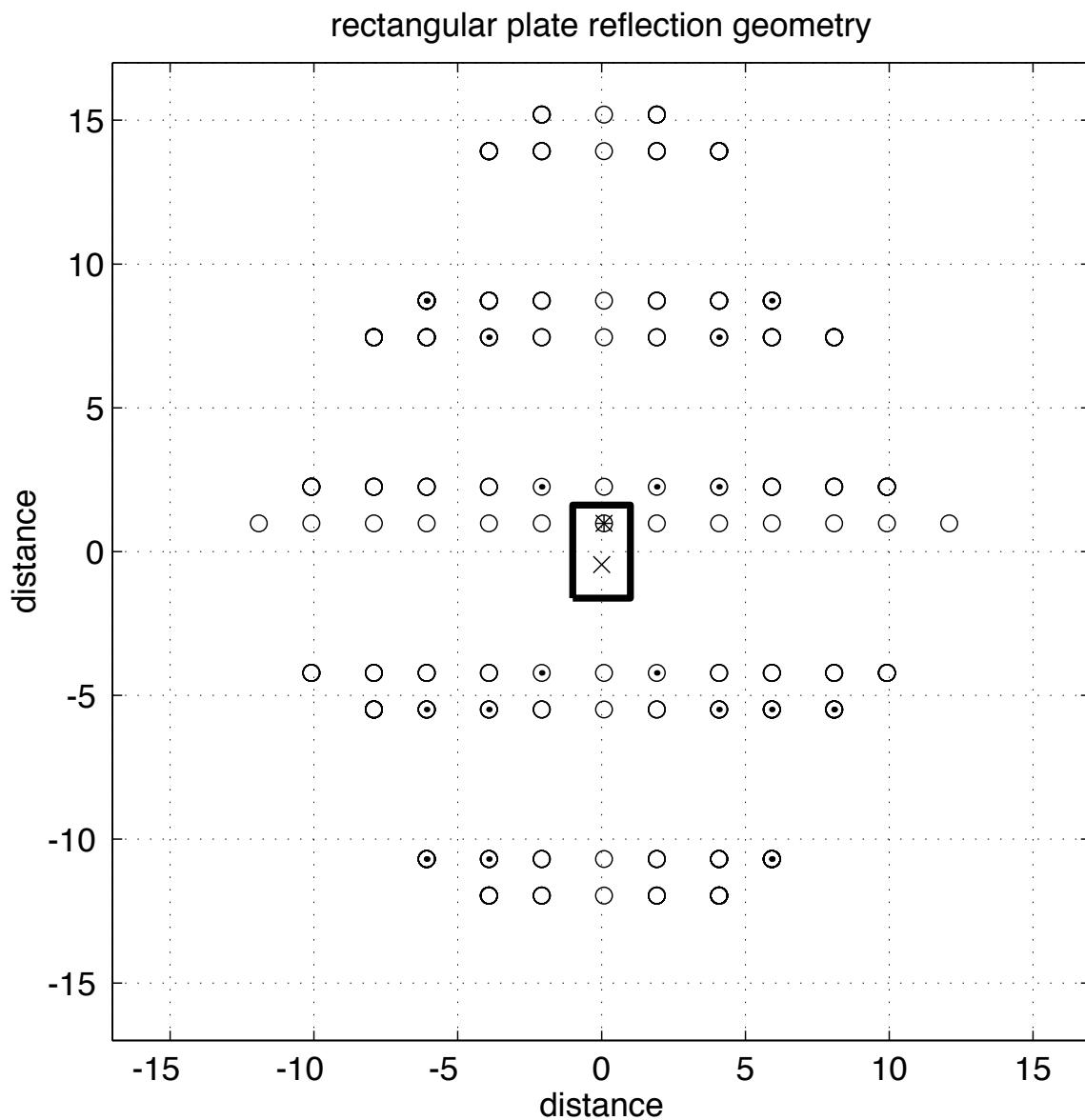
Reflection Visibility.

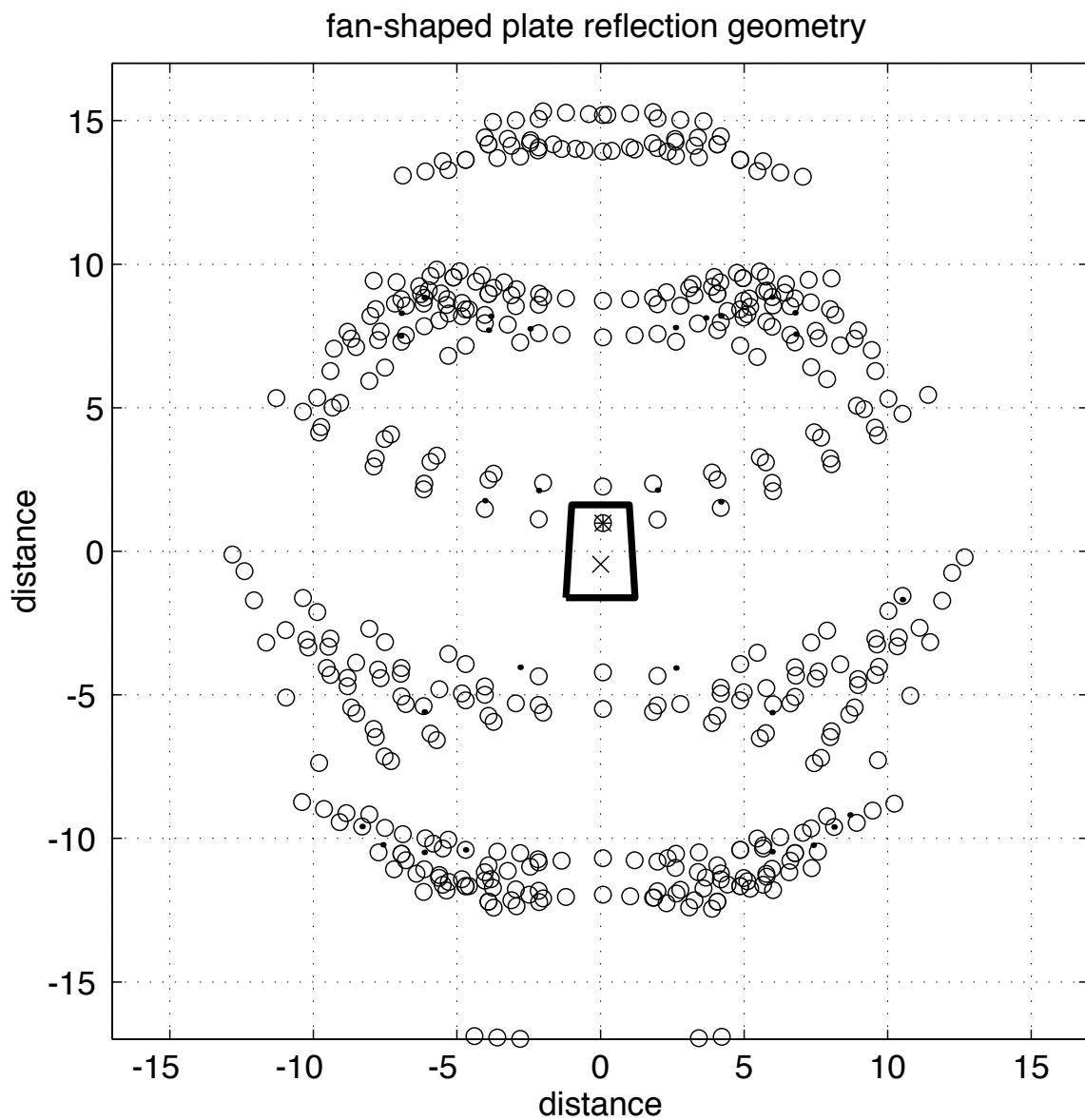


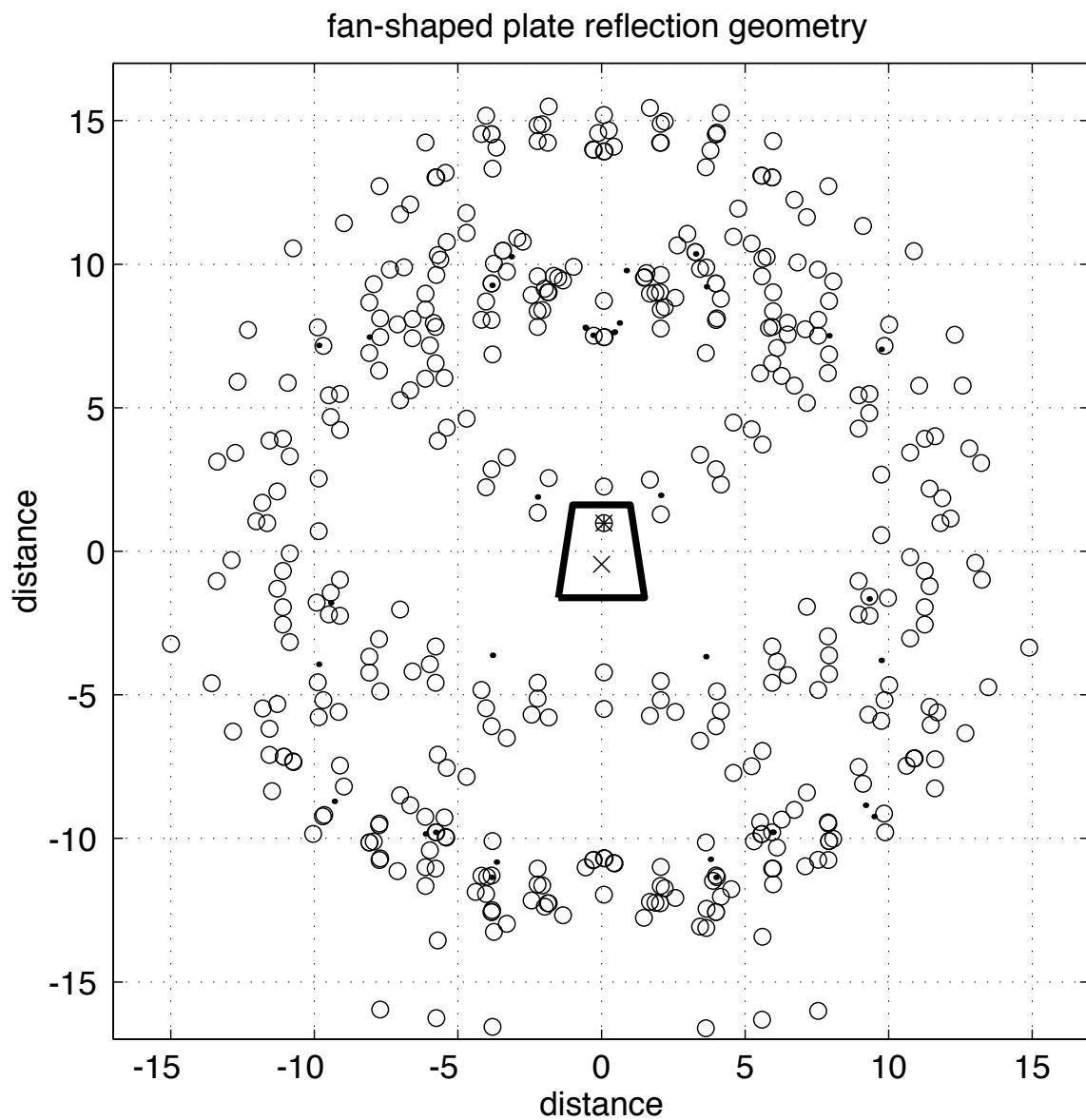
- A reflection is visible if a line connecting it and the listener passes through the reflecting surface.

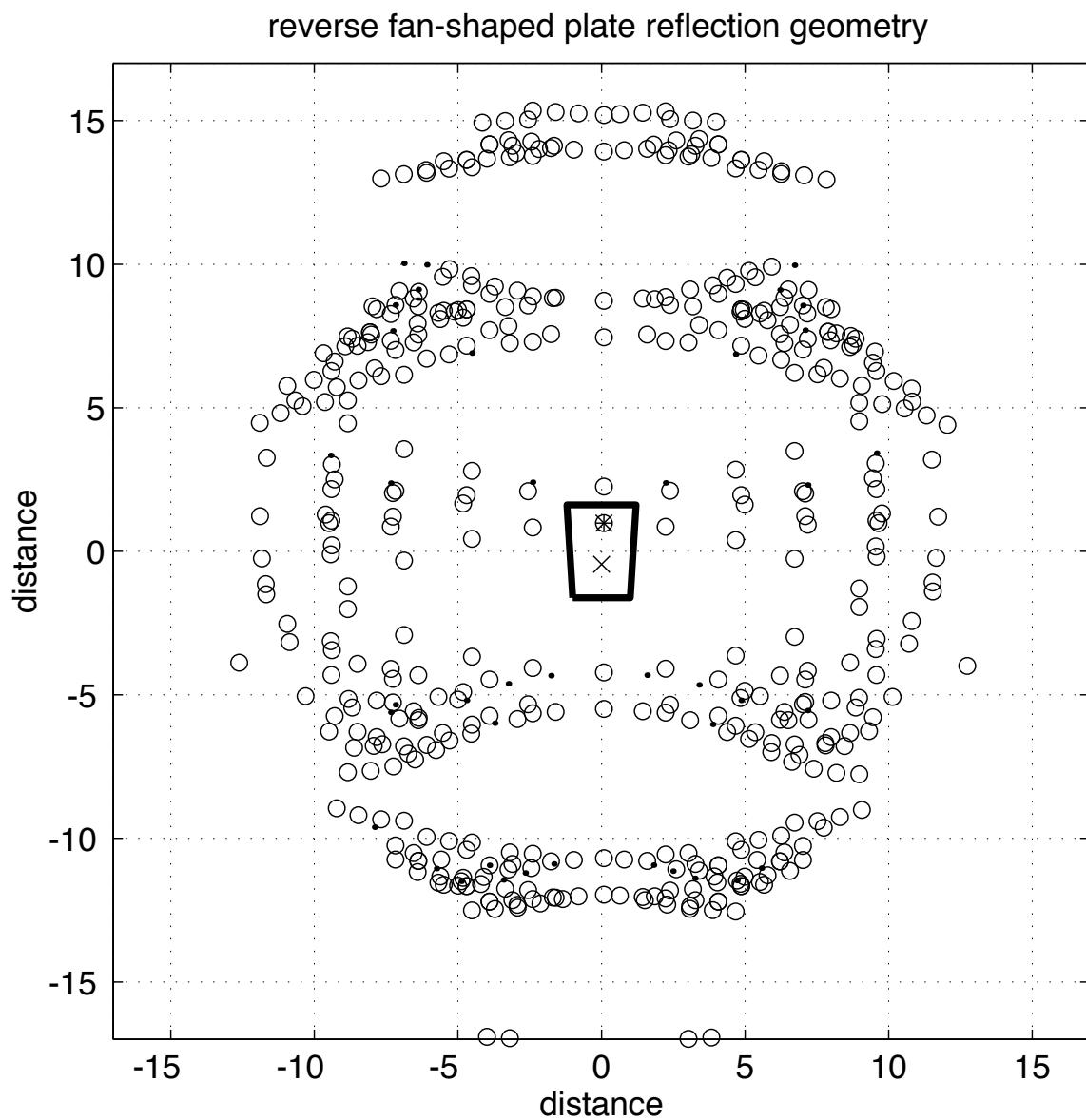


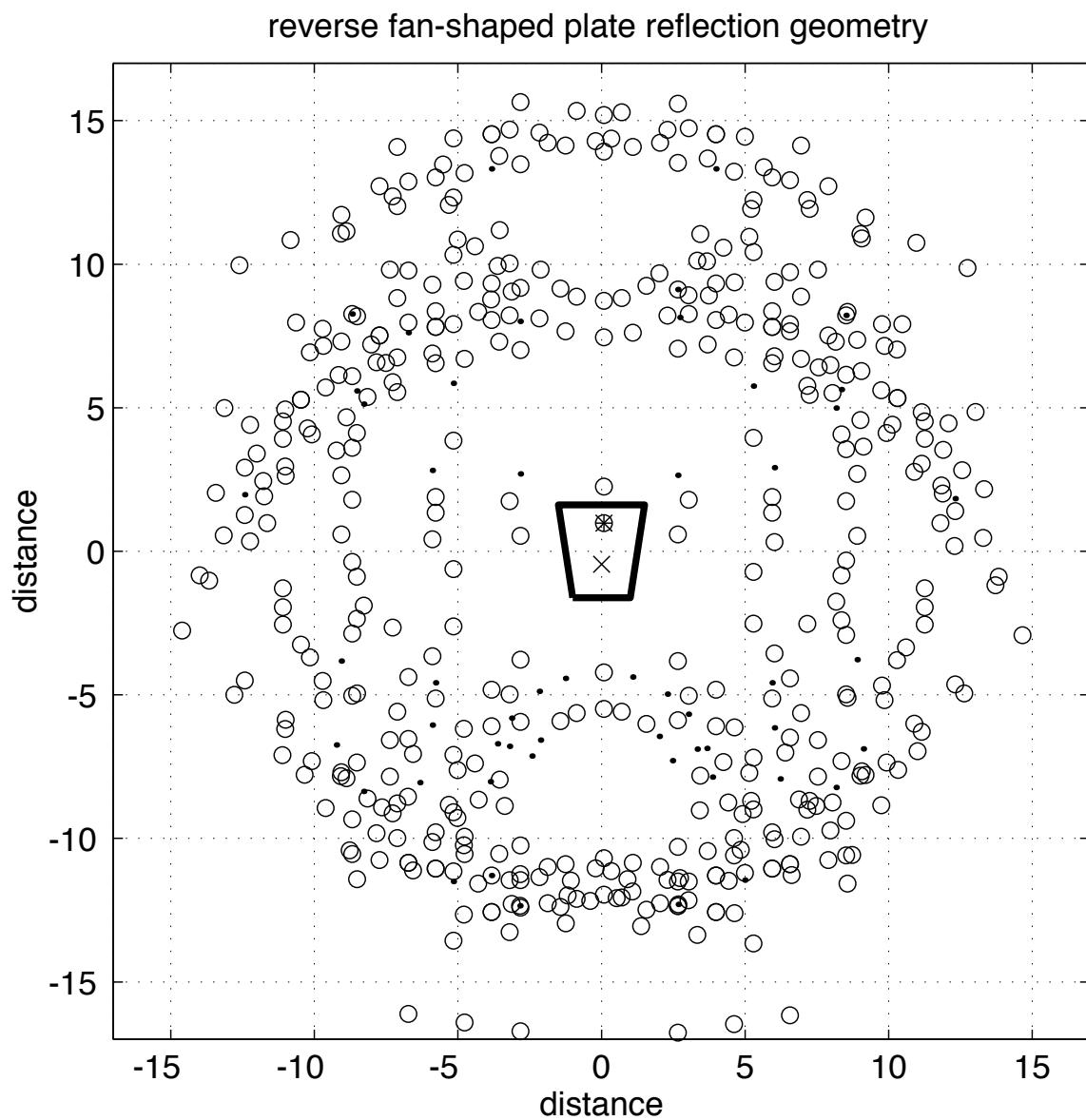
- Visibility must be checked for each reflecting surface along the path from virtual source to listener.
 - To do this, check that the listener ℓ_0 is visible from the virtual source v_n through the last (say, n th) reflecting surface.
 - Then reflect listener ℓ_0 through the last reflecting surface to create the reflected listener ℓ_1 , and check its visibility of v_{n-1} (the virtual source which generated v_n) through the surface which generated v_n .
 - In a similar manner, check the visibility of all virtual sources in the chain generating v_n .

Example Room Reflections.









Impulse Response Synthesis.

- To synthesize an impulse response, each virtual source must be processed according to the physics of the propagation path.
 - Associated with the virtual source-listener distance $r = \|\mathbf{v} - \mathbf{l}\|$ is a time delay equal to the source-listener range scaled by the sound speed in air,

$$\tau = r/c,$$

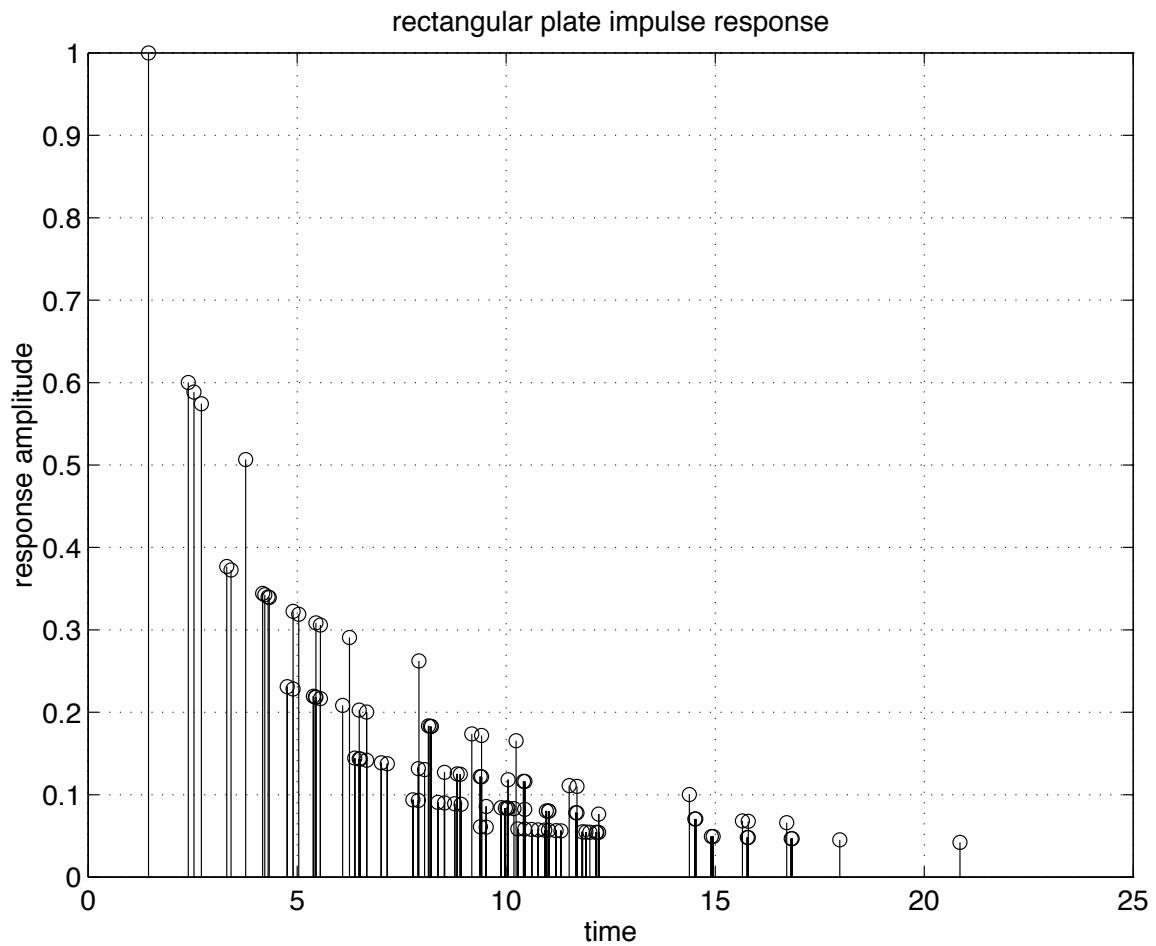
a spherical spreading loss scaling arriving signals by

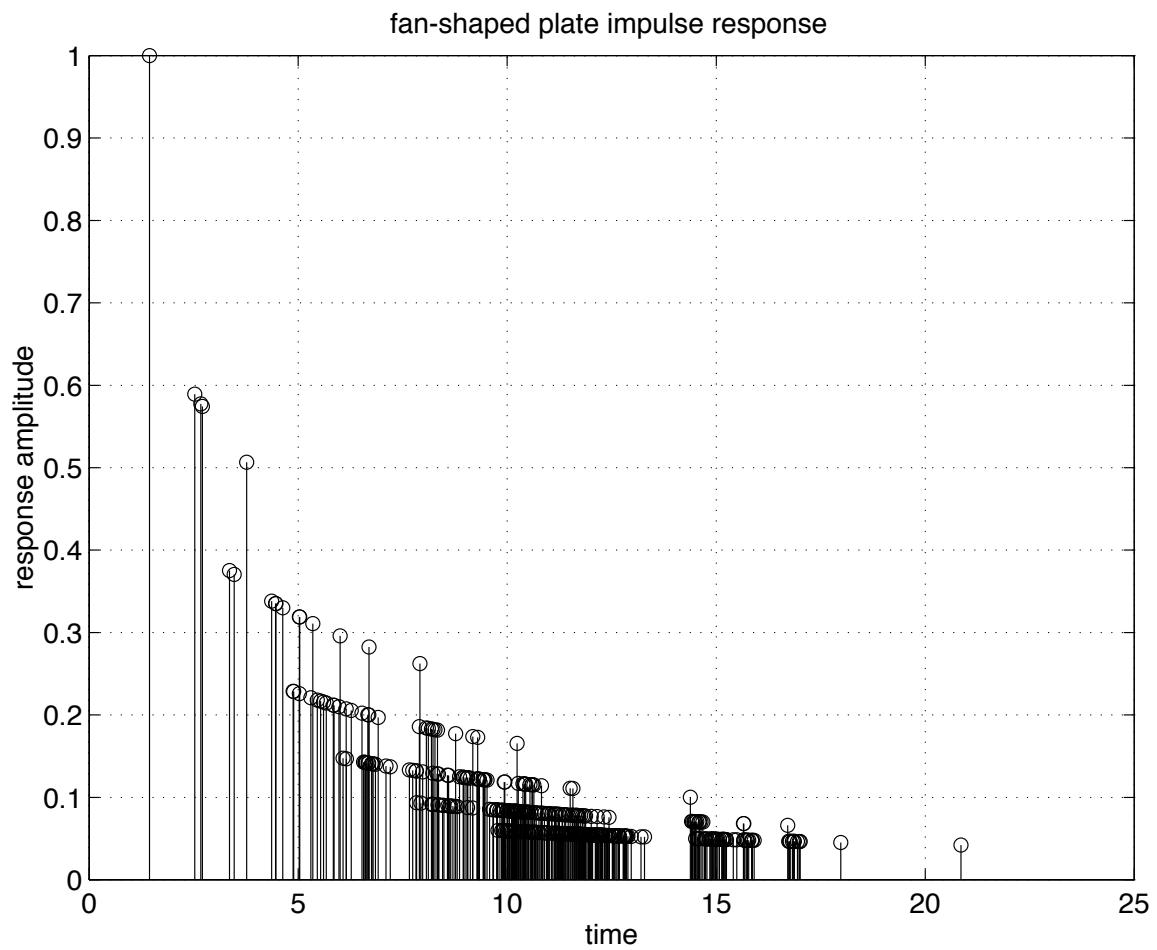
$$\gamma = \frac{1}{1 + r/\rho},$$

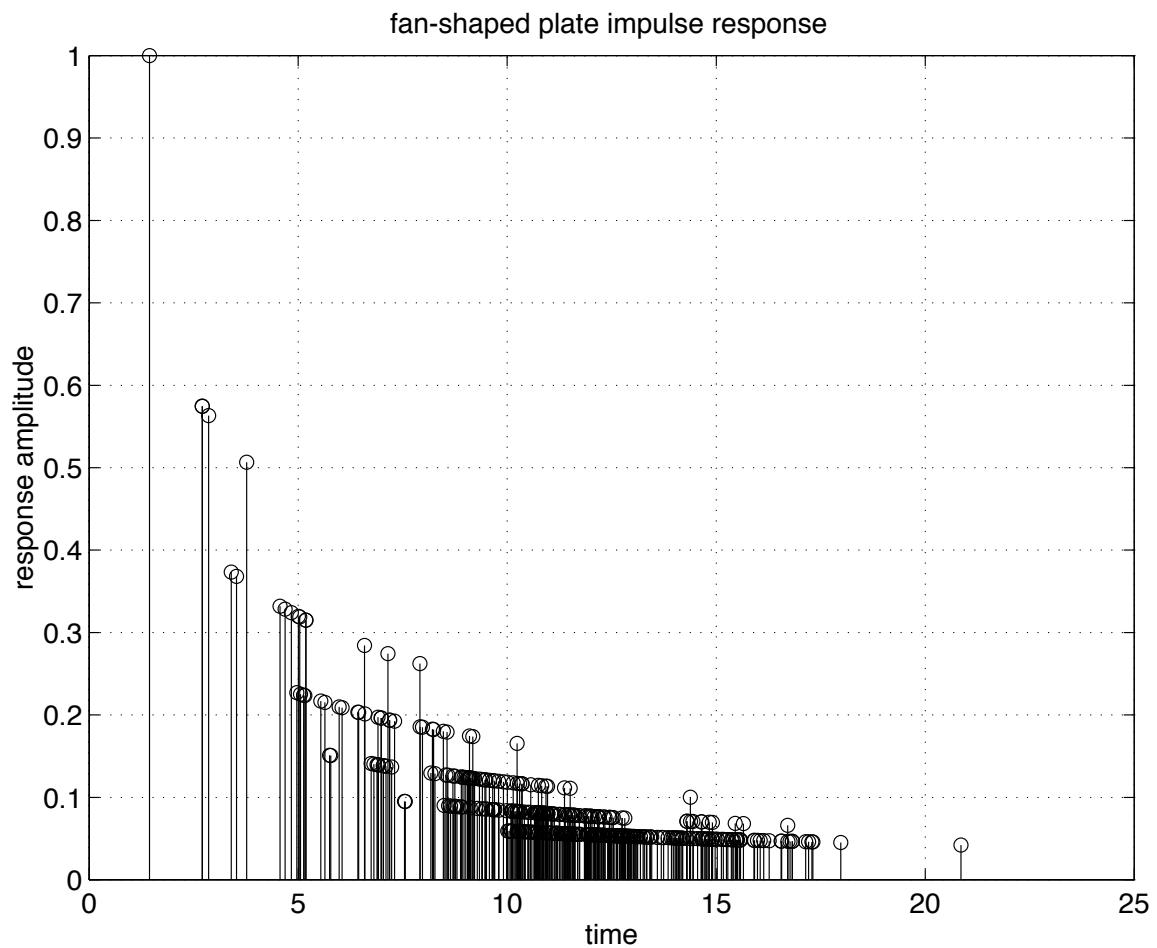
where ρ is roughly the size of the source, and air absorption filtering arriving signals according to

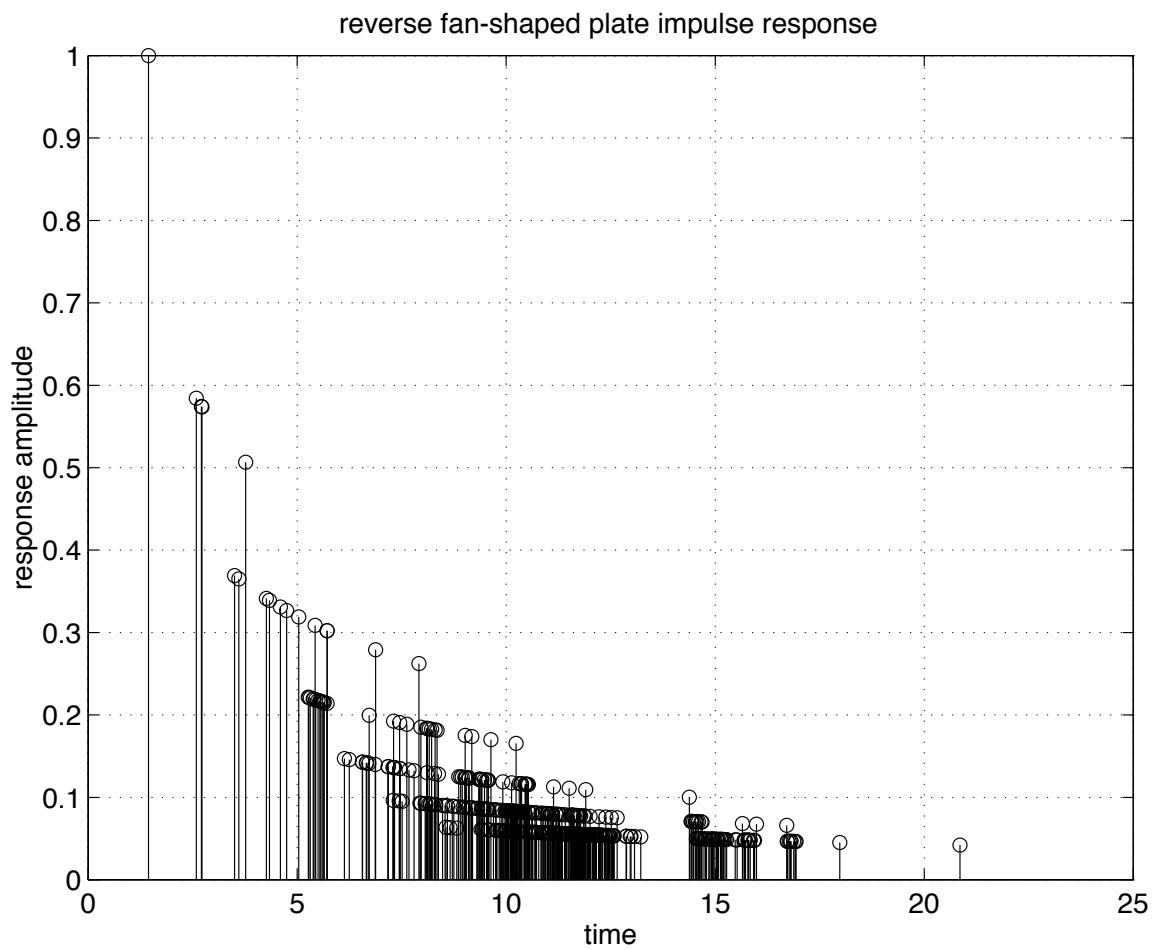
$$\exp(-\alpha(\omega) \cdot r).$$

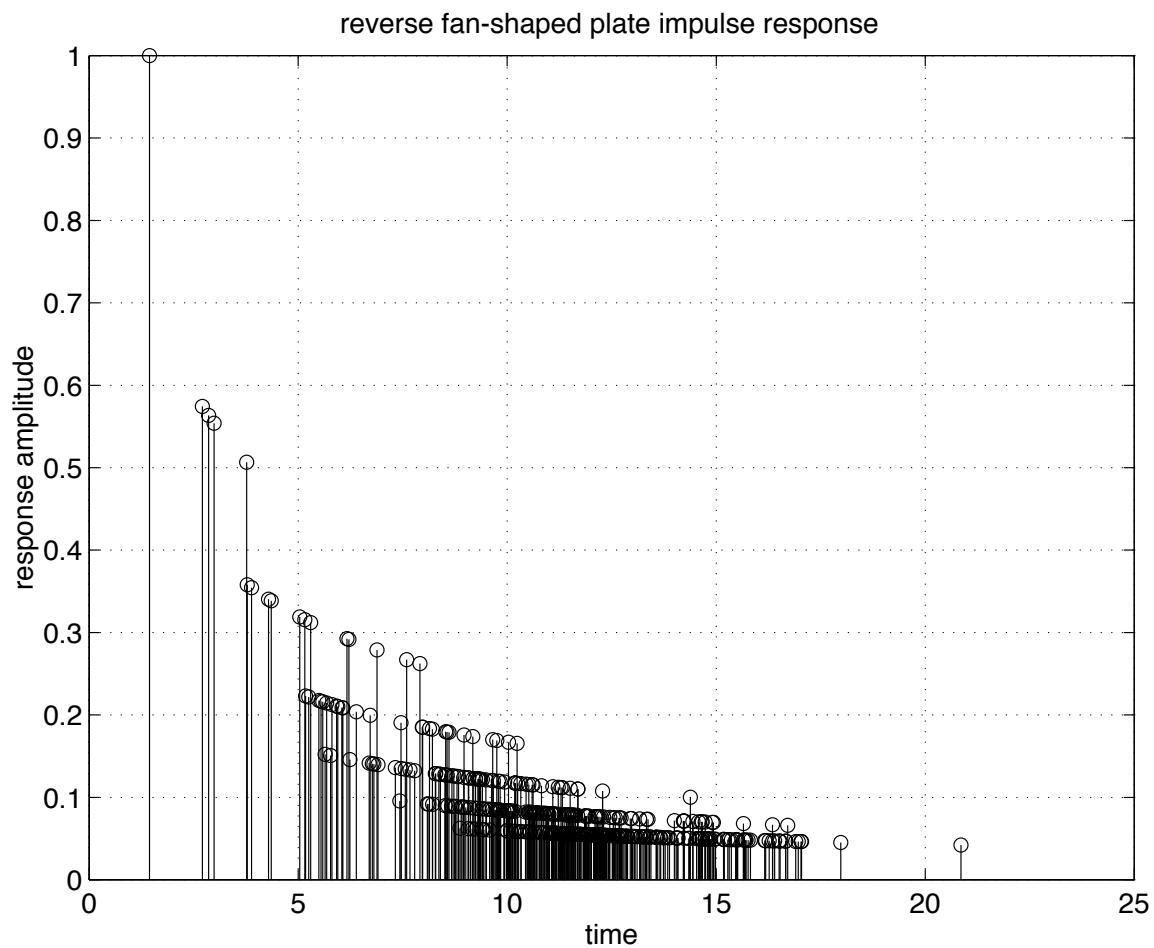
- Virtual sources should be filtered according to materials encountered.
- Finally, virtual sources are also filtered according to any source radiation pattern and listener antenna pattern, as determined by the source-listener direction.

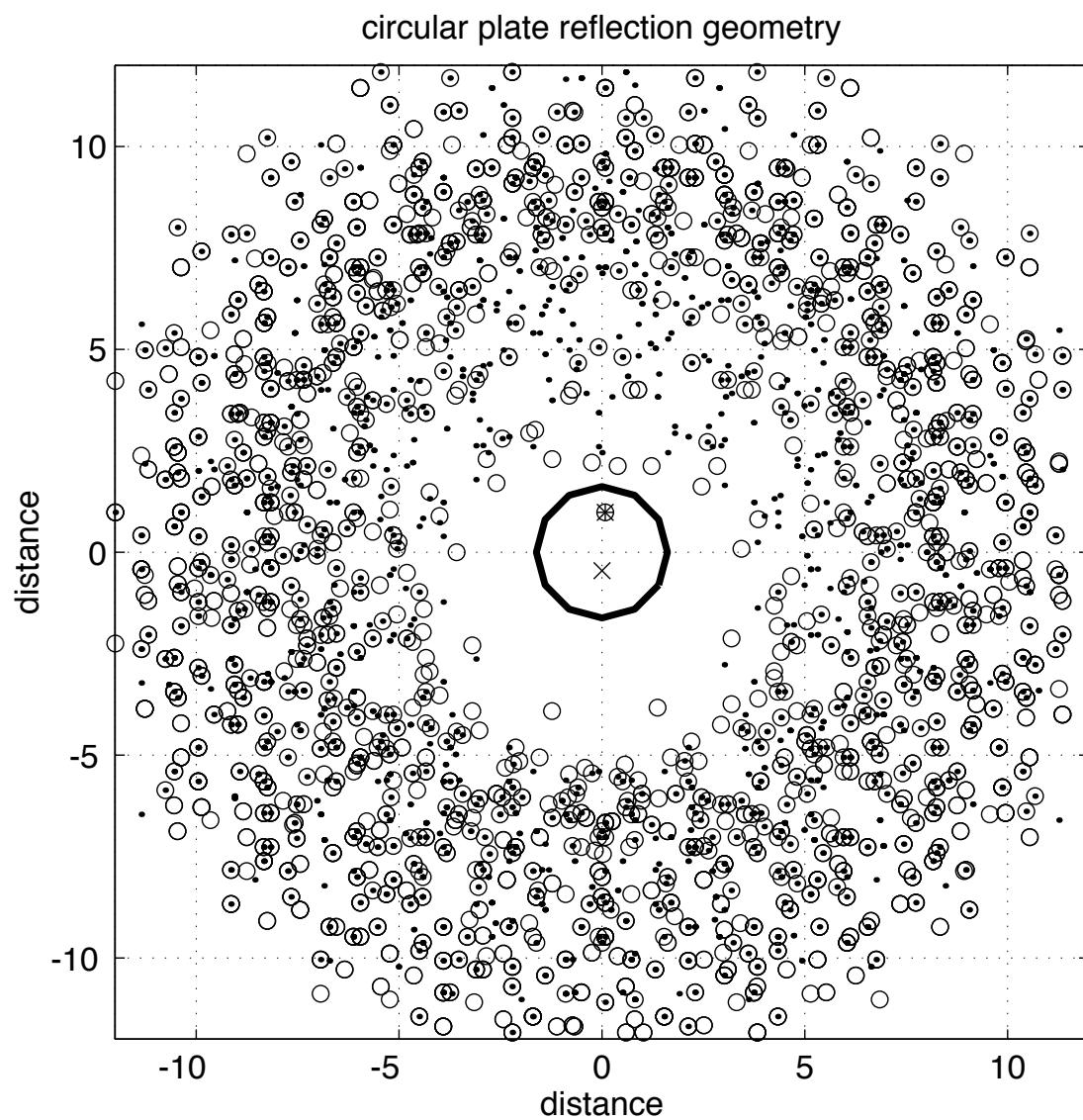
Example Room Impulse Responses.

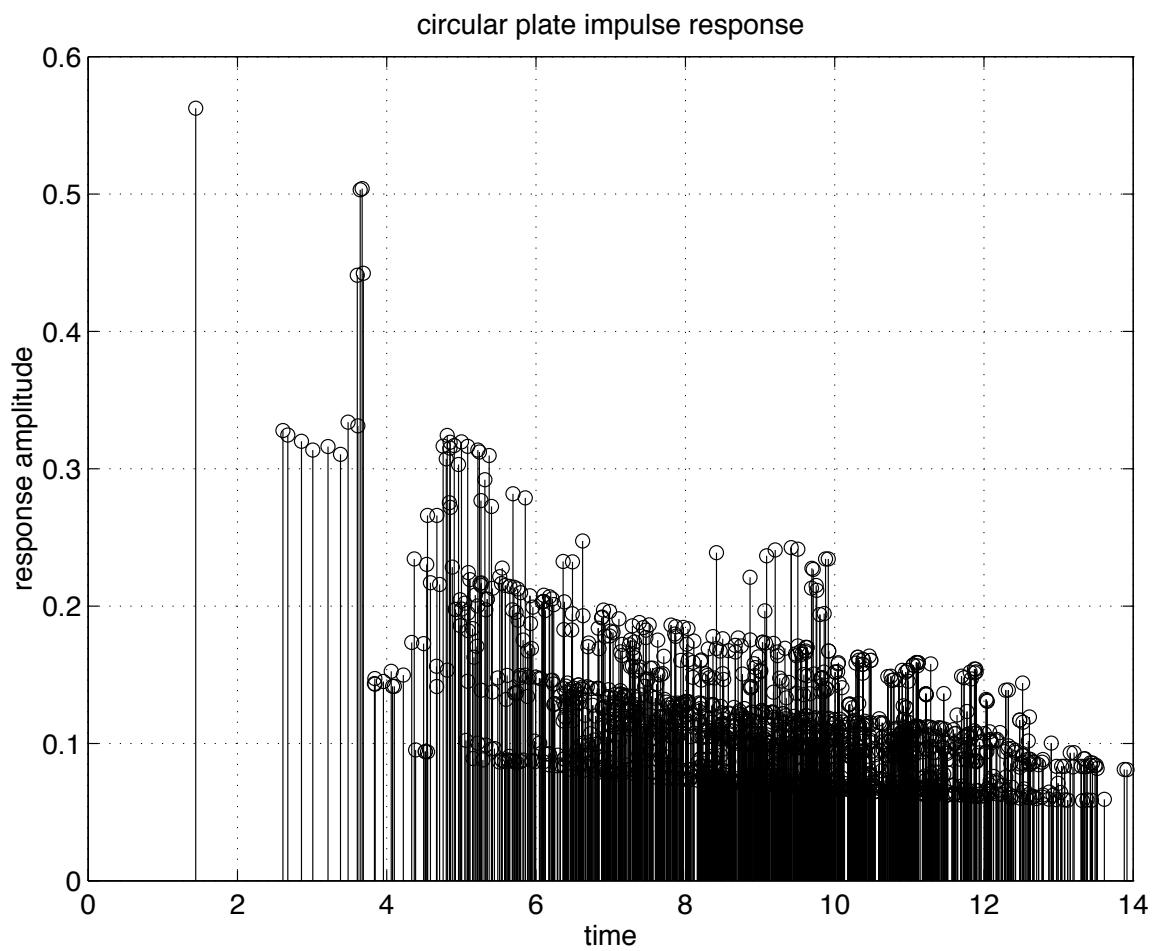






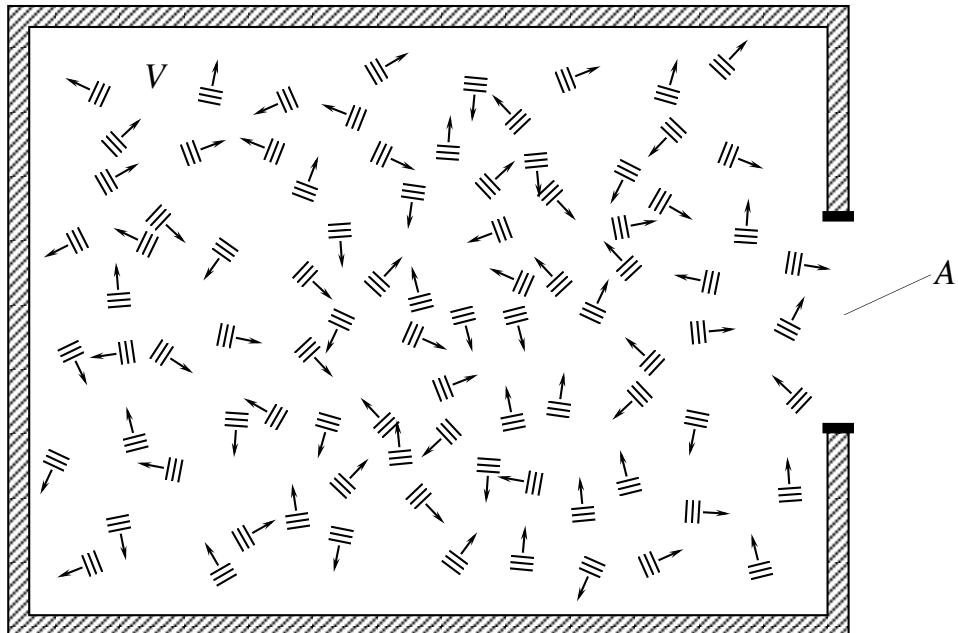


Circular Room Reflection Geometry and Impulse Response

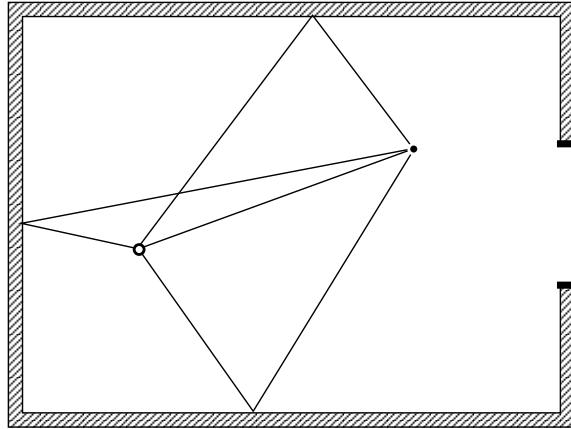


9. The Sabine Theory of Late-Field Reverberation

Energy Density Time Dependence



- Space with volume V :
 - perfectly reflecting surfaces
 - non-absorbing air
- Window or area A .
- Consider what happens after the energy is thoroughly mixed, and the energy density $w(t, x)$, initially a function of time t and position x , is only a function of time. What is $w(t)$?



Room Mixing Time

- To generate m reflections, roughly speaking, source signals must propagate long enough that a sphere with radius equal to the propagation distance contains roughly m times the room volume,

$$\frac{4}{3}\pi \cdot r_{\text{mix}}^3 = mV.$$

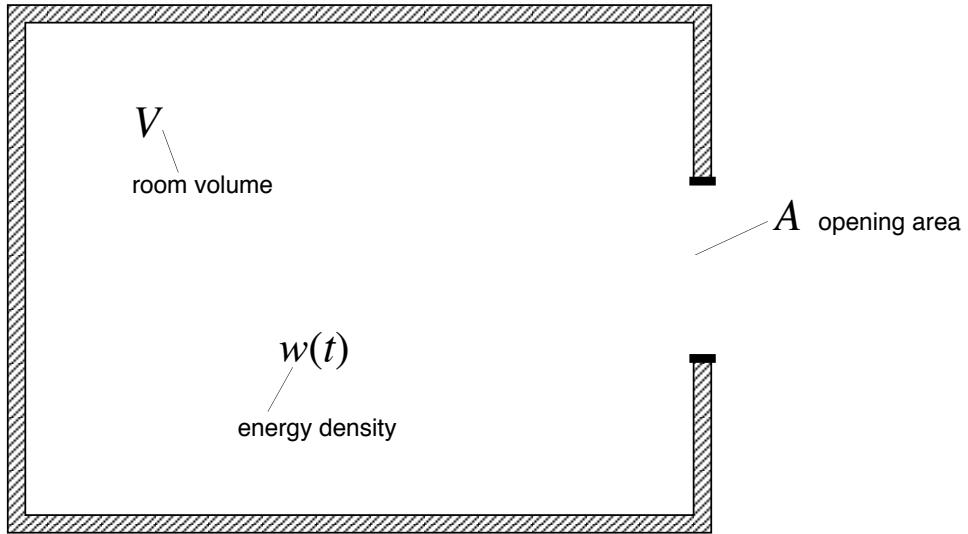
- If m reflections are needed to thoroughly mix a room, then the mixing time τ_{mix} is roughly

$$\tau_{\text{mix}} = r_{\text{mix}}/c, \quad r_{\text{mix}} = \left(\frac{3}{4\pi} \cdot mV\right)^{1/3},$$

where c is the speed of sound in air.

- For $m = 1000$, typical dimension $V^{1/3} = 10$ meters, the mixing time is roughly

$$\tau_{\text{mix}} \approx 200 \text{ milliseconds.}$$



Energy Change during Δt

- Compute the change in total energy during a small interval of time, Δt , assuming the energy density is constant throughout the volume V .
- The total energy at time $t + \Delta t$ is the energy density scaled by the volume,

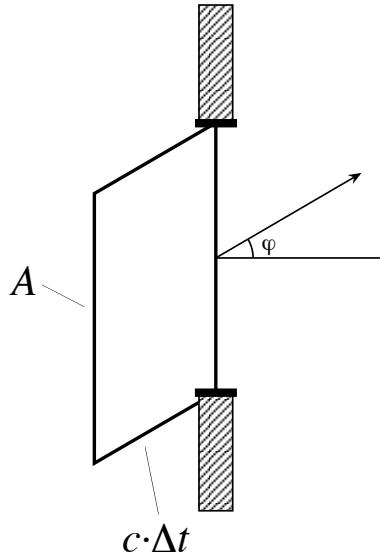
$$V \cdot w(t + \Delta t).$$

- The total energy at time t is similarly

$$V \cdot w(t).$$

- The difference is what leaves through the aperture,

$$V \cdot w(t + \Delta t) = V \cdot w(t) - \gamma(w(t), A, \Delta t).$$



Computing γ : The Two-Dimensional Case

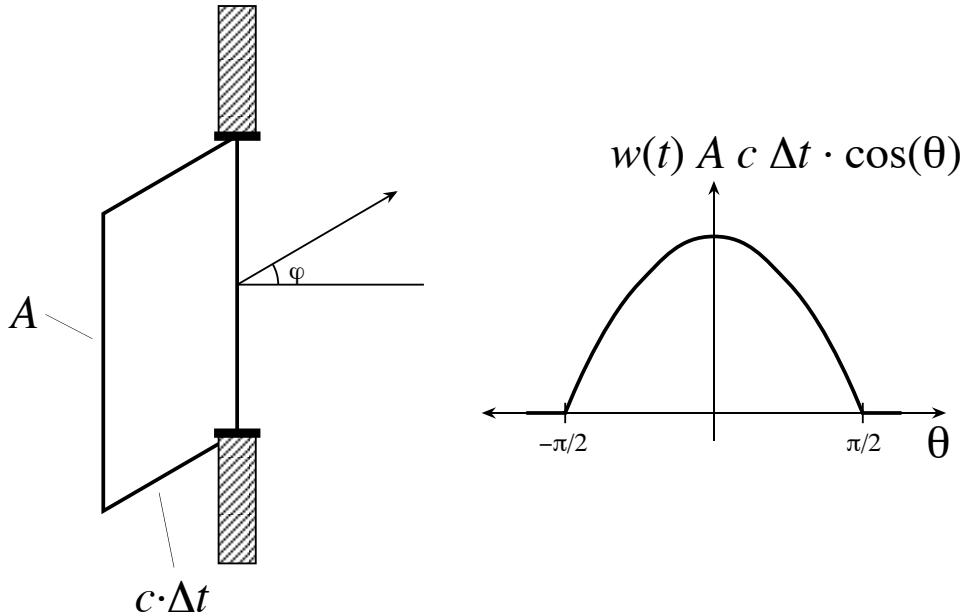
- Consider the 2D case.
- A small volume of energy a distance d from the aperture, and traveling at angle φ relative to the aperture normal, will escape during Δt if its direction of travel takes it through the aperture and

$$d \leq c \cdot \Delta t \cdot \cos \varphi,$$

where c is the speed of sound.

- The total energy leaving along φ is then

$$w(t) \cdot A \cdot c \Delta t \cos \varphi$$



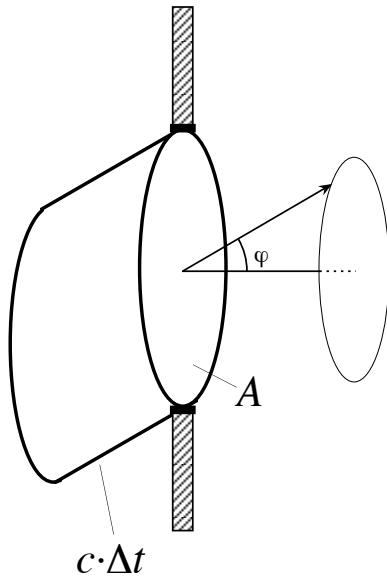
Computing γ : The Two-Dimensional Case

- Under the assumption that the room energy is thoroughly mixed, the probability that a given bit of energy is traveling along φ is

$$\wp(\varphi) = \frac{1}{2\pi}.$$

- The energy lost during Δt is

$$\begin{aligned}\gamma(w(t), A, \Delta t) &= \int_{-\pi/2}^{\pi/2} w(t) \cdot A \cdot c\Delta t \cos \varphi \cdot \wp(\varphi) d\varphi \\ &= \frac{1}{\pi} w(t) \cdot A \cdot c\Delta t.\end{aligned}$$



Computing γ : Circular Aperture

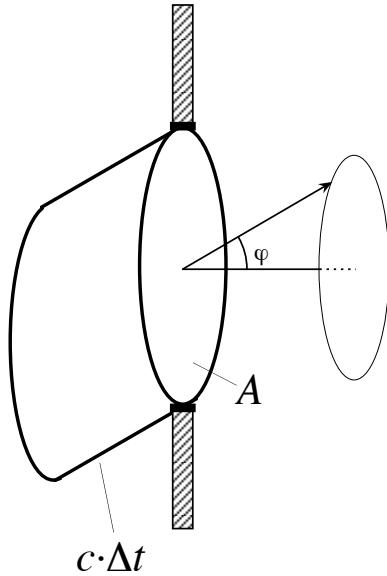
- The energy lost during Δt traveling along direction (φ, θ) is

$$w(t) \cdot A \cdot c\Delta t \cos \varphi.$$

- As a result, the total energy lost during Δt is

$$\begin{aligned} \gamma(w(t), A, \Delta t) \\ = \int_0^{\pi/2} \int_0^{2\pi} w(t) \cdot A \cdot c\Delta t \cos \varphi \cdot \wp(\varphi, \theta) d\theta d\varphi, \end{aligned}$$

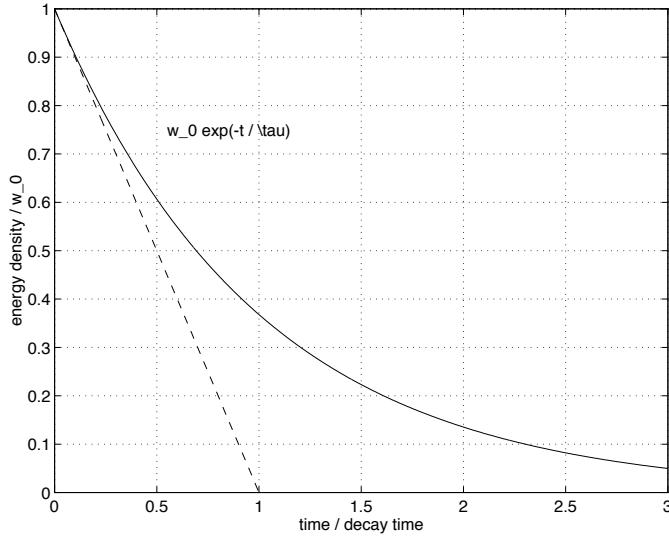
where $\wp(\varphi, \theta)$ is the probability that any bit of energy is traveling along direction (φ, θ) .



Computing γ : Circular Aperture

- Assuming all directions on the sphere are equally likely, $\wp(\varphi, \theta) = \frac{\sin \varphi}{4\pi}$.
- The energy lost during Δt is therefore

$$\begin{aligned}
 \gamma(w(t), A, \Delta t) &= \frac{w(t) \cdot A \cdot c\Delta t}{4\pi} \int_0^{\pi/2} \int_0^{2\pi} \cos \varphi \cdot \sin \varphi d\theta d\varphi, \\
 &= \frac{w(t) \cdot A \cdot c\Delta t}{2} \int_0^{\pi/2} \cos \varphi \cdot \sin \varphi d\varphi, \\
 &= \frac{w(t) \cdot A \cdot c\Delta t}{4}.
 \end{aligned}$$



Energy Density Behavior

- By conservation of energy,

$$V w(t + \Delta t) = V w(t) - g c A w(t) \Delta t,$$

where g is a geometric factor.

- Rearranging terms,

$$\frac{w(t + \Delta t) - w(t)}{\Delta t} = -\frac{g c A}{V} w(t),$$

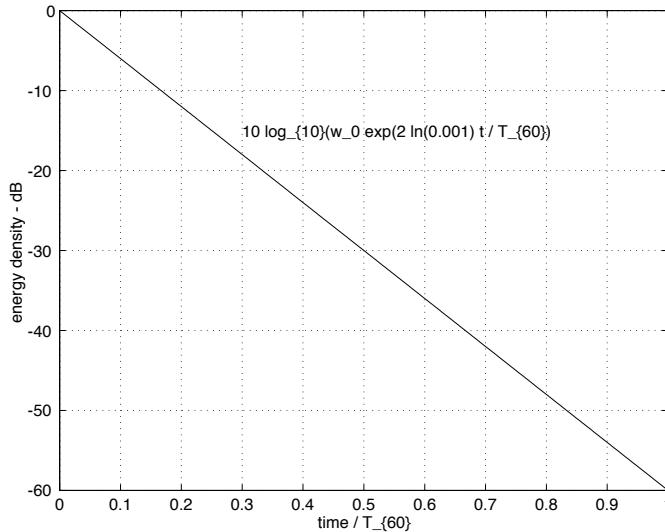
and taking the limit $\Delta t \rightarrow 0$, we have

$$\frac{dw}{dt} = -\frac{1}{\tau} w(t), \quad \tau = \frac{V}{g c A},$$

where τ is the so-called *characteristic decay time*.

- For a well mixed room with $w(t = 0) = w_0$,

$$w(t) = w_0 e^{-t/\tau}.$$



Energy Density Behavior

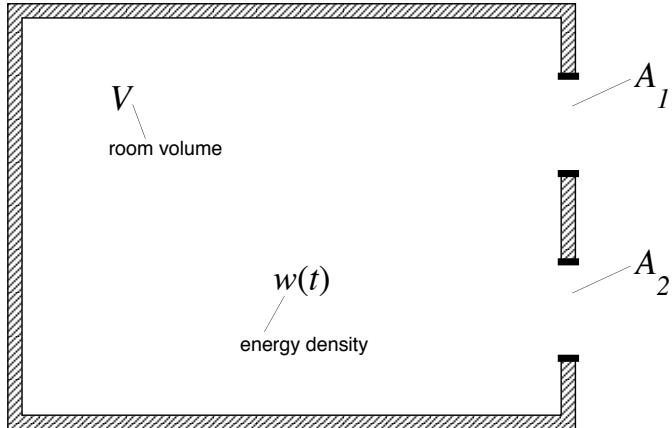
- The energy density follows an exponential decay to zero,
- $$w(t) = w_0 e^{-t/\tau}, \quad \tau = \frac{V}{g c A}.$$
- The time constant is proportional to room volume, and inversely proportional to window area.
 - The *reverberation time* or T_{60} is the time it takes the energy density to decay to 60 dB below its initial value,

$$w(t) = w_0 e^{-t/\tau} = w_0 e^{2 \ln(0.001) t / T_{60}}.$$

We have

$$T_{60} = -2 \ln(0.001) \frac{V}{g c A} \approx (0.161 \text{ s/m}) \frac{V}{A}.$$

Extensions: Absorbing Materials and Air



Multiple Windows

- In the presence of multiple noninteracting windows, the energy losses add,

$$V w(t + \Delta t) = V w(t) - \gamma(w(t), \Delta t, A_1, \dots, A_N),$$

and

$$\gamma = \left(\sum_i A_i \right) g c \cdot w(t).$$

- The reverberation time is accordingly reduced,

$$T_{60}(A_i, i = 1, \dots, N) = -2 \ln(0.001) \frac{V}{g c \sum_i A_i}.$$

Materials Patches vs. Windows

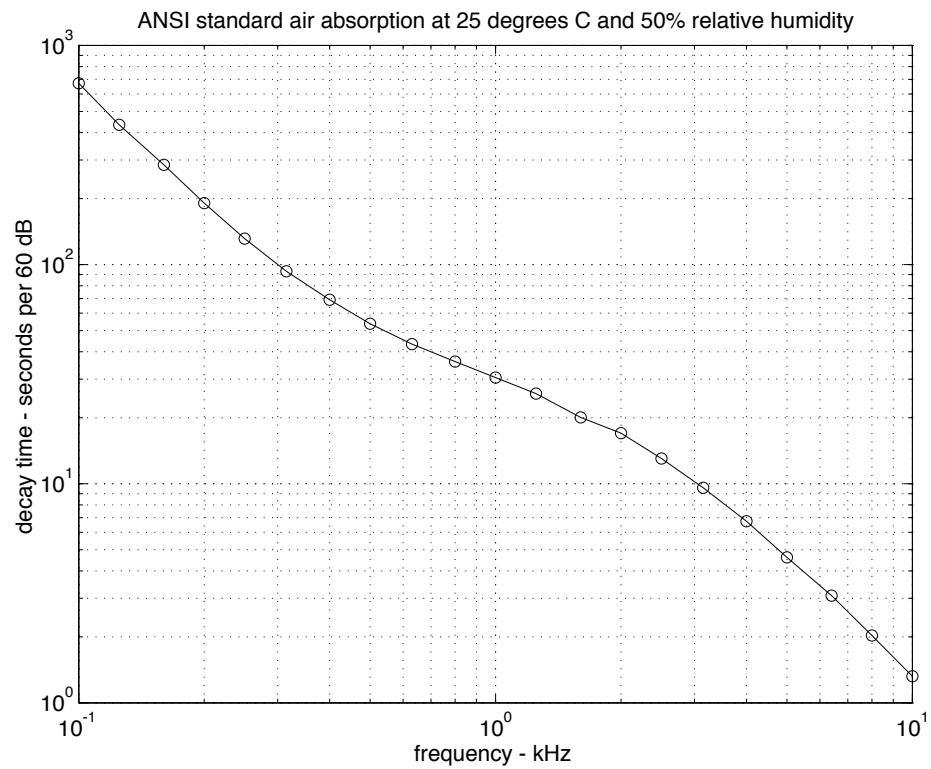
material	Absorption Coefficients, S					
	125	250	500	1000	2000	4000
marble	0.01	0.01	0.01	0.01	0.02	0.02
brick	0.03	0.03	0.03	0.04	0.05	0.07
concrete block	0.36	0.44	0.31	0.24	0.39	0.25
plywood	0.28	0.22	0.17	0.22	0.10	0.11
cork	0.14	0.25	0.40	0.25	0.34	0.21
glass window	0.35	0.25	0.18	0.12	0.07	0.04
drapery	0.10	0.25	0.46	0.60	0.56	0.52
carpet	0.02	0.06	0.14	0.37	0.66	0.65
hardwood	0.15	0.11	0.10	0.07	0.06	0.07
grass	0.11	0.26	0.60	0.69	0.92	0.94

- Room surfaces typically absorb a portion of impinging acoustic energy. The portion depends on the surface material and is a function of frequency.
- To compute the reverberation time associated with a patch of material, its area a is scaled by its absorbing power S ,

$$A_{\text{effective}} = a \cdot S.$$

- Note that $S = 1$ for an open window, and $S = 0$ for a perfectly reflecting wall.
- The reverberation time is then

$$T_{60} = -2 \ln(0.001) \frac{V}{gc \sum_i a_i S_i}.$$



Air Absorption

- Air absorbs sound at a frequency-dependent rate proportional to the distance traveled.
- During a time interval Δt , every unit volume of energy density experiences the same attenuation,

$$\exp\{-\alpha(\omega)\Delta t\}.$$

Air Absorption

- In a room of volume V with uniform energy density $w(t)$, the total energy absorbed via air absorption during the time interval Δt is

$$\begin{aligned}\gamma_{\text{air}} &= V w(t) \cdot [1 - \exp\{-\alpha(\omega)\Delta t\}], \\ &\approx V w(t) \cdot \alpha(\omega)\Delta t,\end{aligned}$$

assuming $\Delta t \ll 1$.

- Again, by conservation of energy,

$$V w(t + \Delta t) = V w(t) - \gamma_{\text{surfaces}} - \gamma_{\text{air}},$$

and, substituting for the absorbed energy terms,

$$V w(t + \Delta t) = V w(t) - \left(g c \sum_i A_i + V \alpha(\omega) \right) w(t) \Delta t.$$

- Rearranging, and taking the interval Δt to zero,

$$\frac{dw}{dt} = - \left(\frac{g c \sum_i A_i}{V} + \alpha(\omega) \right) w(t).$$

- Again, we have an exponential decay, this time with characteristic decay time

$$\tau = \frac{V}{g c \sum_i A_i + V \alpha(\omega)}.$$

Reverberation Time as a Function of Room Size

- Consider a room with volume V , absorbing materials with effective areas A_i , and characteristic dimension $\ell = V^{1/3}$.
- The reverberation time is

$$\begin{aligned} T_{60}(\omega) &= -2 \ln(0.001) \frac{1}{gc \sum_i A_i/V + \alpha(\omega)}, \\ &= -2 \ln(0.001) \frac{1}{gc\sigma(\omega)/\ell + \alpha(\omega)}, \end{aligned}$$

where the $\sigma(\omega)$ is the sum of effective areas, normalized by ℓ^2 .

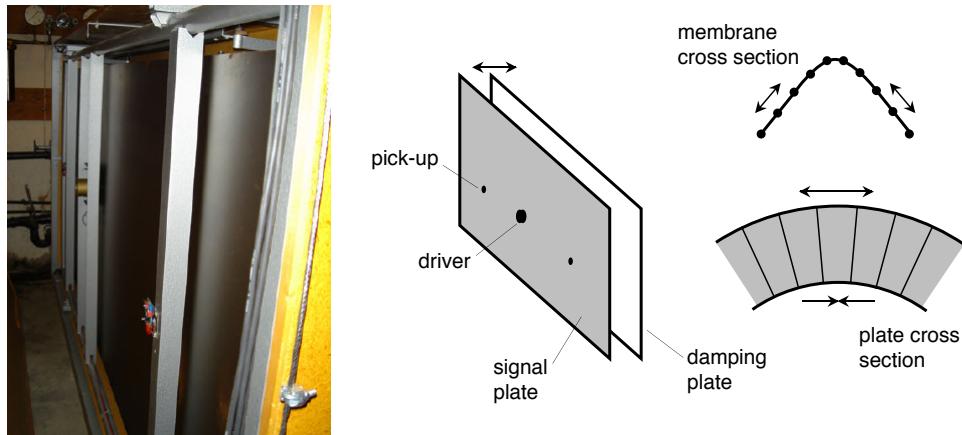
- Note that the reverberation time is the harmonic mean of the materials and air reverberation times,

$$T_{60}(\omega) = \left[\frac{1}{T_{60\text{surfaces}}(\omega)} + \frac{1}{T_{60\text{air}}(\omega)} \right]^{-1}.$$

This means that in any frequency band, the shorter reverberation times dominate. Also, air absorption provides an upper limit on reverberation time as a function of frequency.

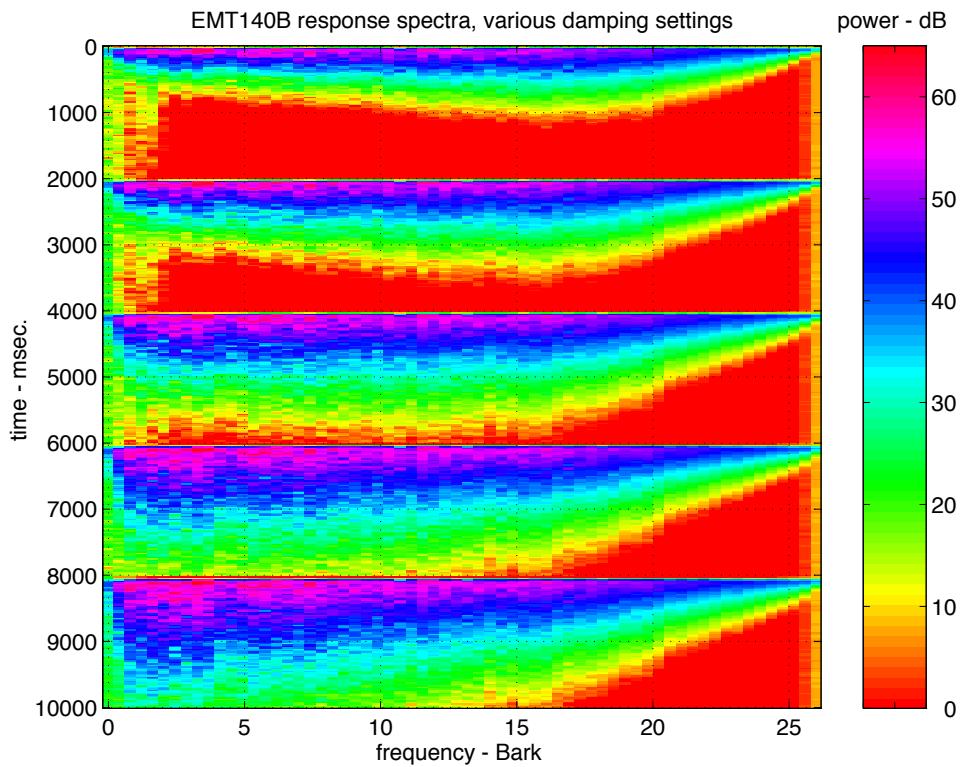
- As the room gets larger, the contribution of the air becomes more important: Larger rooms will have longer reverberation times with darker tails.

Frequency-Dependent T_{60} Example: Plate Reverberator

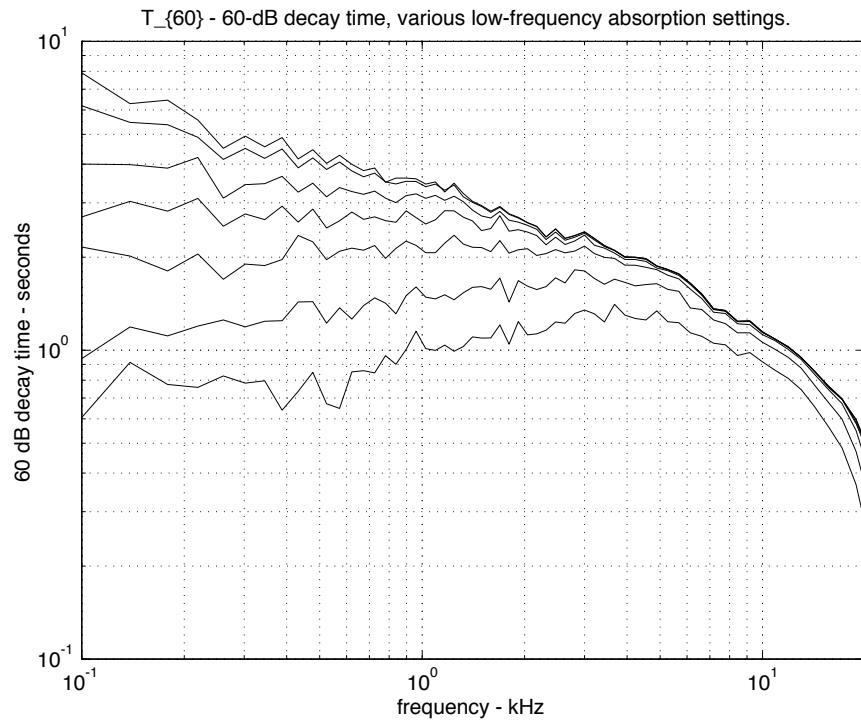


- Tensioned plates support wave propagation.
- Along a transverse wave, restoring forces are provided by extension of the plate on one side and compression of the plate on the other side.
- These forces are different than those at play in a membrane, in which the material is sufficiently thin that the restoring force results from material stretching. As a result wave propagation on a plate is dispersive, with high frequencies traveling faster than low frequencies.
- Thermoviscous losses and edge reflections dominate the low-frequency T_{60} , whereas radiation losses contribute mainly to the high-frequency decay rate.
- In the EMT140 plate reverberator, a damping plate is positioned near the signal plate and provides additional low-frequency losses.

EMT140 Plate Reverberator



- EMT140 plate reverberator spectrograms at various damping settings.

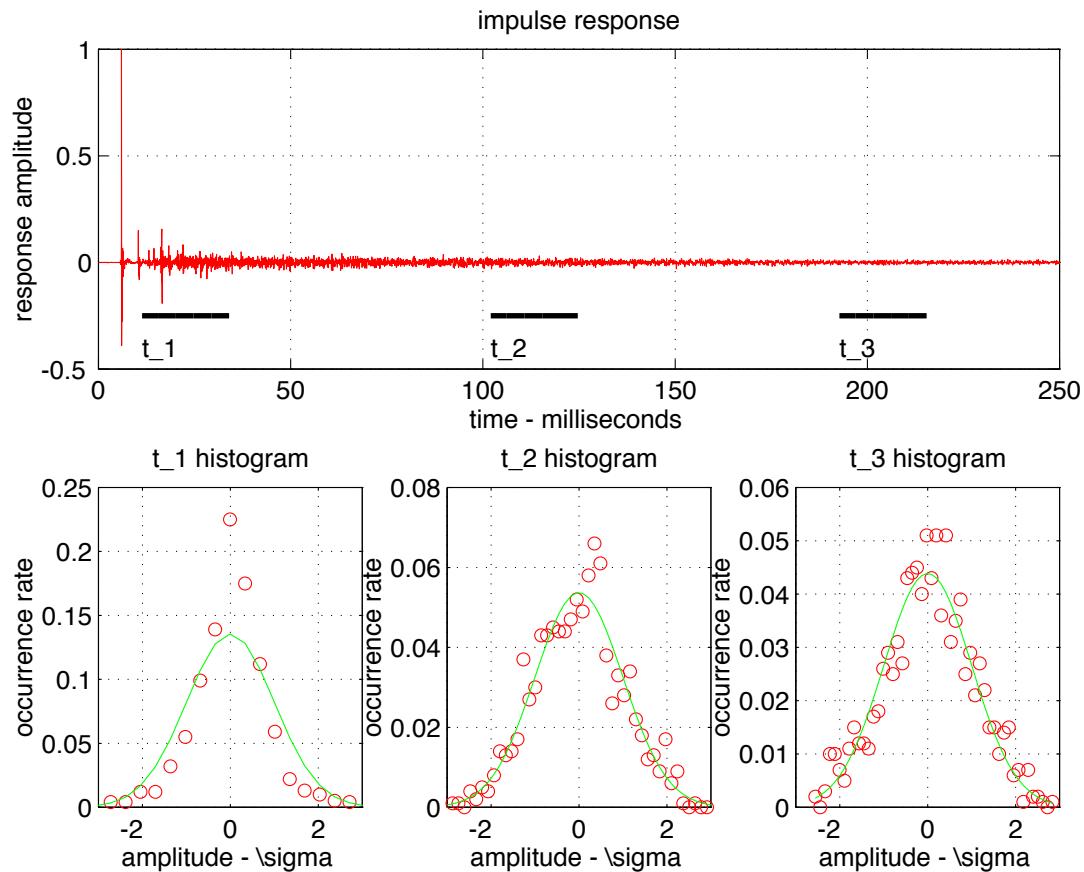


Frequency-Dependent T_{60} Example

- EMT140 plate reverberator T_{60} at various damping settings.
- Sound example: EMT140 impulse response set.

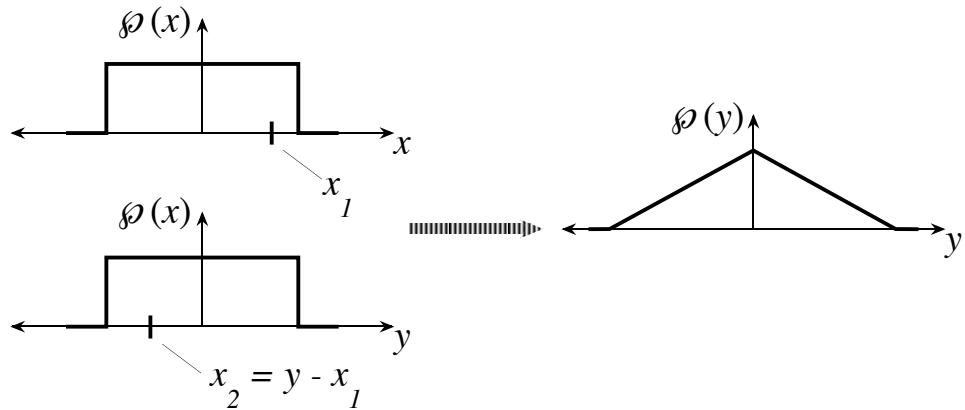
10. Reverberation Statistics and Impulse Response Model

Statistical Nature of the Late Field



- It appears that any segment of the *late field* is approximately colored Gaussian noise.

The Central Limit Theorem.



- The sum of a sufficient number of i.i.d. random variables is approximately Gaussian.
- Rationale:
 - Consider the sum of two independent random variables x_1 and x_2 , both drawn from $\varphi(x)$.
 - The probability that the sum takes on a value $y = x_1 + x_2$ is the probability that x_1 takes on a particular value times the probability that x_2 is

$$x_2 = y - x_1,$$

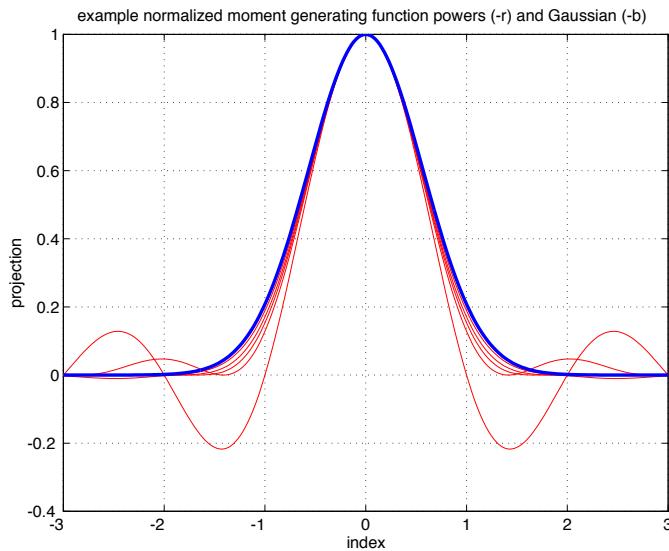
integrated over all possible values x_1 .

$$\varphi(y = x_1 + x_2) = \int \varphi(x_1) \cdot \varphi(y - x_1) \cdot dx_1.$$

- Note that this is the convolution of the probability densities for x_1 and x_2 ,

$$\varphi(y = x_1 + x_2) = \varphi(x_1) * \varphi(x_2).$$

The Central Limit Theorem.



- The *moment generating function* $\mu(t)$ of a probability density $\varphi(x)$ is the Fourier Transform of the probability density,

$$\mu(t) = \mathcal{F}\{\varphi(x)\}.$$

- Therefore, the sum of n such i.i.d. random variables has as its moment generating function $\mu^n(t)$.
- Because the probability density is real, positive and integrates to one, the moment generating function is symmetric, and achieves a maximum of one at $t = 0$.
- In a small neighborhood about $t = 0$, then

$$\mu(t) \approx (1 - \nu^2 t^2).$$

- Raising $\mu(t)$ to the n th power, for large n we see that it—and therefore its generating density—approximates a Gaussian,

$$\mu(t)^n \approx (1 - (n\nu^2 t^2)/n)^n \approx \exp\{-n\nu^2 t^2\}.$$

Late-Field Reverberation Impulse Response Model

- The late-field spectral envelope is an equalized exponential decay,

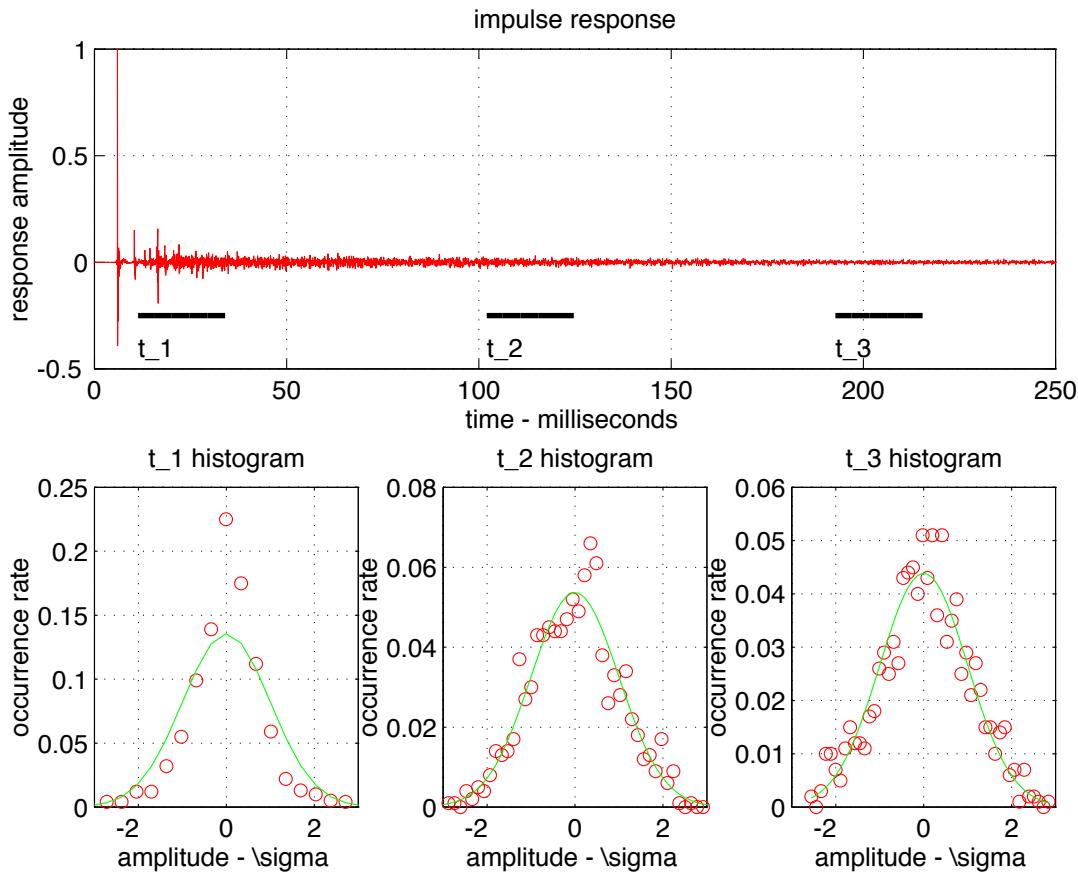
$$P(\omega, t) = |q(\omega)|^2 \exp\{-t/\tau(\omega)\}.$$

- The late field is Gaussian noise.
- Therefore, late-field reverberation is well approximated by white Gaussian noise $n(t)$ convolved with a time-varying filter,

$$h(t) = [q(\omega) \cdot \exp\{-t/\tau(\omega)\}] * n(t).$$

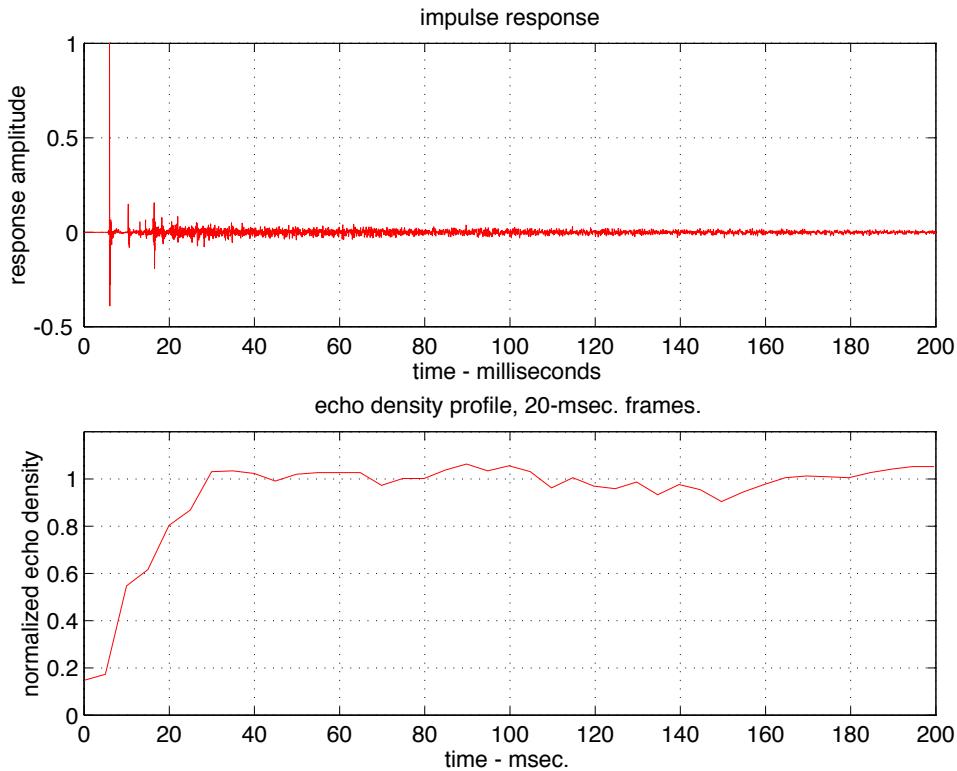
11. Reverberation Impulse Response Analysis

Echo Density Analysis



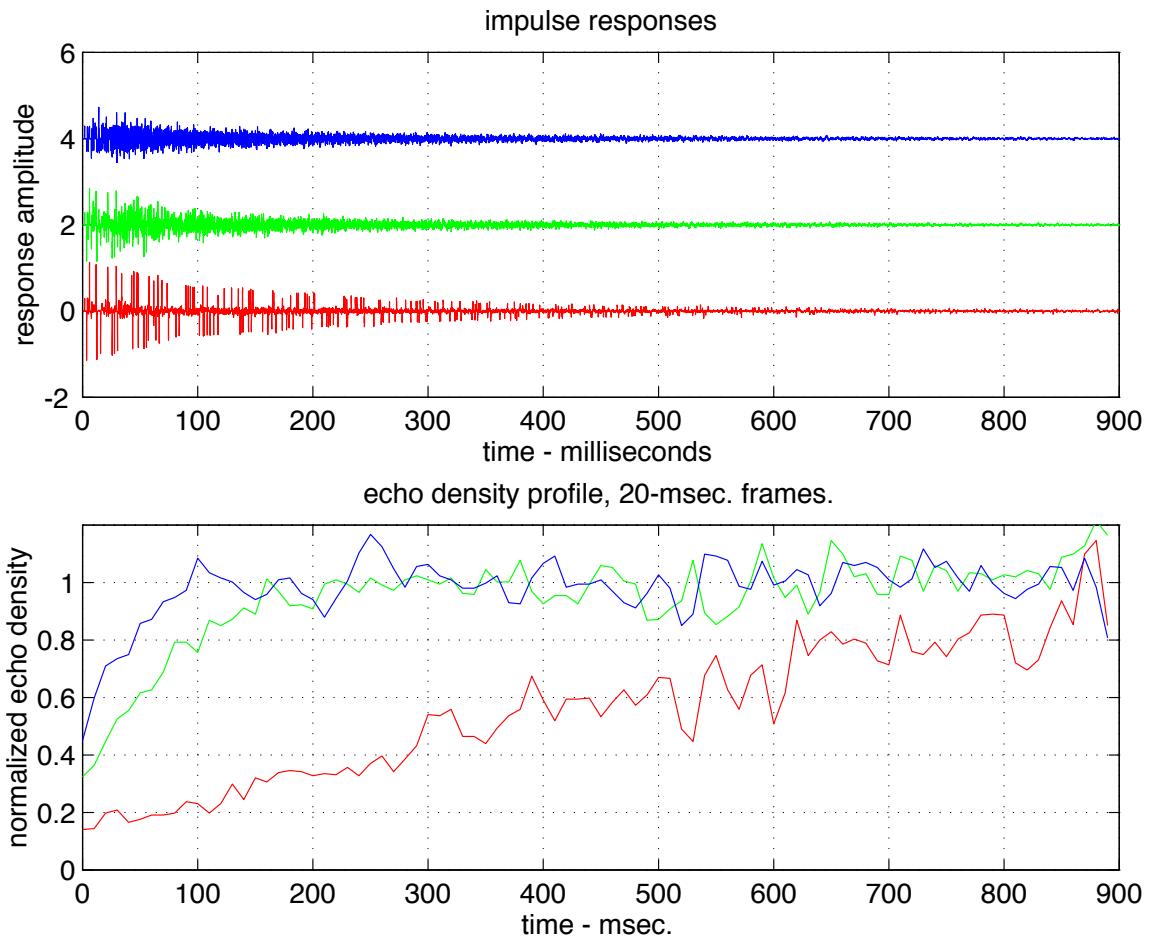
- Once the room is fully mixed, any room impulse response segment is well approximated by colored Gaussian noise.
- Note that the segment containing well separated reflections is kurtotic (more peaked about the center than a Gaussian of the same standard deviation).

Echo Density Profile



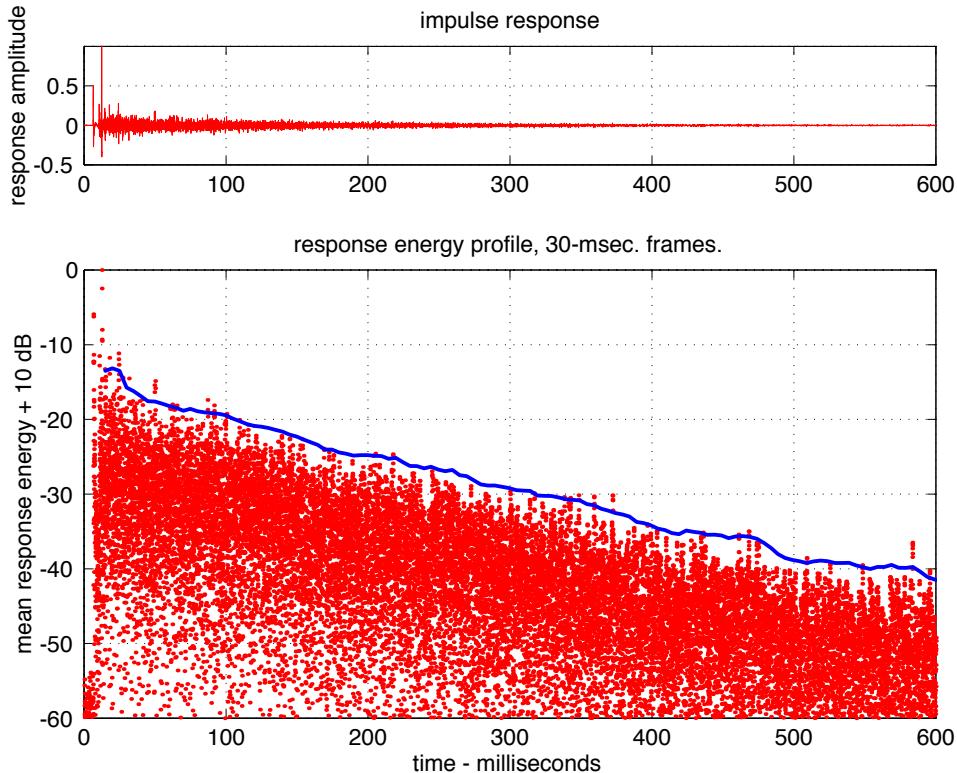
- A simple measure of impulse response echo density as a function of time is the percentage of samples in a frame which are outside a standard deviation from the mean.
- In the late field, this measure is expected to be that of a Gaussian, about 30%.
- In the presence of early reflections, the measure is expected to be lower, as the energy (and therefore standard deviation) is concentrated in the reflections.
- Note that this measure is insensitive to the segment spectrum—it simply looks at temporal concentrations of impulse response energy.

Echo Density Profile: Example



- Impulse responses and corresponding echo density profiles are shown for an artificial reverberator at several “diffusion” control settings.
- Sound example: artificial reverberator impulse responses, different diffusion settings.

Energy Envelope Analysis



Impulse Response Energy Envelope

- Reverberation impulse responses are nonstationary noise processes with exponentially decaying variances.
- How do you estimate the variance of a noise process?
 - Average sample variances: For an averaging width β , the *smoothed energy envelope* of an impulse response $h(t)$ is

$$P(t; \beta) = \frac{1}{\beta} \sum_{n=t-\beta/2}^{t+\beta/2} h^2(n).$$

Impulse Response Energy Envelope

- In computing the energy envelope,

$$P(t; \beta) = \frac{1}{\beta} \sum_{n=t-\beta/2}^{t+\beta/2} h^2(n),$$

it is important to pick the averaging window to be wide enough to suppress noise, but not so wide that the estimate is biased.

- For a late-field decay,

$$\mathbb{E}\{h^2(t)\} = w_0 \exp(-t/\tau).$$

- In a neighborhood of t_0 , $t \in [t_0 - \beta/2, t_0 + \beta/2]$ we have

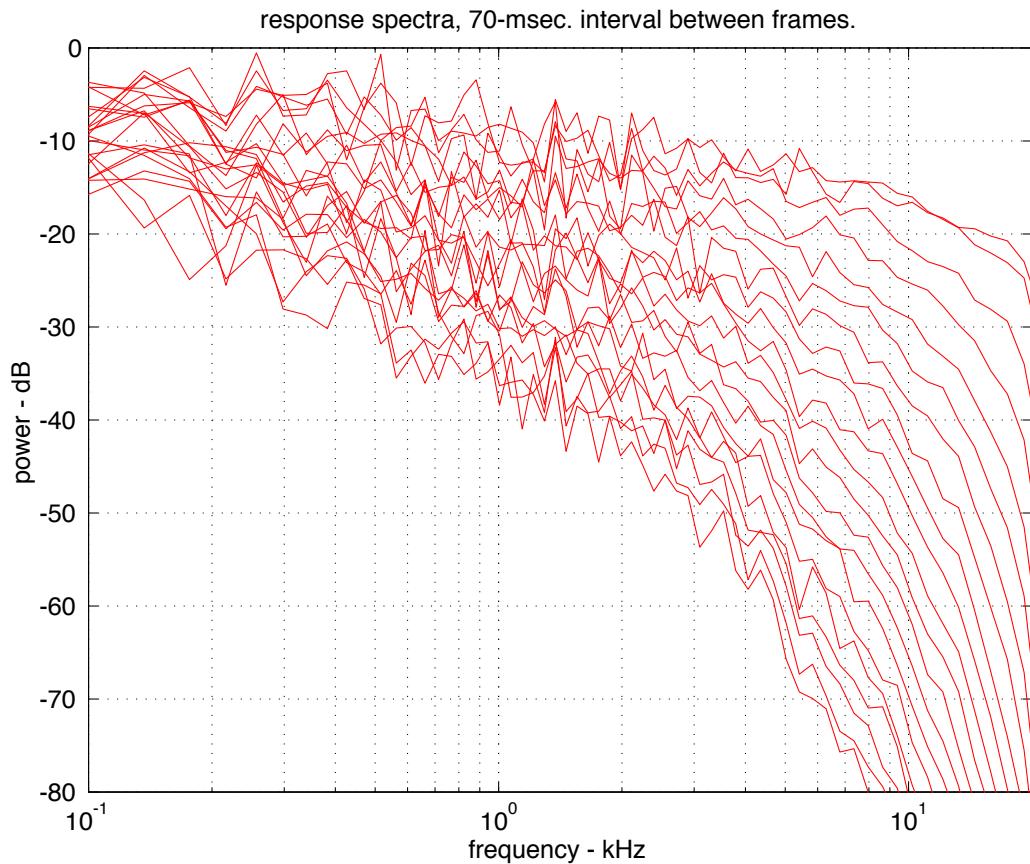
$$\exp(-t/\tau) \approx \exp(-t_0/\tau) \cdot (1 - (t - t_0)/\tau).$$

provided $\beta \ll \tau$.

- Accordingly,

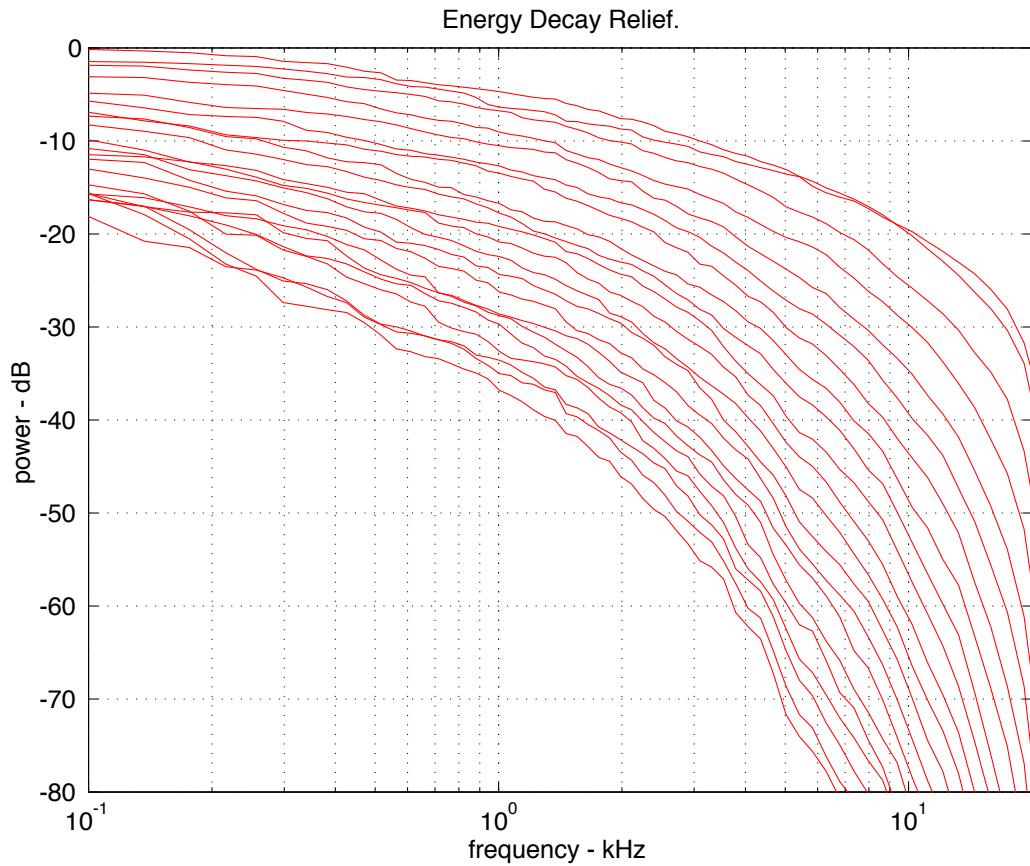
$$\mathbb{E}\{P(t; \beta)\} = w_0 \exp(-t/\tau)$$

when $\beta \ll \tau$.



Impulse Response Time-Frequency Analysis

- The short-time Fourier transform $P(\omega, t)$ —the Fourier transform of successive impulse response segments—provides an estimate of average power as a function of frequency.



Impulse Response Time-Frequency Analysis

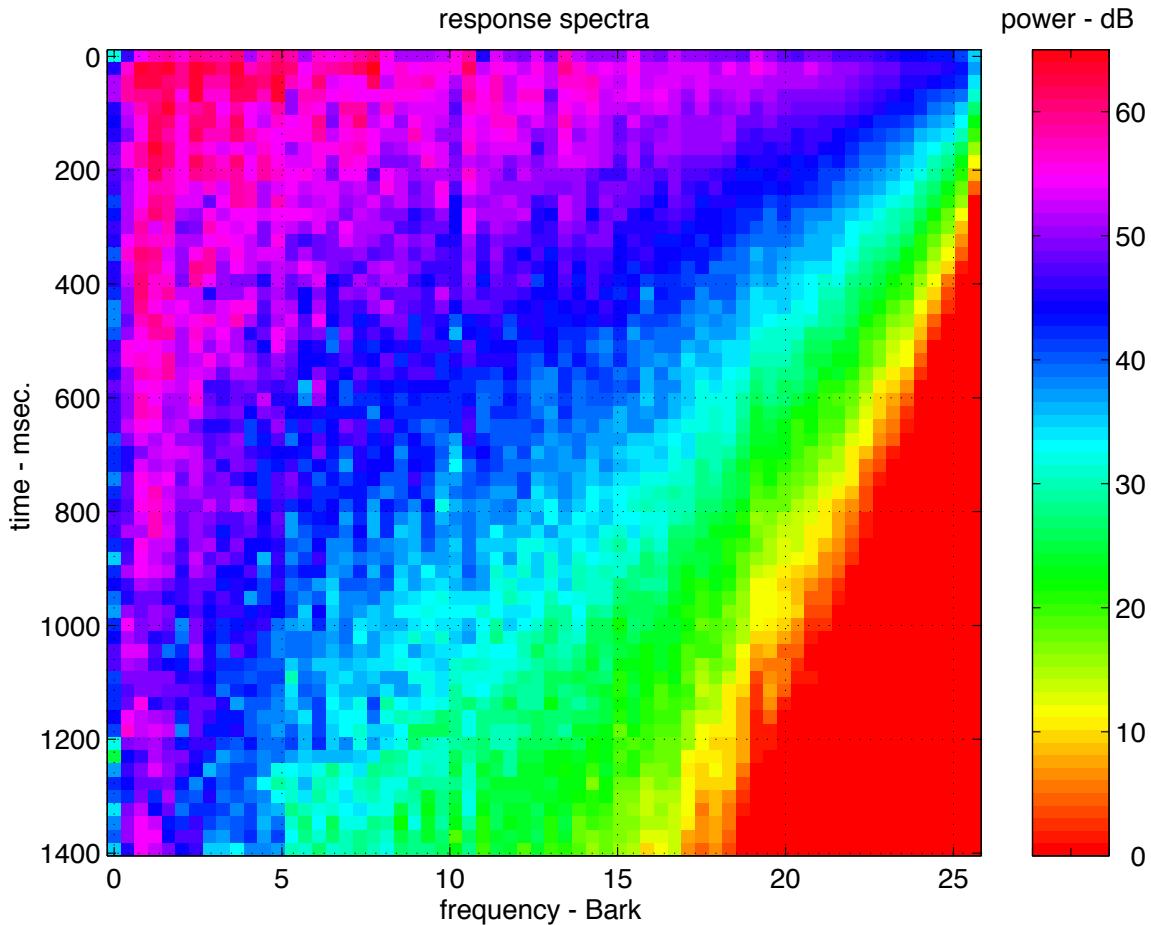
- The Energy Decay Relief (EDR) is the integral of the spectrogram $P(\omega, \tau)$,

$$\text{EDR}(\omega, t) = \int_t^\infty P(\omega, \tau) d\tau,$$

and provides a smooth alternative to the STFT.

- Note that if the power spectrum $P(\omega, \tau)$ is an exponential, its integral will have the same decay rates.

Impulse Response Spectrogram Model

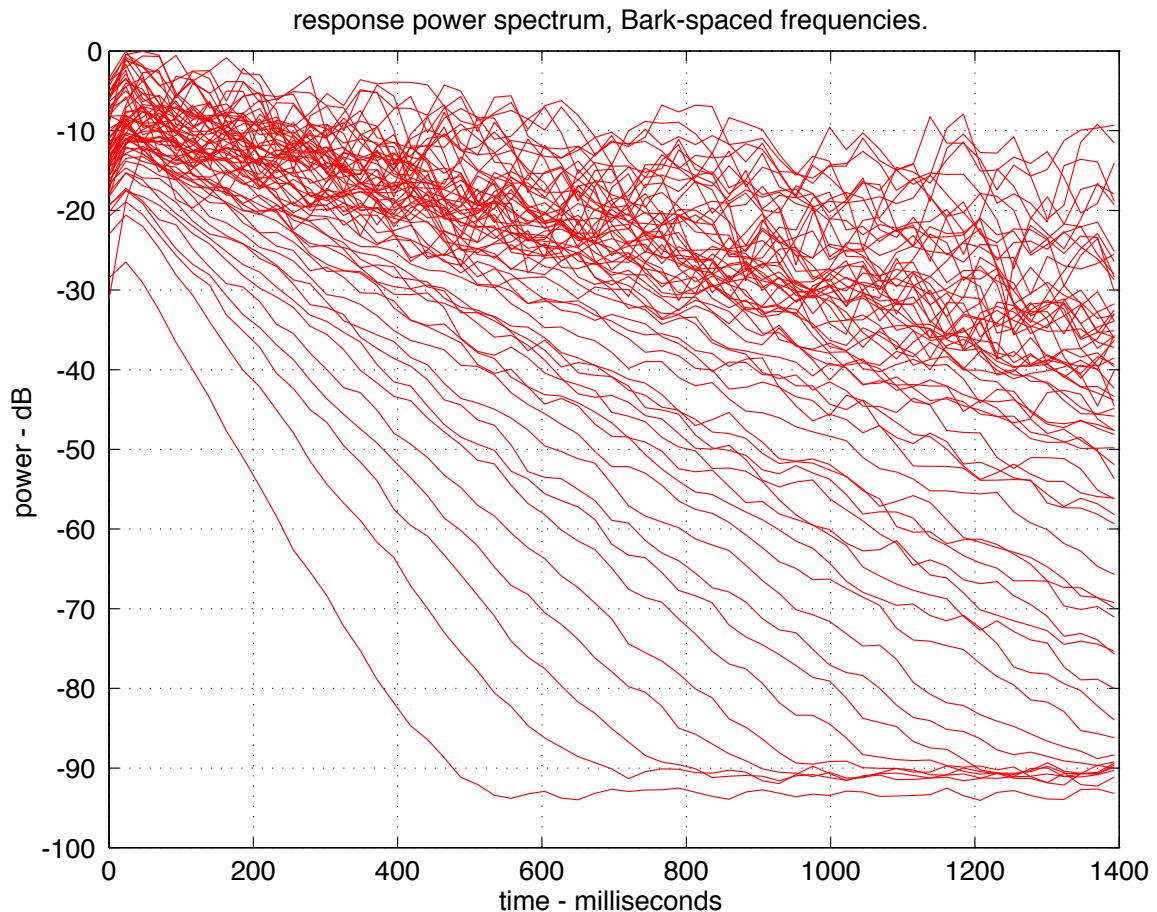


- The measured late-field spectrogram $\tilde{P}(\omega, t)$ is modeled as a nonstationary noise process with mean

$$\mathbb{E}\{\tilde{P}(\omega, t)\} = |q(\omega)|^2 \exp(-2t/\tau(\omega)),$$

where $|q(\omega)|$ is the magnitude of the initial late-field equalization and $\tau(\omega)$ is the late-field decay rate.

Equalization and Reverberation Time Estimation



- The unknown equalization $|q(\omega)|$ and decay rate $\tau(\omega)$ may be separately estimated at each frequency, and form sufficient statistics if estimated accurately.

Equalization and Reverberation Time Estimation

- Denote by $\boldsymbol{\theta}$ the unknown log equalization and decay rate,

$$\boldsymbol{\theta} = \begin{bmatrix} \ln |q(\omega)| \\ 1/\tau(\omega) \end{bmatrix},$$

by $\boldsymbol{\eta}_{\boldsymbol{\theta}}$ the stack of hypothesized band powers at times t_0, \dots, t_N for a given log equalization and decay rate $\boldsymbol{\theta}$,

$$\boldsymbol{\eta}_{\boldsymbol{\theta}} = \begin{bmatrix} \ln |q(\omega)| - t_0/\tau(\omega) \\ \vdots \\ \ln |q(\omega)| - t_N/\tau(\omega) \end{bmatrix},$$

and by $\tilde{\boldsymbol{\eta}}$ the similar stack of measured log spectra.

- We then have

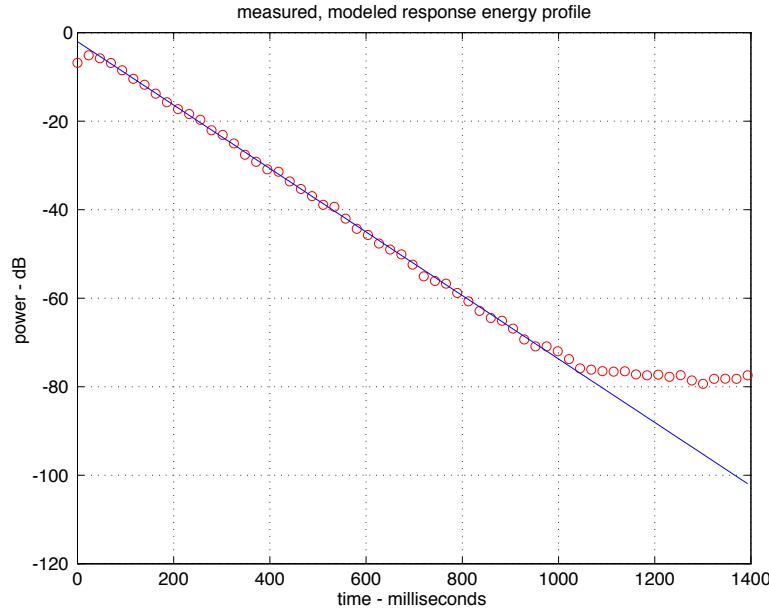
$$\tilde{\boldsymbol{\eta}}(\omega, t_i) = \boldsymbol{\eta}_{\boldsymbol{\theta}}(\omega, t_i) + \epsilon(\omega, t_i),$$

where $\epsilon(\omega, t_i)$ is an *equation error* representing the discrepancy between the measured and hypothesized log spectra.

- The unknown initial equalization and decay rate $\boldsymbol{\theta}$ are estimated as those minimizing the sum of square equation errors,

$$\hat{\boldsymbol{\theta}} = \operatorname{argmin}_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) = [\boldsymbol{\eta}_{\boldsymbol{\theta}} - \tilde{\boldsymbol{\eta}}]^{\top} [\boldsymbol{\eta}_{\boldsymbol{\theta}} - \tilde{\boldsymbol{\eta}}].$$

Equalization and Reverberation Time Estimation



- Note that the equation error is linear in the unknown log equalization and decay rate θ , making its sum of squares easily minimized.
- Denoting by B the basis

$$B = [1 \ -t],$$

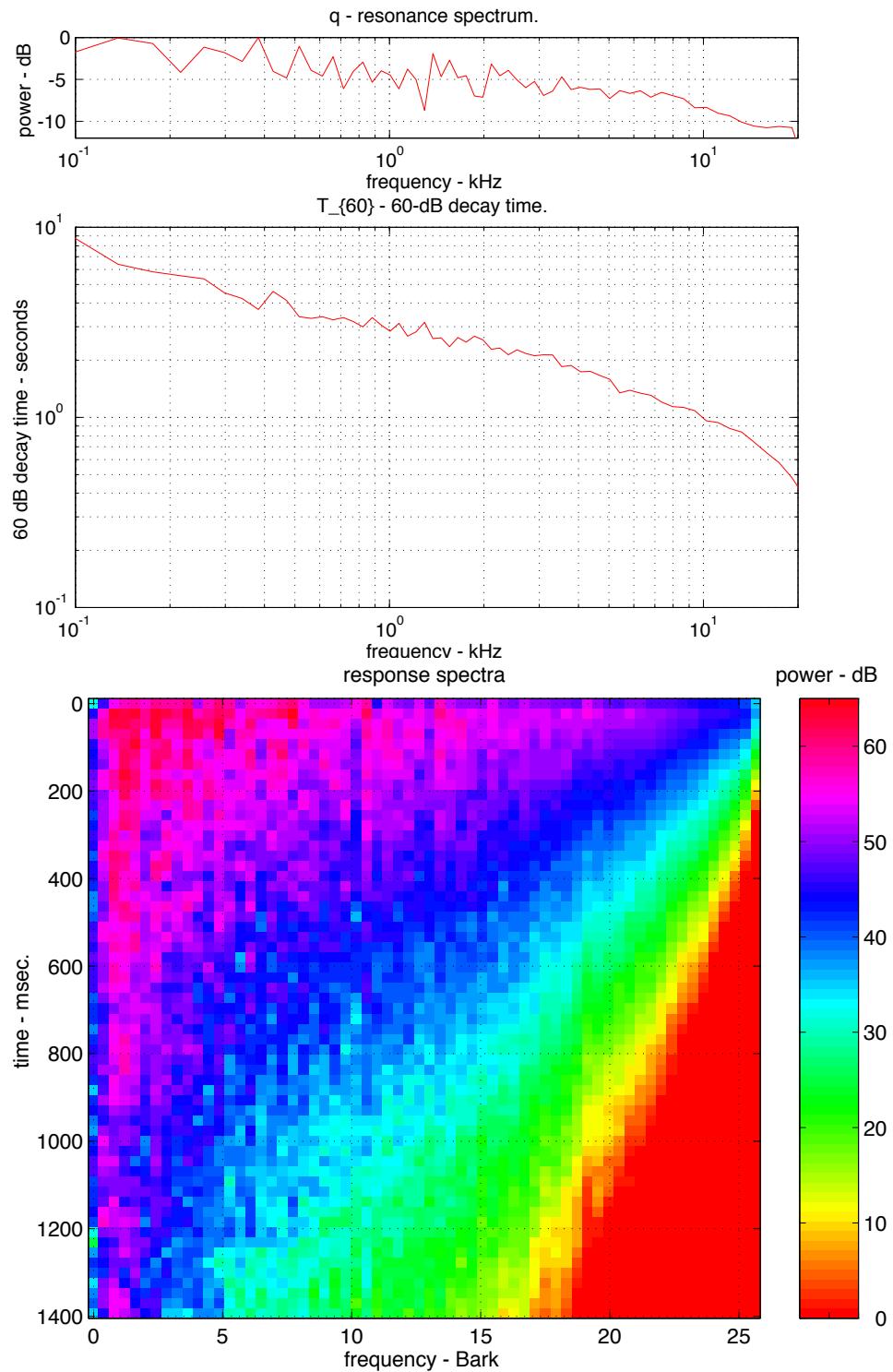
the log equalization and decay rate estimate is

$$\hat{\theta} = (B^\top B)^{-1} B^\top \tilde{\eta},$$

with the estimated dB energy as a function of time being the projection onto B of the measured dB band power as a function of time,

$$\hat{\eta} = B\hat{\theta} = B(B^\top B)^{-1} B^\top \tilde{\eta}.$$

Equalization and Reverberation Time Estimation Example



12. Reverberation Psychoacoustics

Psychoacoustic Parameters

- Source parameters.
 - direction, distance
 - brilliance, warmth
- Room parameters.
 - clarity
 - intimacy, remoteness
 - running and late reverberance
 - heaviness, liveliness
 - spaciousness, envelopment

Psychoacoustic Parameters

- *Direction* and *distance* describe the source position, and are related to the direct path direction and the relative direct path and reflected energy levels.
- Source *brilliance* and source *warmth* are determined by the early energy equalization.
- *Intimacy* describes the remoteness or intimacy of the space. It is determined by the delay between the arrival of the direct path and reflected energy; the shorter the delay the more intimate the room.
- Whether a room appears muddy or clear is indicated by *clarity*, the ratio of early energy to late energy: the greater the portion of energy arriving within 50 to 80 milliseconds of the direct path, the clearer the source.
- Room *reverberance* describes whether the room is wet or dry, live or dead—the sense that the room prolongs source sounds. Reverberance has two aspects: *running reverberance* is heard during continuous sounds, whereas *late reverberance* is heard during breaks in the source signal. Running reverberance is specified by the early decay time (EDT), six times the time taken for source signals to decay to 10 dB below their initial level. Late reverberance is expressed as T_{60} , the time needed for source signals to decay 60 dB below their initial level.
- Room *heaviness* and *liveliness* are specified by the variation in T_{60} as a function of frequency, livelier rooms having a longer high-frequency T_{60} .
- *Spaciousness* and *envelopment* depend on the percentage of early energy arriving laterally and the crosscorrelation of late-field energy heard at the left and right ears of a listener. Large lateral energy portions and small crosscorrelations correspond to spacious, enveloping environments.

13. Impulse Response Synthesis and Convolutional Reverberation

Reverberation Impulse Response Synthesis

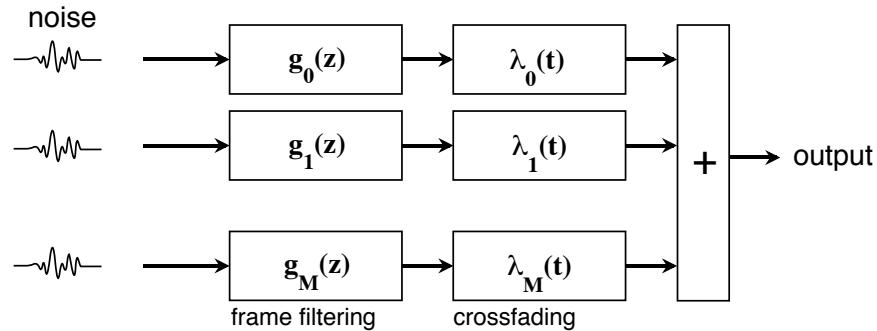
- Reverberation impulse responses may be synthesized by adding direct path, early reflection, and late field components,

$$h(t) = \gamma_D \delta(t) + \gamma_E h_E(t - \tau_E) + \gamma_L h_L(t - \tau_L).$$

- The early reflection component may be estimated based on the geometry.
- The late field is a time-varying convolution of the late-field spectral envelope with a Gaussian noise sequence,

$$h_L(t) = [q(\omega) \cdot \exp\{-t/\tau(\omega)\}] * n(t).$$

Late Field Synthesis Using Frame Filters



- Design a set of frame filters to match the desired spectral envelope at a set of frame times t_i ,

$$t_i = i\beta, \quad i = 0, 1, \dots, M - 1,$$

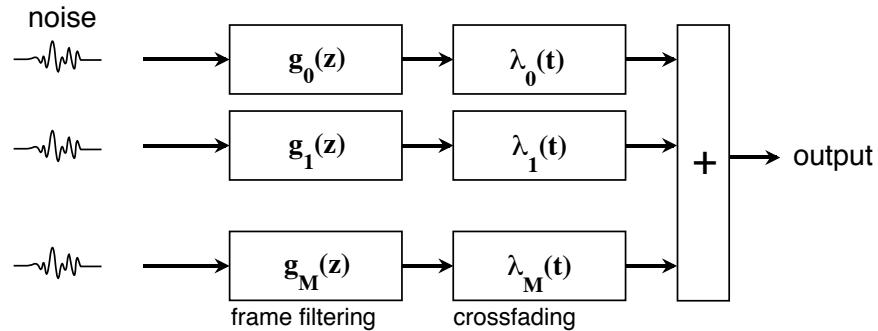
for some interval β ,

$$|g_i(\omega)|^2 = |q(\omega)|^2 \exp\{t_i/\tau(\omega)\};$$

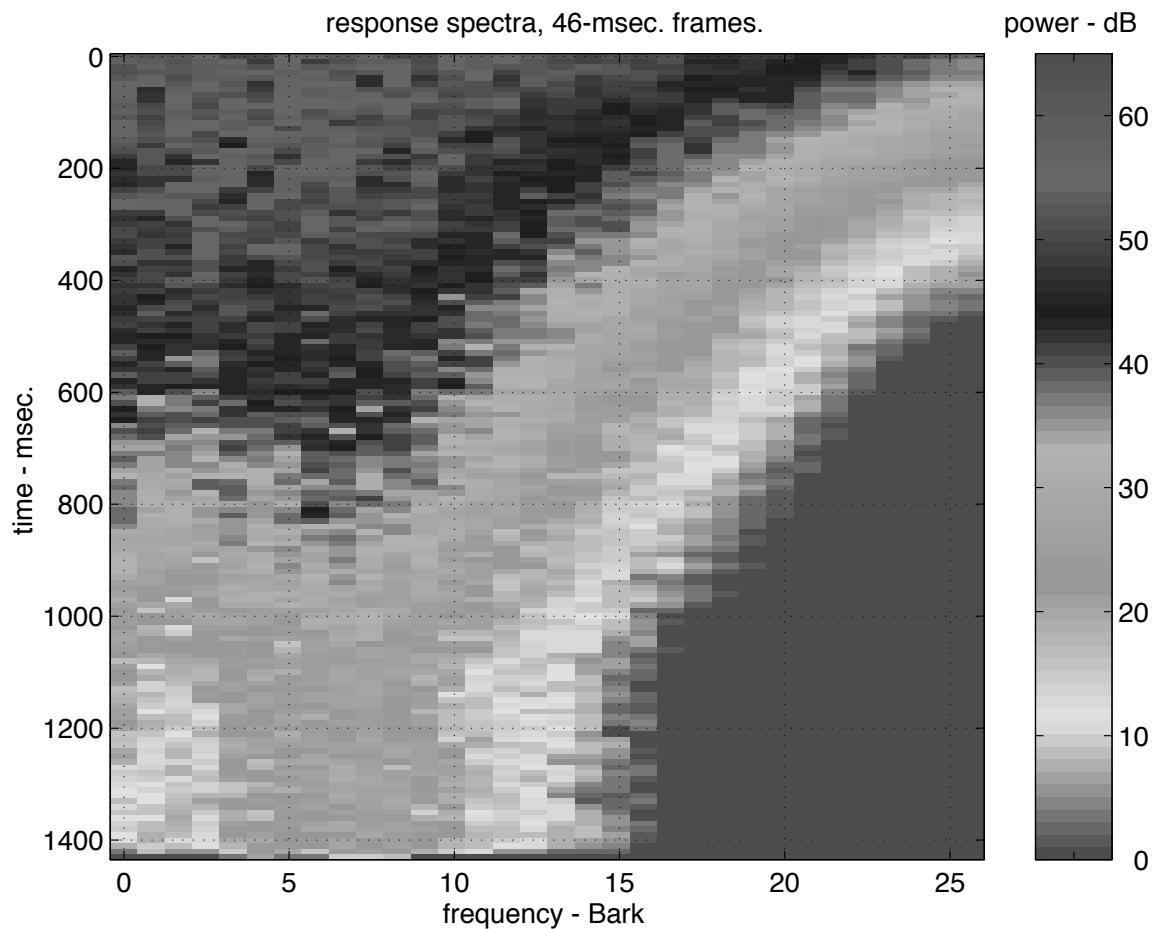
filter phase is unimportant.

- Apply a separate, independent white Gaussian noise sequence $n_i(t)$ to each filter $g_i(t)$ in the sequence to create M filtered sequences.
 - Be sure that each frame filter is driven by the noise sequence at least for the duration of its impulse response before its output is used.
 - A reasonable guide for the filter update time β is the JND between successive filters.
- Form the synthesized late field $h_L(t)$ by cosine crossfading between successive filtered noise sequences.

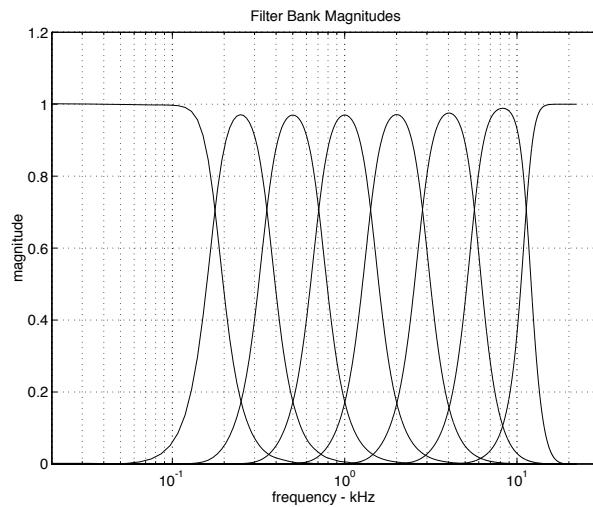
Late Field Synthesis Using Frame Filters



- The power of a filtered white noise sequence is the square filter magnitude, $|g(\omega)|^2$.
- Also, the power spectrum of the sum of independent noise sequences is simply the sum of the power spectra—phase is not an issue if the sequences are independent.
- Therefore, the crossfade between cosine windowed independent sequences gives a crossfaded “power spectrum.”

Late Field Synthesis Using Frame Filters

Late Field Synthesis via Windowed Filter Bank Output

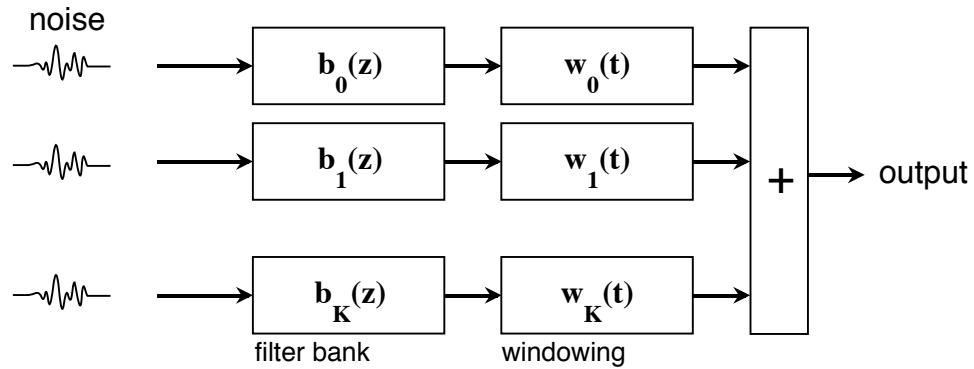


- Design a filter bank using power complementary filters $b_k(t)$, $k = 0, 1, \dots, K$,

$$\sum_{k=0}^K |b_k(\omega)|^2 = 1.$$

(It turns out that squared Butterworth filters have this property.)

Late Field Synthesis via Windowed Filter Bank Output

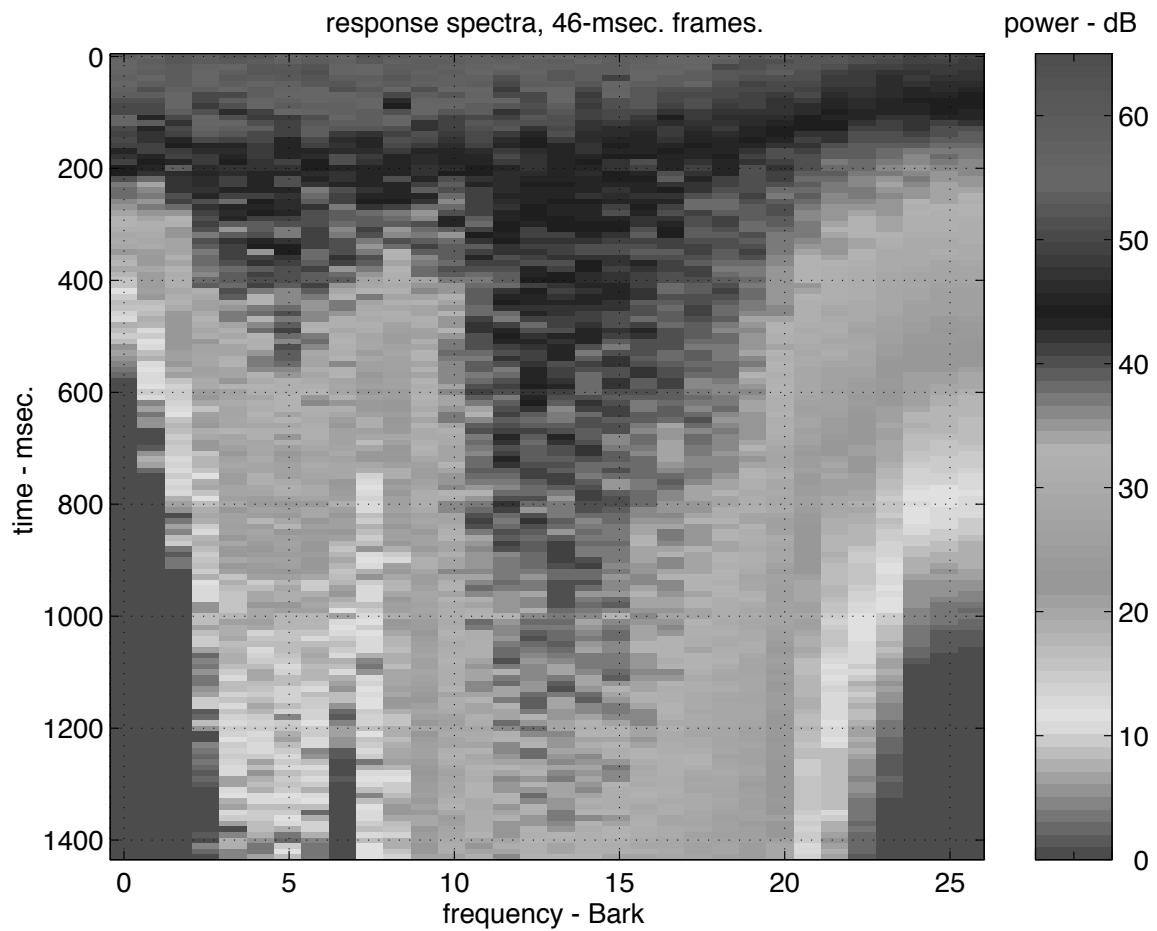


- Apply a separate, independent white Gaussian noise sequence $n_k(t)$ to each filter $b_k(t)$ in the sequence to create K filtered sequences.
- Window each filtered sequence according to

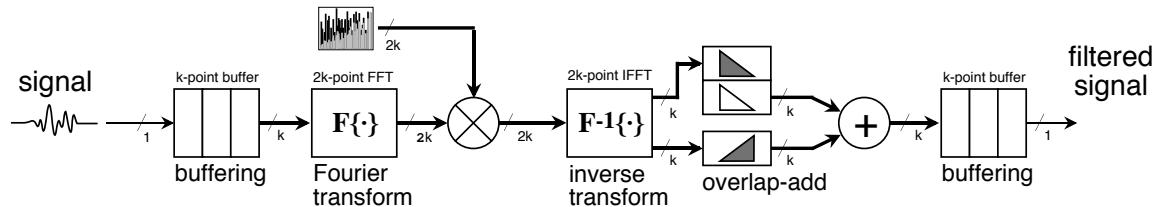
$$w_k(t) = |q(\omega_k)| \exp\{-t/\tau(\omega_k)\},$$

where ω_k is the band center for the k th filter.

- Sum the windowed filter outputs.

Late Field Synthesis via Windowed Filter Bank Output

Low-Latency Convolution



- Implementing an N -tap FIR Filter using overlap-add FFT-based convolution requires $o(12 \log_2 N)$ operations per sample output—significantly less than the $o(N^2)$ operations required for direct time-domain convolution.
- FFT methods, however, have a computational latency: Samples are processed in blocks and must be buffered. The processing is may be designed to be completed when the next block of data is available, making the total latency $2N$ samples.
- For reverberation impulse responses which can be many seconds long, a low-latency computationally efficient convolution method is desired.

Low-Latency Convolution

- How do you retain the computational benefits of FFT processing, but have a reasonable latency?
- Divide the impulse response into two sections,

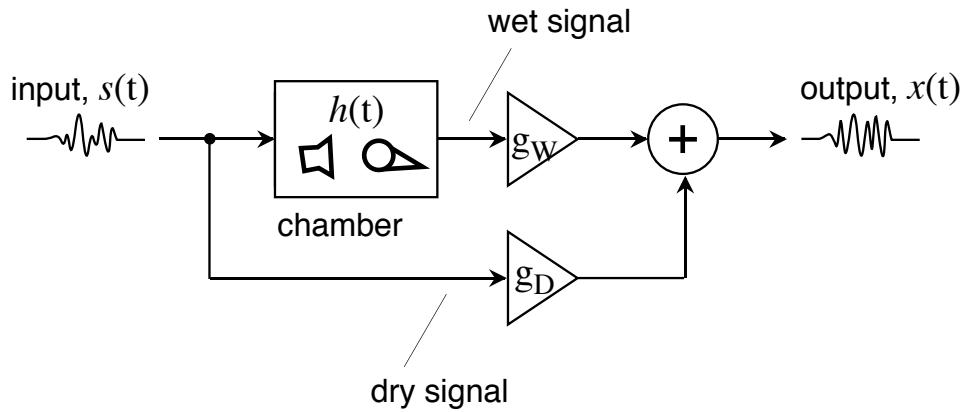
$$h(t) = h_1(t) + \delta(t - N/2) * h_2(t),$$

and use FFT techniques to separately process each impulse response section.

- The computational cost increases, but the latency is cut in half.
- To achieve a given latency, iteratively split the first section to two.
 - At each iteration, the latency is halved, and the total computation modestly increased.
 - In this way, shorter impulse response blocks are used to give low latency, and larger ones for computational efficiency.

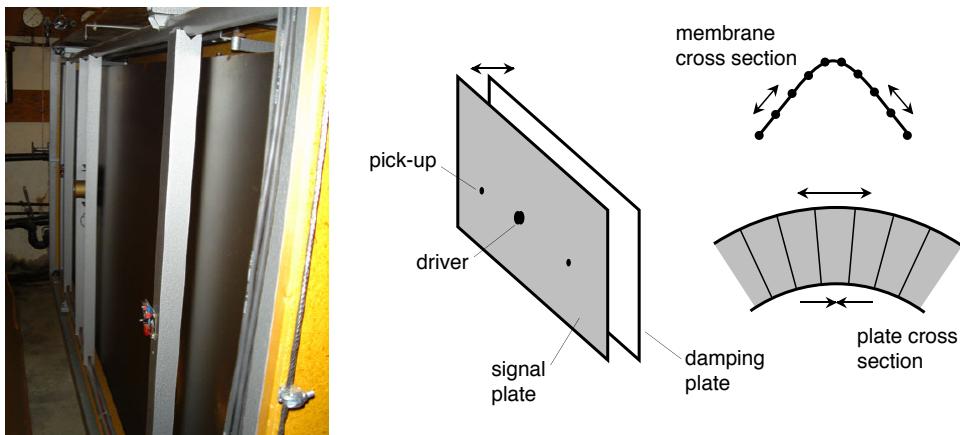
14. Mechanical and Acoustic Reverberation

Chambers

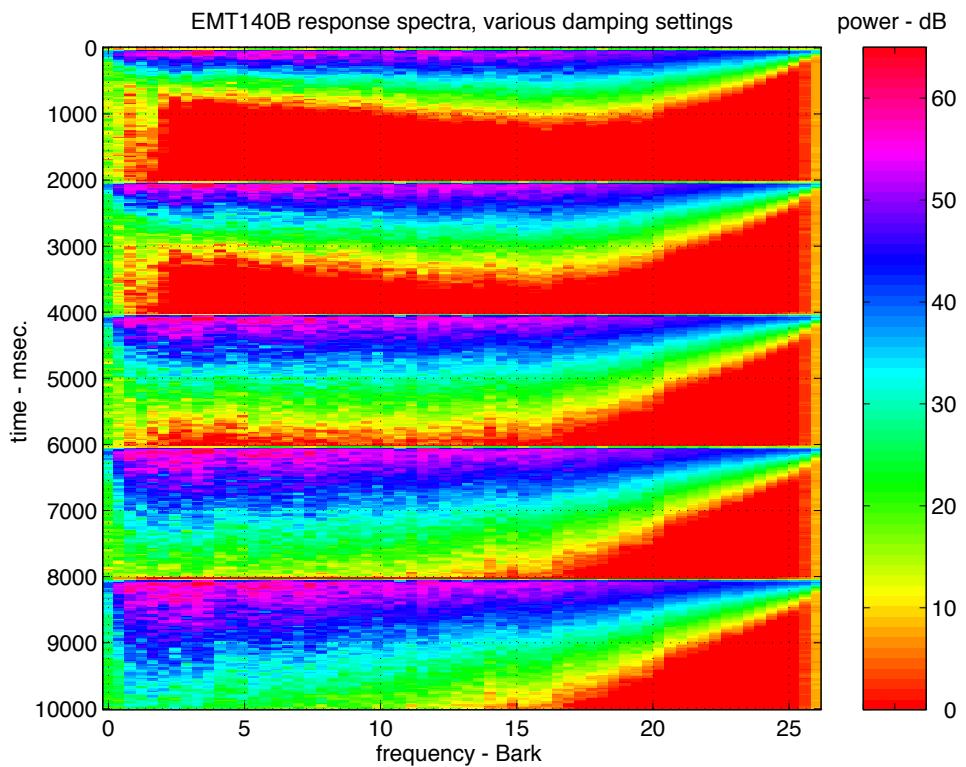


- Reverberation can be added to a track by playing the track in a chamber and recording the response.
- Wet and dry signal levels and equalization are controlled at the mixing board; other reverberation parameters are determined by the chamber acoustics.

Plates

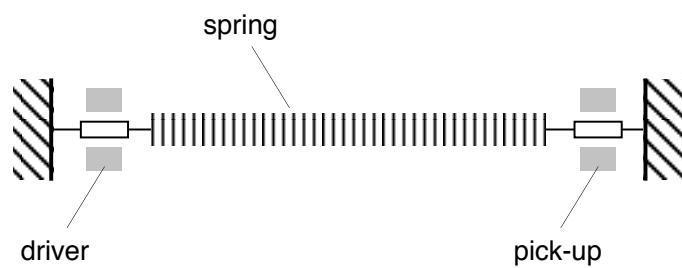
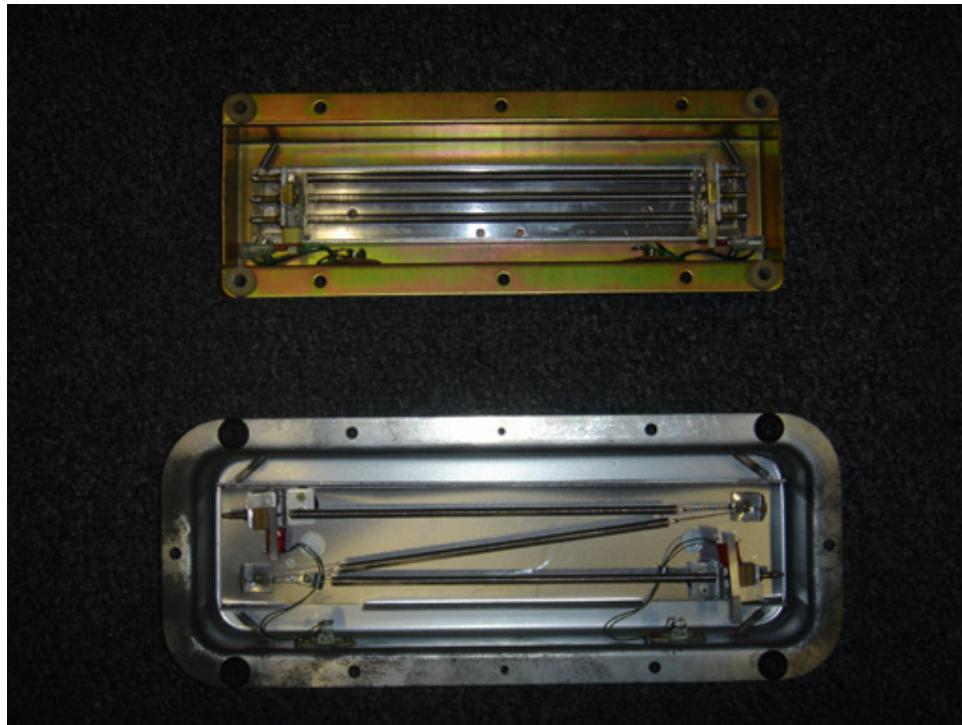


- Tensioned plates support wave propagation.
- Along a transverse wave, restoring forces are provided by extension of the plate on one side and compression of the plate on the other side.
- These forces are different than those at play in a membrane, in which the material is sufficiently thin that the restoring force results from material stretching. As a result wave propagation on a plate is dispersive, with high frequencies traveling faster than low frequencies, and multimodal, with different propagation modes coupling into each other.
- Thermoviscous losses and edge reflections dominate the low-frequency T_{60} , whereas radiation losses contribute mainly to the high-frequency decay rate.
- In the EMT140 plate reverberator, a damping plate is positioned near the signal plate and provides additional low-frequency losses.

EMT140 Plate Reverberator

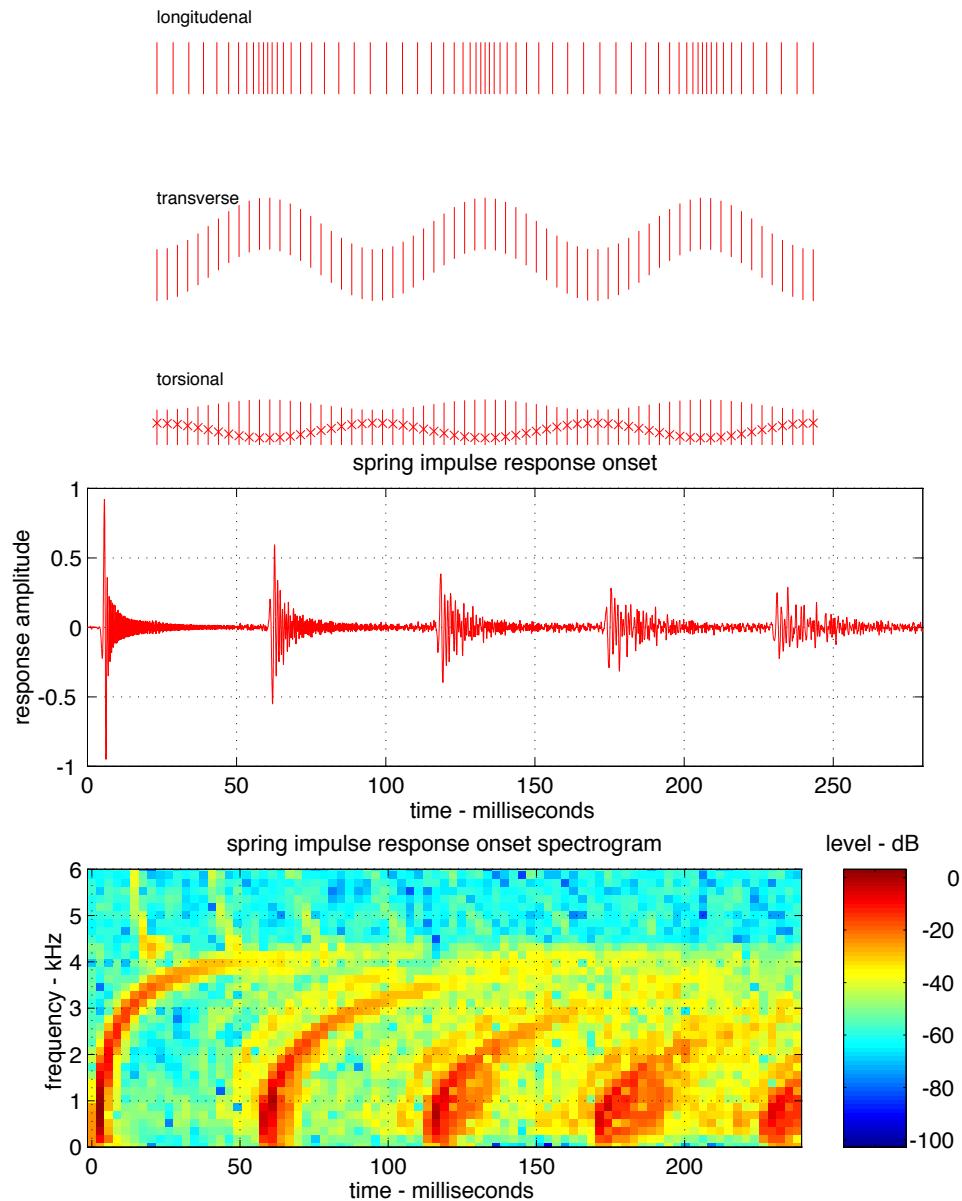
- EMT140 plate reverberator spectrograms at various damping settings.

Spring Reverberators

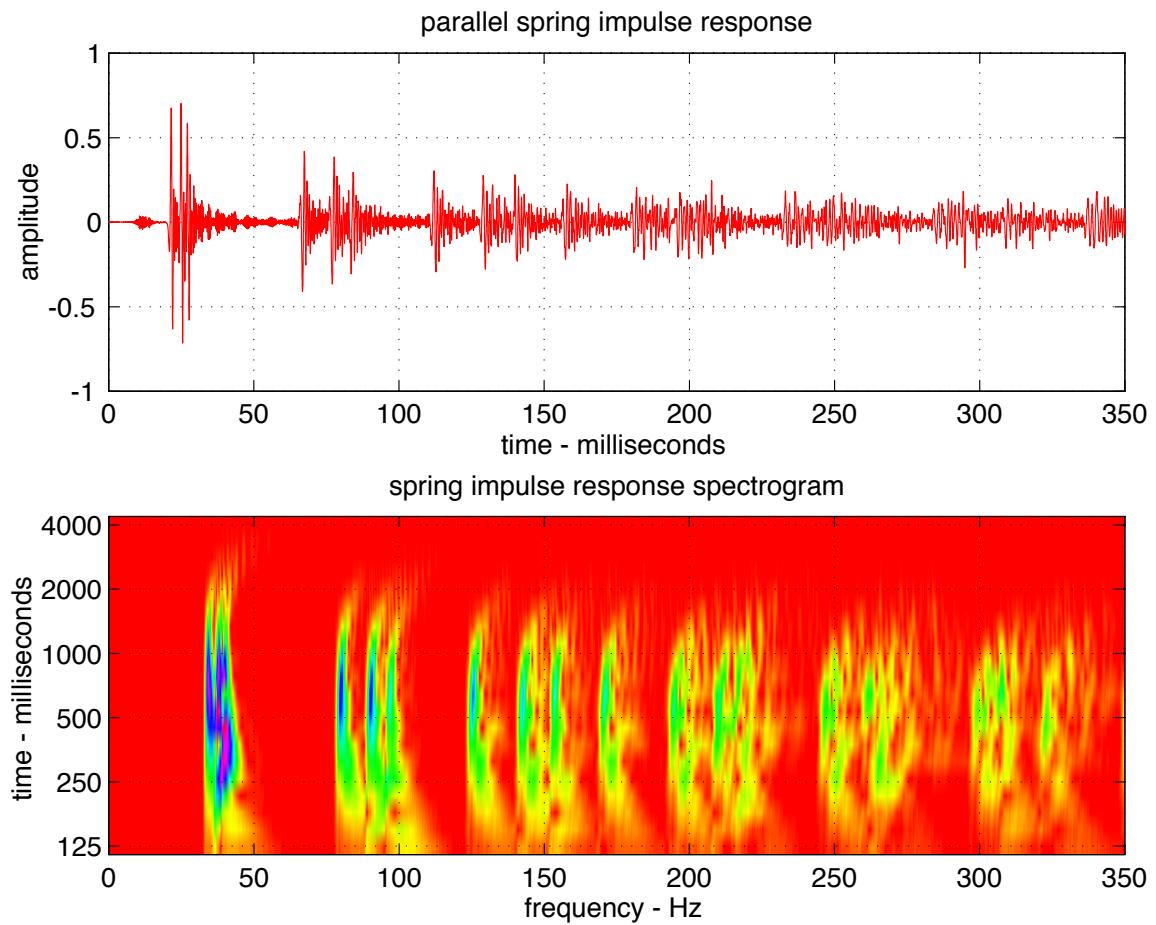


- Modern spring reverberators are fitted with magnetic transducers to excite and detect torsional waves on the spring.

Spring Propagation

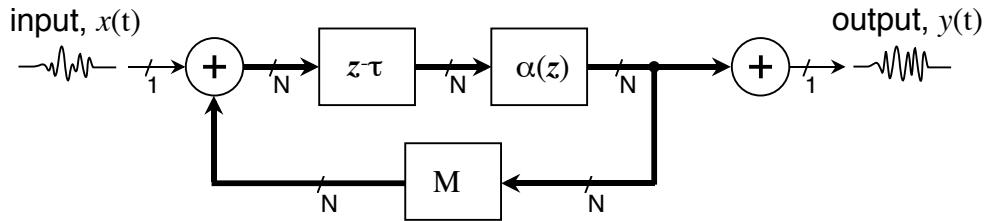


- Tensioned springs support dispersive wave propagation, giving spring reverberators a unique sound.
- Energy lost during propagation and at reflecting boundaries and scattering junctions determine decay rates.

Spring Impulse Response

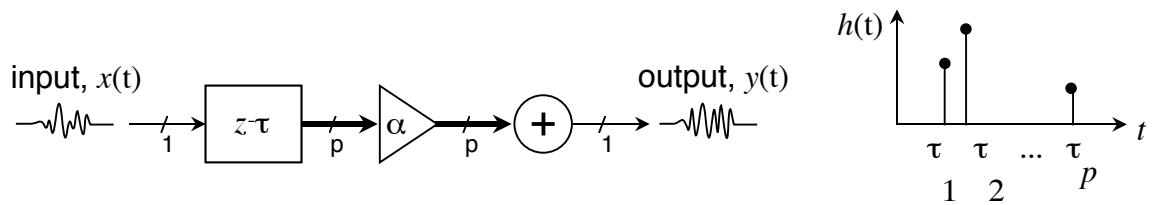
15. Feedback Delay Network Reverberators

Outline



- Delay-line-based echo.
- Delay line with feedback, T_{60} control.
- Parallel and series allpass reverberation structures.
- Feedback delay network reverberators.

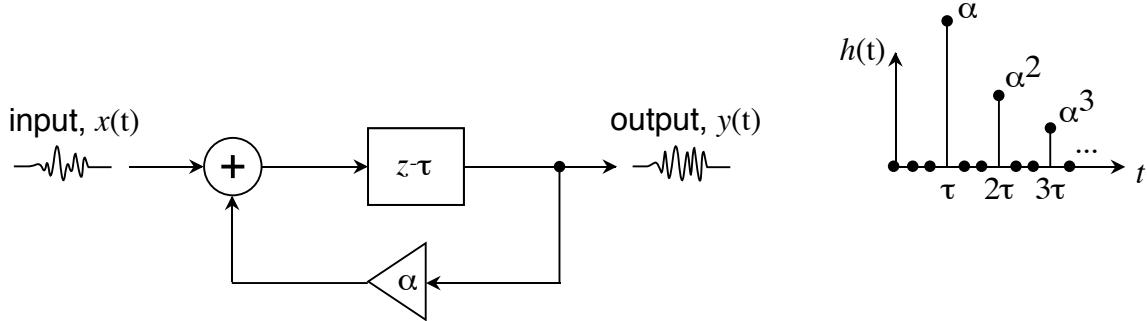
Echo



- A delay line with multiple taps may be used to implement a series of echos,

$$h(t) = \sum_{i=0}^p \alpha_i \delta(t - \tau_i).$$

Allpass and Comb Filters



Comb Filter

- In a *comb filter*, the input and a portion α of the output are combined and delayed τ samples to form the output.
- A first-order comb filter has difference equation

$$x(t) = s(t - \tau) + \alpha x(t - \tau),$$

and transfer function

$$H(z) = \frac{z^{-\tau}}{1 - \alpha z^{-\tau}}.$$

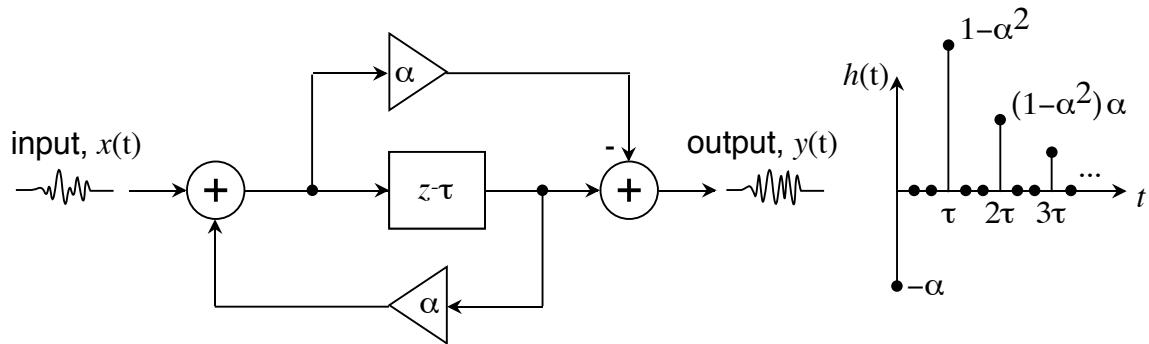
Its impulse response is

$$h(t) = \sum_{i=0}^{\infty} \alpha^{i-1} \delta(t - i \cdot \tau).$$

- The comb filter derives its name from its magnitude transfer function, which looks comb-like,

$$|H(e^{j\omega})|^2 = \frac{1}{1 + \alpha^2 - 2\alpha \cos(\tau\omega)}.$$

Allpass Filter



- An *allpass filter* has the structure of the comb filter with a portion of the input appearing directly at the output.
- A first-order allpass filter has difference equation

$$x(t) = s(t - \tau) + \alpha[x(t - \tau) - s(t)],$$

and transfer function

$$H(z) = \frac{-\alpha + z^{-\tau}}{1 - \alpha z^{-\tau}}.$$

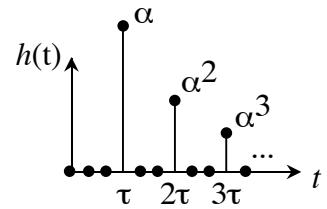
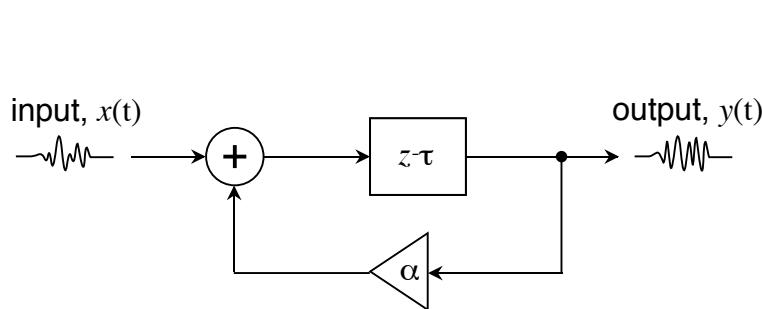
Its impulse response is

$$h(t) = \sum_{i=0}^{\infty} \alpha^{i-1} \delta(t - i \cdot \tau).$$

- The allpass filter gets its name from the fact that its magnitude transfer function is one,

$$|H(e^{j\omega})| = 1.$$

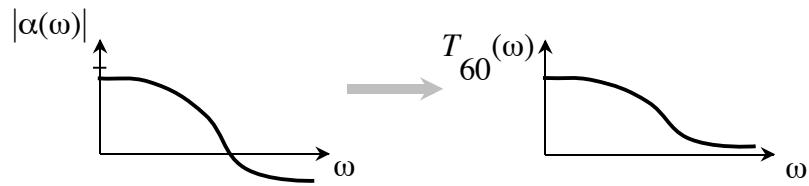
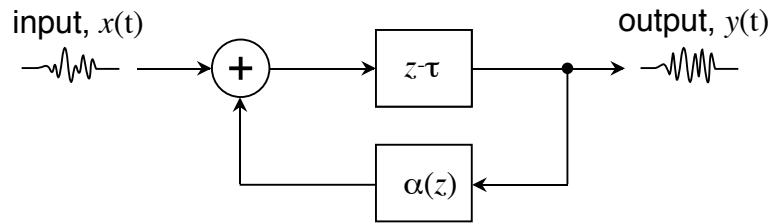
Comb and Allpass Filter Decay Rate



- With a delay of τ samples at a sampling rate of f_s , and feedback portion α , the comb (or allpass) T_{60} is

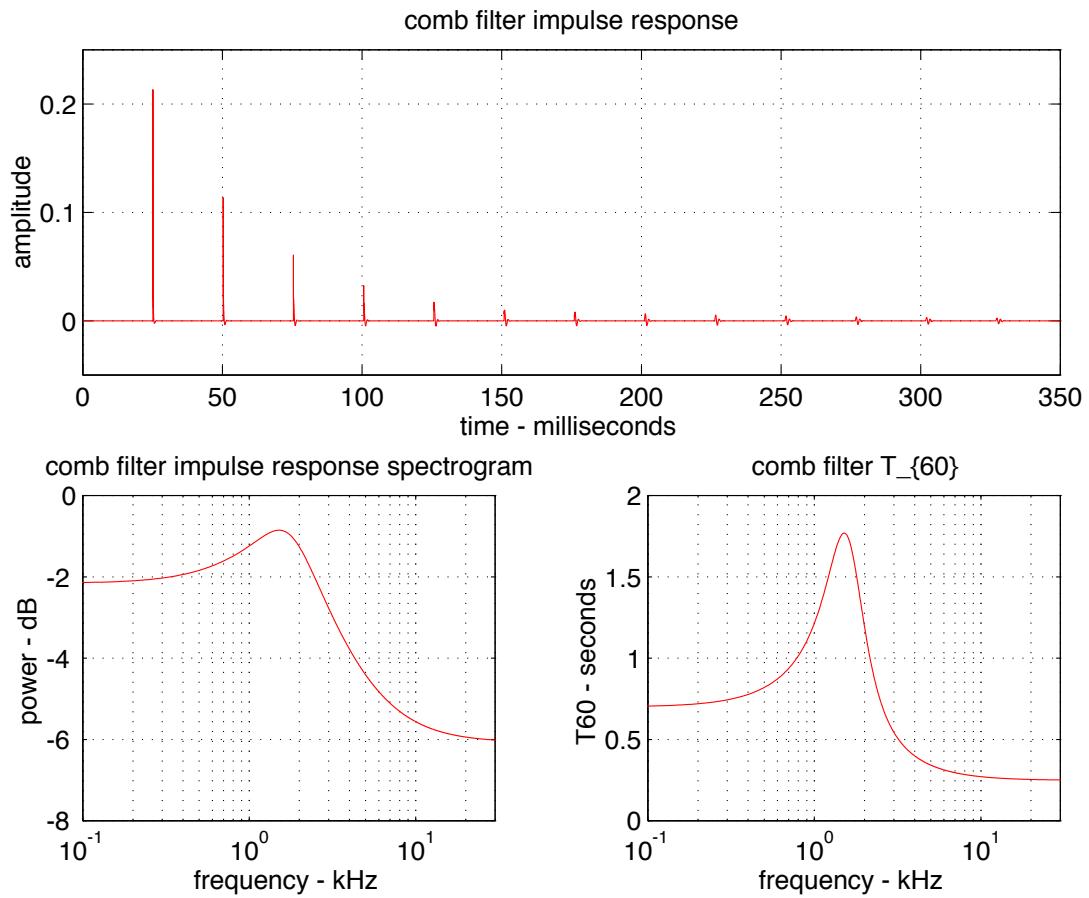
$$T_{60} = -\frac{60 \cdot \tau / f_s}{20 \log_{10} |\alpha|}.$$

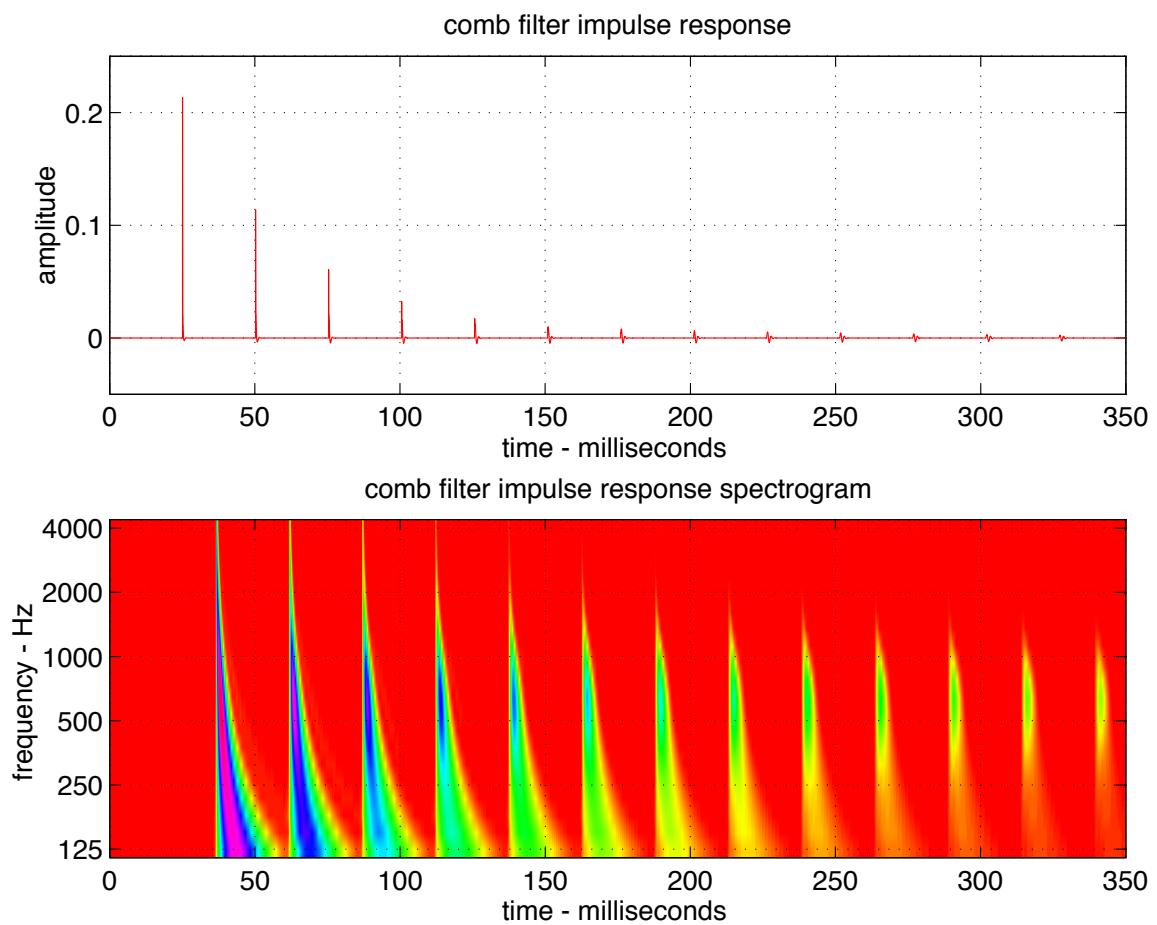
Frequency-Dependent Decay Rate



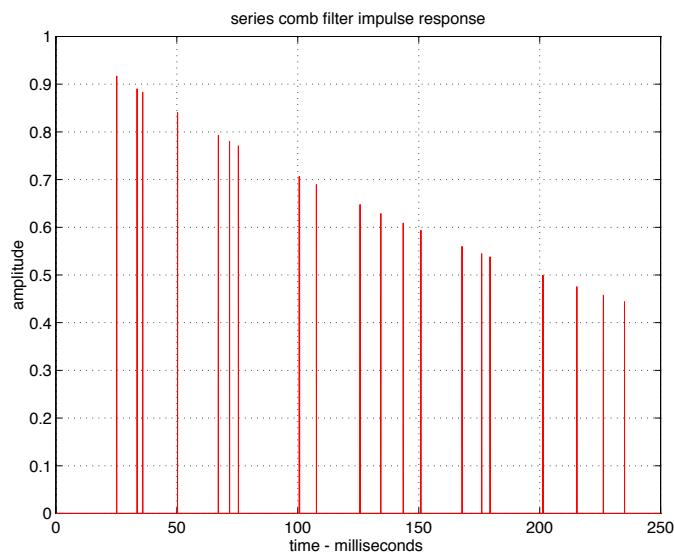
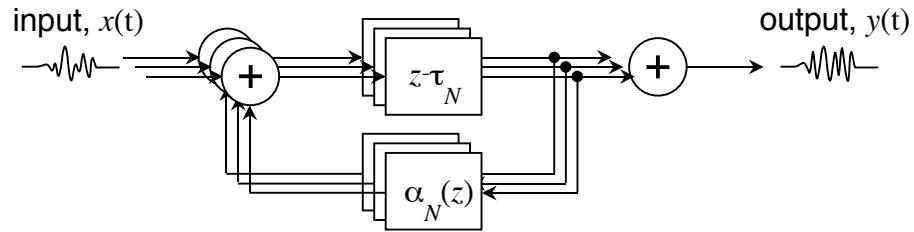
- When the feedback takes the form of a filter, $\alpha(z)$, the decay rate becomes frequency-dependent,

$$T_{60}(\omega) = -\frac{60 \cdot \tau / f_s}{20 \log_{10} |\alpha(\omega)|}.$$

Frequency-Dependent Decay Rate Example

Frequency-Dependent Decay Rate Example

Feedback Delay Networks

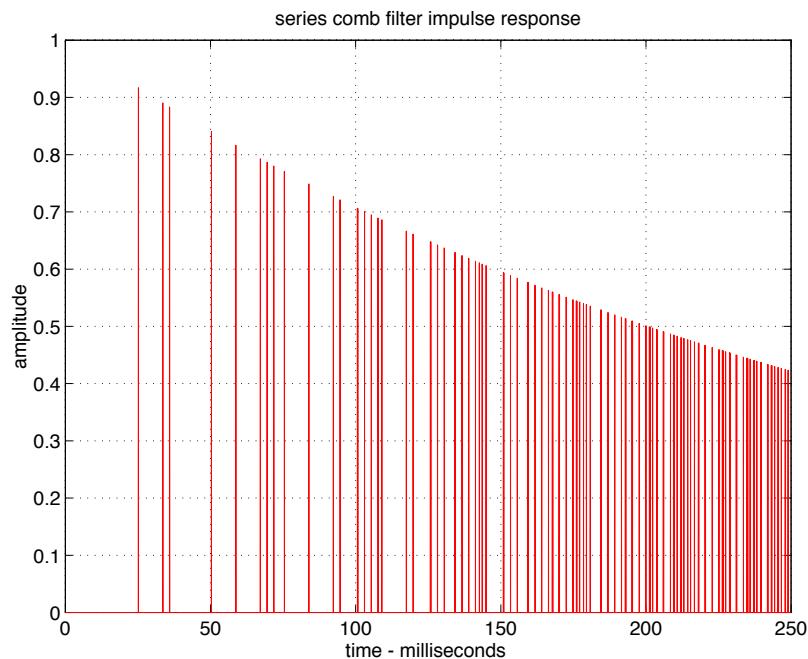
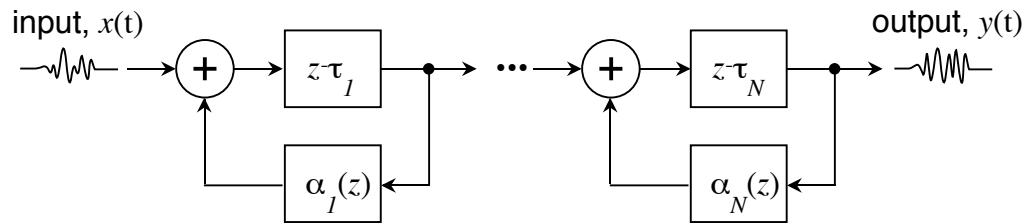


Parallel Comb Filter Network

- By placing a set of N comb filters with incommensurate lengths in parallel, the echo density is increased roughly by a factor of N .
- Note that by designing each feedback filter so as to account for its corresponding delay line length, all comb structures will have the same decay rate,

$$|\alpha_i(\omega)| = \exp\{\ln(0.001) \tau / [f_s \cdot T_{60}(\omega)]\}.$$

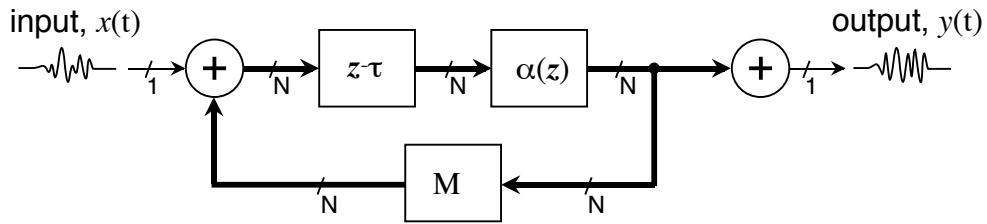
Series Comb Filter Network



- By cascading comb filters of incommensurate lengths, the echo density is greatly increased, with each echo generating a series of echos in the subsequent stage.
- Note that by designing each feedback filter to have a given decay rate, the cascade will have that decay rate,

$$|\alpha_i(\omega)| = \exp\{\ln(0.001) \tau / [f_s \cdot T_{60}(\omega)]\}.$$

Feedback Delay Network

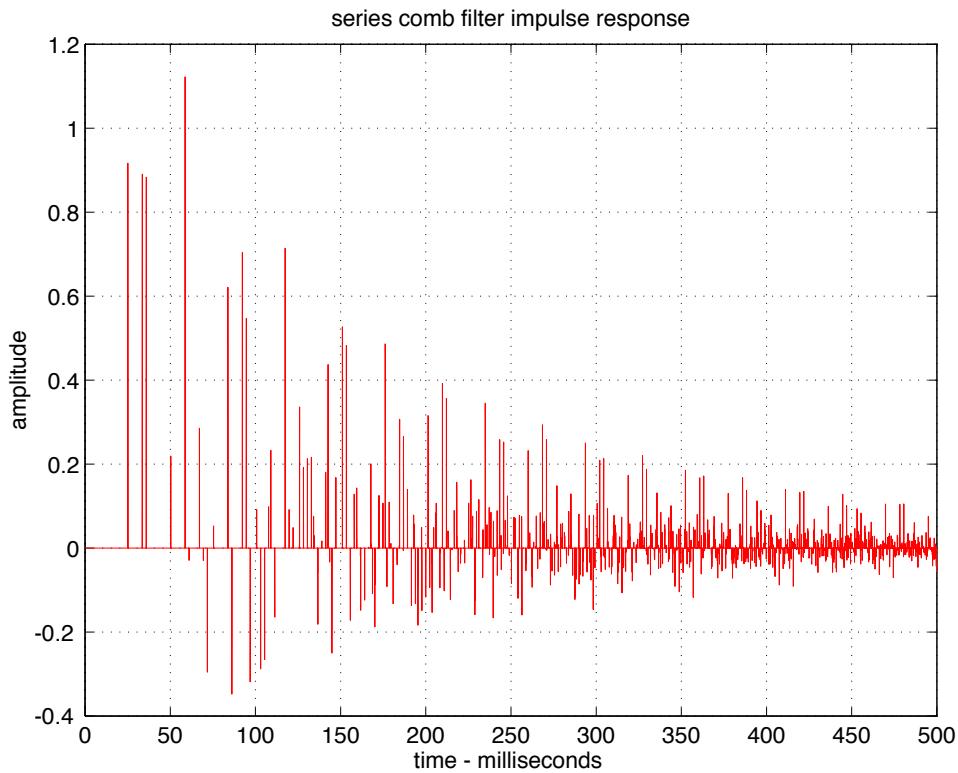


- The structures above are instances of a general structure called a *feedback delay network*, in which states are separately delayed and filtered, and mixed before being fed back.
- A parallel comb filter network results from an identity mixing matrix, $M = I$.
- Adding off-diagonal elements,

$$M = \begin{bmatrix} 1 & 1 & & 0 \\ & \ddots & \ddots & 1 \\ 0 & & \ddots & 1 \end{bmatrix},$$

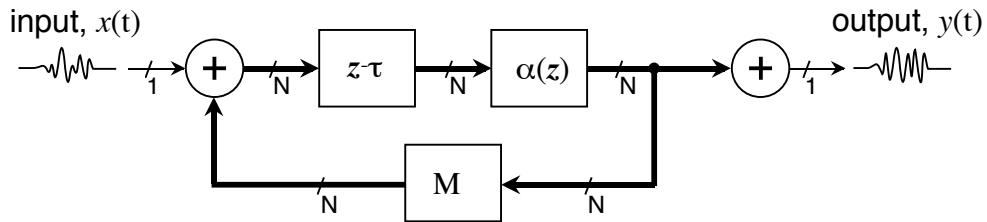
gives a series combination.

Feedback Delay Network



- By designing the feedback filters to give the same desired decay rate when used with an identity mixing matrix, any unitary mixing matrix M will be combining signals with like spectra, and the design decay rate will be preserved.
- By choice of mixing matrix, the rate of echo density increase may be controlled.

Feedback Delay Network Design Method



To design a feedback delay network for simulating late-field reverberation, do the following:

- Remove the feedback filtering and mixing matrix, and choose a set of delay line lengths such that an impulse input to the FDN results in a noise-like sequence which sounds rather white.
- Pick a mixing matrix to achieve the desired rate of echo density increase.
- Design feedback filters and overall equalization according to the given decay rate as a function of frequency.

Equalization and Filtering

16. LaPlace and Fourier Transforms; Bode Plots

Linear Systems

Relationship between LaPlace and Fourier Transforms:

LaPlace:

$$X(s) = \int_0^{\infty} x(t)e^{-st} dt$$

- for study of lumped systems

Fourier:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt$$

- for study of signals

Differences between transforms:

- limits of integration
- kernel

If we assign $s = j\omega$:

- both kernels the same

- s is *complex frequency* $\sigma + j\omega$

System Analysis:

- let $x(t)$ be the impulse response of the system
- for realizable system:

$$x(t) = 0, \quad t < 0$$

- integration limits of $[0, \infty]$ → look at causal systems

Can Fourier Transform be used to study impulse responses?

- F.T. is invertible, unique, and can represent any signal $x(t)$ if:
 - $\int |x(t)| < \infty$ (global condition)
 - $x(t)$ has a finite number of maxima, minima and discontinuities in any finite region (local condition)
- for 'causal' signals, limits of integration do not matter
- F.T. of the impulse response is the frequency response of the system
- if desired, entire s-plane representation can be recovered from the Fourier Transform

Why use generalized frequency s for lumped circuits?

- compact representation:

for circuit with L capacitors and M inductors and any number of resistors:

$$H(s) = \frac{\sum_0^N b_n s^n}{1 + \sum_1^N a_n s^n}$$

where $N = L + M$

- compare to Fourier Transform:
 - F.T. gives an ∞ -dimensional vector
 - Laplace Transform gives quotient of order N polynomials

Poles and Zeros:

$$H(s) = \frac{\sum_0^N b_n s^n}{1 + \sum_1^N a_n s^n} = \frac{\prod_N (s - z_n)}{\prod_N (s - p_n)}$$

where the right-hand side is obtained by factoring the numerator and denominator polynomials.

- at zeros: $H(s) = 0 \rightarrow 'o'$
- at poles: $H(s) = \infty \rightarrow 'x'$
- away from poles and zeros, $|H(s)|$ behaves like a 'stretched membrane'

Study of poles and zeros reveals multiple properties of system:

- can be used to predict stability of system
 - all poles in left-half-plane \rightarrow stable system
 - Routh stability test gives necessary conditions for stability:
 - i. all coefficients have the same sign
 - ii. all coefficients between highest and lowest power are nonzero unless all even- or odd- order coefficients are zero
- can be used to determine properties of phase

example: if all poles and zeros are in left-half plane, system is minimum-phase: $\int_0^\tau |h(t)|^2 dt$ is maximized for any τ for all

filters sharing $|H(s)|$. System thus has the “most compact” impulse response.

- pole, zero locations used in control / feedback theory to predict stability of systems using feedback.

Summary:

- For study of lumped systems, LaPlace Transform gives a compact representation and provides insight that cannot be obtained with Fourier Transforms.
- impulse response is causal for most systems of interest.
- for lumped systems, $H(s)$ is a rational function

→ use Laplace Transform for system analysis

Signal Analysis:

- want bilateral form of integral in case signal is not time-limited or non-causal
- ω -axis is complete for signals
- if “delay” or noise is present, signal cannot be represented by a rational function (ratio of polynomials) in s .

Summary:

Unless a signal is tightly coupled to a system being studied (e.g. its impulse response), generalizing to complex frequency is not likely

to be of any benefit. For non-causal systems, bi-lateral integration is necessary to include information for $t < 0$.

→ use Fourier Transform for generic signal analysis

Inverse Laplace Transform:

$$x(t) = \frac{1}{2\pi j} \int_C X(s)e^{st}ds$$

this integration is “difficult”. if there are poles within the contour, residues have to be computed.

But, if the system is proper, *i.e.*, if

$$H(s) = \frac{\sum_0^M b_m s^m}{1 + \sum_1^N a_n s^n}, \quad M < N,$$

then as $s \rightarrow \infty$, $H(s) \rightarrow 0$, so the contribution from the edge of the contour is zero.

Then,

$$x(t) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} X(s)e^{st}ds$$

where c is a positive constant, can have $c \rightarrow 0$. In this case, the Inverse Laplace Transform becomes similar to the Inverse Fourier Transform:

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

Unilateral \mathcal{L} Transforms:

$$\begin{aligned}
 \delta(t) &\iff 1 \\
 u(t) &\iff \frac{1}{s} \\
 e^{-\alpha t} &\iff \frac{1}{s + \alpha} \\
 t^n &\iff \frac{n!}{s^{n+1}} \\
 t^n e^{-\alpha t} &\iff \frac{n!}{(s + \alpha)^{n+1}} \\
 \sin(\omega_0 t) &\iff \frac{\omega_0}{s^2 + \omega_0^2} \\
 \cos(\omega_0 t) &\iff \frac{s}{s^2 + \omega_0^2} \\
 e^{-\alpha t} \cos(\omega_0 t) &\iff \frac{s + \alpha}{(s + \alpha)^2 + \omega_0^2}
 \end{aligned}$$

Unilateral \mathcal{L} Transform Theorems:

$$\begin{aligned}
 ax_1(t) + bx_2(t) &\iff aX_1(s) + bX_2(s) \\
 x(t - T)u(t - T) &\iff X(s)e^{-sT}, \quad T > 0 \\
 tx(t) &\iff -\frac{dX(s)}{ds} \\
 e^{-\alpha t}x(t) &\iff X(s + \alpha) \\
 x(at) &\iff \frac{1}{a}X\left[\frac{s}{a}\right], \quad a > 0 \\
 \frac{dx(t)}{dt} &\iff sX(s) - x(0) \\
 \int_0^t x(\tau)d\tau &\iff \frac{X(s)}{s} \\
 x(0) &\iff \lim_{s \rightarrow \infty} sX(s) \\
 x(\infty) &\iff \lim_{s \rightarrow 0} sX(s) \\
 \int_0^t x_1(\tau)x_2(t - \tau)d\tau &\iff X_1(s)X_2(s)
 \end{aligned}$$

Fourier Transforms

$$\begin{aligned}
\delta(t) &\iff 1 \\
1 &\iff \delta(f) \\
e^{-\pi(t/\tau)^2} &\iff \tau e^{-\pi(f\tau)^2} \\
e^{-\alpha t} u(t) &\iff \frac{1}{\alpha + j2\pi f}, \quad \alpha > 0 \\
u(t) &\iff \frac{\delta(f)}{2} + \frac{1}{j2\pi f} \\
\text{sgn}(t) &\iff \frac{1}{j\pi f} \\
\frac{1}{\pi t} &\iff -j\text{sgn}(f) \\
e^{-\alpha|t|} &\iff \frac{2\alpha}{\alpha^2 + (2\pi f)^2} \\
e^{j2\pi f_0 t} &\iff \delta(f - f_0) \\
\sin(2\pi f_0 t) &\iff \frac{\delta(f - f_0) - \delta(f + f_0)}{2j} \\
\cos(2\pi f_0 t) &\iff \frac{\delta(f - f_0) + \delta(f + f_0)}{2} \\
e^{j2\pi f_0 t} u(t) &\iff \frac{\delta(f - f_0)}{2} + \frac{1}{j2\pi} \left[\frac{1}{f - f_0} \right] \\
\sin(2\pi f_0 t) u(t) &\iff \frac{\delta(f - f_0) - \delta(f + f_0)}{4j} + \frac{1}{2\pi} \left[\frac{f}{f_0^2 - f^2} \right] \\
\cos(2\pi f_0 t) u(t) &\iff \frac{\delta(f - f_0) + \delta(f + f_0)}{4} + \frac{1}{j2\pi} \left[\frac{f}{f^2 - f_0^2} \right] \\
\dot{\delta}(t) &\iff j2\pi f \\
te^{-\alpha t} u(t), \alpha > 0 &\iff \frac{1}{(\alpha + j2\pi f)^2} \\
t &\iff \frac{j\dot{\delta}(f)}{2\pi}
\end{aligned}$$

Bode Plots:

- Find $|H(s)|$, $\angle H(s)$

$$|H(s)| = \frac{\prod |s - z_i|}{\prod |s - p_i|}$$

for frequency response, $s = j\omega$.

example:

$$H(s) = \frac{a}{s + a}$$

- at $\omega = a$:

$$|H| = \left| \frac{a}{(j+1)a} \right| = \left| \frac{1}{1+j} \right| = \frac{\sqrt{2}}{2} = -3dB$$

- one octave either side of a :

$$|H| = \frac{4}{\sqrt{5}} \approx -1dB$$

- for $\omega \gg a$,

$$|H| \approx \frac{a}{\omega}$$

for $\angle H$:

$$\angle H(s) = \sum \angle(s - z_i) - \sum \angle(s - p_i)$$

if $H(s) = \frac{a}{s+a}$,

- for $\omega \ll a$:

$$\angle H(j\omega) = 0$$

- for $\omega = a$:

$$\angle H(j\omega) = -45^\circ$$

- for $\omega \gg a$:

$$\angle H(j\omega) = -90^\circ$$

- an asymptotic line plotted over log of frequency which goes through -45 degrees at $w = a$ and hits the -90 and 0 degree asymptotes one decade above and below a will be within 6 degrees of $\angle H(j\omega)$.

Bode Plots for systems with multiple poles and zeros:

- for log-log magnitude, contributions from separate poles and zeros “add”.
- for phase, contributions from separate poles and zeros also add.

Circuit Properties:

- RL or RC circuits are referred to as 'relaxation' circuits. All poles are on the real axis in the left-half-plane.
- LC circuits are passive, lossless, and marginally stable. All poles are on the $j\omega$ axis.

note: 'passive' indicates that the net energy which can be extracted from the system cannot exceed a finite bound, which depends on the initial conditions of the system.

'stable' indicates either that a bounded input always produces a bounded output, or that the effects at the output of small perturbations on the input will remain small, or both.

'lossless' indicates that the system does not absorb energy. The output energy is equal to the input energy.

- for circuits with real-valued elements, all poles and zeros are either real, or occur in complex-conjugate pairs.

$$(s + a + bj)(s + a - bj) = s^2 + 2as + (a^2 + b^2)$$

Biquad Filters:

$$H(s) = \frac{b_2 s^2 + b_1 s + b_0}{a_2 s^2 + a_1 s + a_0}$$

- numerator and denominator polynomials are both quadractic, hence 'biquad'
- system has two poles and two zeros: for real coefficients, these can be complex-conjugate pairs, or real-valued.
- any real-coefficient transfer function $Y(s)$ can be realized by a cascade of real-coefficient biquads, $\prod B_n(s)$.

Example: Resonant peaking filter

$$H(s) = \frac{\frac{s}{\omega_0}}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q} \left(\frac{s}{\omega_0}\right) + 1}$$

- zero at $s = 0$
- for $Q > 0.5$, complex-conjugate poles, with real part at $-\omega_0/2Q$, and radius ω_0 . Stable for positive Q.
- for $\omega \ll \omega_0$, $H(j\omega) \approx \frac{j\omega}{\omega_0}$.
- for $\omega \gg \omega_0$, $H(j\omega) \approx \frac{\omega_0}{j\omega}$.
- asymptotic gains intersect at $\omega = \omega_0$, at zero dB. Gain of system at $\omega = \omega_0$ is equal to Q.
- in log-log space, transfer function is symmetric about $\omega = \omega_0$.
- asymptotic low-frequency phase is 90 degrees.

- asymptotic high-frequency phase is -90 degrees.
- phase is zero at $\omega = \omega_0$.
- where gain is 3dB below peak,

$$\frac{\omega}{\omega_0} = -\frac{1}{2Q} + \sqrt{\frac{1}{4Q^2} + 1}$$

for $\omega < \omega_0$. Phase here is 45 degrees.

- for large Q , bandwidth between half-power (-3dB) points is ω_0/Q .

Eigenfunctions of LTI systems:

recall the convolution integral:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)d\tau$$

let $x = e^{st}$:

$$y(t) = \int_{-\infty}^{\infty} h(\tau)e^{s(t-\tau)}d\tau = \left[\int_{-\infty}^{\infty} h(\tau)e^{-s\tau}d\tau \right] e^{st} = H(s)e^{st},$$

since

$$H(s) = \int_{-\infty}^{\infty} h(t)e^{-st}dt$$

so e^{st} is an eigenfunction of the system, with $H(s)$ as the eigenvalue.

17. Transform Mechanics

Inverse Laplace Transform

$$\frac{1}{2\pi j} \int_C X(s)e^{st} ds$$

Residue Theorem

if

$$f(s) = \prod \frac{1}{s - a_i}$$

then

$$\text{Res}_i = \lim_{s \rightarrow a_i} (s - a_i) \cdot f(s)$$

$$\int_C f(s) ds = 2\pi j \cdot \sum \text{Res}$$

\forall residues within C .

Example: $X(s) = 1/(s + a)$

$$\frac{1}{2\pi j} \int_C X(s)e^{st} ds = \frac{1}{2\pi j} \cdot 2\pi j \cdot \lim_{s \rightarrow -a} e^{st} = e^{-at}$$

Example:

$$X(s) = \frac{\omega_0}{(s + j\omega_0)(s - j\omega_0)}$$

$$\begin{aligned}\sum \text{Res} &= \lim_{s \rightarrow j\omega_0} \frac{\omega_0}{s + j\omega_0} e^{j\omega_0 t} + \lim_{s \rightarrow -j\omega_0} \frac{\omega_0}{s - j\omega_0} e^{-j\omega_0 t} \\ &= \frac{1}{2j\omega_0} (e^{j\omega_0 t} - e^{-j\omega_0 t}) \\ &= \sin(\omega_0 t)\end{aligned}$$

Example:

$$X(s) = \frac{s}{(s + j\omega_0)(s - j\omega_0)}$$

$$\begin{aligned}\sum \text{Res} &= \lim_{s \rightarrow j\omega_0} \frac{s}{s + j\omega_0} e^{j\omega_0 t} + \lim_{s \rightarrow -j\omega_0} \frac{s}{s - j\omega_0} e^{-j\omega_0 t} \\ &= \frac{1}{2\omega_0} (e^{j\omega_0 t} + e^{-j\omega_0 t}) \\ &= \cos(\omega_0 t)\end{aligned}$$

For proper systems which are stable:

$$H(s) = \frac{\sum_{i=0}^{i=n} b_i s^i}{\sum_{k=0}^{k=m} a_k s^k}, \quad m > n$$

$$\int_{\text{arc}} = 0$$

so we are integrating along $s = j\omega$.

Partial Fraction Expansion:

Example:

$$\begin{aligned} H(s) &= \frac{2s + 3}{s^2 + 3s + 2} \\ &= \frac{a}{s + 1} + \frac{b}{s + 2} \\ &= \frac{a(s + 2) + b(s + 1)}{s^2 + 3s + 2} \end{aligned}$$

$$a(s + 2) + b(s + 1) = 2s + 3$$

$$a + b = 2$$

$$2a + b = 3$$

$$\rightarrow a = 1, \quad b = 1$$

$$H(s) = \frac{1}{s + 1} + \frac{1}{s + 2}$$

Differentiation Theorem

$$\mathcal{L}f(t) = \int_0^\infty e^{-st} f(t) dt$$

$$\begin{aligned}\mathcal{L}f'(t) &= \lim_{a \rightarrow \infty} \int_0^a e^{-st} f'(t) dt \\ &= \lim_{a \rightarrow \infty} \left[e^{-st} f(t)|_0^a + s \int_0^a e^{-st} f(t) dt \right] \\ &= \lim_{a \rightarrow \infty} \left[e^{-sa} f(a) - f(0) + s \int_0^a e^{-st} f(t) dt \right] \\ &= s\mathcal{L}f(t) - f(0)\end{aligned}$$

Linear Systems

- Differentiation and delay both linear operators

Differentiation:

$$\frac{d}{dt}(a(t) + b(t)) = \frac{d}{dt}a(t) + \frac{d}{dt}b(t)$$

Delay: if

$$\begin{aligned}c(t) &= a(t) + b(t) \\ c(t - t_0) &= a(t - t_0) + b(t - t_0)\end{aligned}$$

- Exponentials are eigenfunctions of linear systems

in general,

$$a \cdot x(t) + b \cdot \frac{d}{dt}x(t) + c \cdot x(t - t_0) \neq d \cdot x(t)$$

but if $x(t) = Ae^{(\sigma+j\omega)t}$, the equality holds:

$$\frac{d}{dt}Ae^{(\sigma+j\omega)t} = (\sigma + j\omega) \cdot Ae^{(\sigma+j\omega)t}$$
$$Ae^{(\sigma+j\omega)(t-t_0)} = e^{(\sigma+j\omega)(-t_0)} \cdot Ae^{(\sigma+j\omega)t}$$

For Exponentials:

Any sum of differentiated and time-displaced signals is \propto to the original signal.

Question:

How to arrange exponentials on s-plane or z-plane? Where do $e^{j\omega t}$ (freq axis) go?

s-plane: 's' is differentiator

$$\frac{d}{dt}e^{j\omega t} = j\omega e^{j\omega t}$$

so if $s = j\omega$, it will 'differentiate' a pure exponential: the exponential that 'lives' at point s on the plane is e^{st}

- $s = j\omega \rightarrow$ frequency axis becomes imaginary axis

z-plane: ' z^{-1} ' is the unit delay operator

$$e^{j\omega(t-T)} = e^{-j\omega T} \cdot e^{j\omega t}$$

so if $z = e^{j\omega T}$, it will 'delay' a pure exponential: the exponential that 'lives' at point z on the plane is $e^{(\ln(z^{2\pi\omega t}) + jn2\pi)}$.

- frequency axis becomes unit circle (with aliasing)

Example: Where is DC on the s - and z - planes?

s-plane: d/dt of constant = $0 \cdot$ constant, so $s = 0 \rightarrow$ DC.

z-plane: delayed constant = $1 \cdot$ constant, so $z = 1 \rightarrow$ DC.

Question: Why is z^{-1} the delay operator, rather than z ?

- makes poles/zeros easier to plot
 - for causal, stable systems, poles will be within unit circle
 - for minimum phase causal systems, zeros will be within unit circle

Mapping the s -plane to the z -plane:

Bilinear Transform: if s is the differentiation operator, then $1/s$ is the integration operator.

$$y = \int x dt$$

Discrete-time calculation using 'trapezoidal' integration:

$$y(n) = y(n-1) + \frac{x(n) + x(n-1)}{2} \cdot T$$

$$y(n) - y(n-1) = x(n) + x(n-1) \cdot \frac{T}{2}$$

$$Y \cdot (1 - z^{-1}) = X \cdot \frac{T}{2} \cdot (1 + z^{-1})$$

$$\frac{1}{s} \rightarrow \frac{Y}{X} = \left(\frac{T}{2} \right) \frac{1 + z^{-1}}{1 - z^{-1}}$$

$$s \rightarrow \left(\frac{2}{T} \right) \frac{1 - z^{-1}}{1 + z^{-1}}$$

Impulse Invariance Transform: make the discrete-time filter impulse response equal to the sampled impulse response of the continuous-time filter.

for parallel single-pole systems:

$$H(s) = \sum \frac{A_k}{s - s_k}$$

$$H(z) = \sum \frac{T_d A_k}{1 - e^{s_k T_d} \cdot z^{-1}}$$

- pole at s_k in s -plane moves to $e^{s_k T_d}$ in the z -plane, for sampling interval T_d .
- preserves order
- must do partial fraction expansion to yield impulse-invariant response.
- frequency response is aliased, since impulse response of unknown bandwidth is sampled at interval T_d .

For higher-order transfer functions

$$H(s) = \frac{b_n s^n + \dots + b_0}{a_n s^n + \dots + a_0}$$

doing the same mapping $s = j\omega \rightarrow z = e^{j\omega}$ is called the 'matched- z transform' and is not impulse invariant.

18. Bode Plots; Peak and Shelf Filters

Linear Systems, Continued

Bode Plots:

- Find $|H(s)|$, $\angle H(s)$

$$|H(s)| = \frac{\prod |s - z_i|}{\prod |s - p_i|}$$

for frequency response, $s = j\omega$.

example:

$$H(s) = \frac{a}{s + a}$$

- at $\omega = a$:

$$|H| = \left| \frac{a}{(j+1)a} \right| = \left| \frac{1}{1+j} \right| = \frac{\sqrt{2}}{2} = -3dB$$

- one octave either side of a :

$$|H| = \frac{4}{\sqrt{5}} \approx -1dB$$

- for $\omega \gg a$,

$$|H| \approx \frac{a}{\omega}$$

for $\angle H$:

$$\angle H(s) = \sum \angle(s - z_i) - \sum \angle(s - p_i)$$

if $H(s) = \frac{a}{s+a}$,

- for $\omega \ll a$:

$$\angle H(j\omega) = 0$$

- for $\omega = a$:

$$\angle H(j\omega) = -45^\circ$$

- for $\omega \gg a$:

$$\angle H(j\omega) = -90^\circ$$

- an asymptotic line plotted over log of frequency which goes through -45 degrees at $\omega = a$ and hits the -90 and 0 degree asymptotes one decade above and below a will be within 6 degrees of $\angle H(j\omega)$.

Bode Plots for systems with multiple poles and zeros:

- for log-log magnitude, contributions from separate poles and zeros “add”.
- for phase, contributions from separate poles and zeros also add.

Bode Plot Example #1

$$H(s) = \frac{100}{s + 100}$$

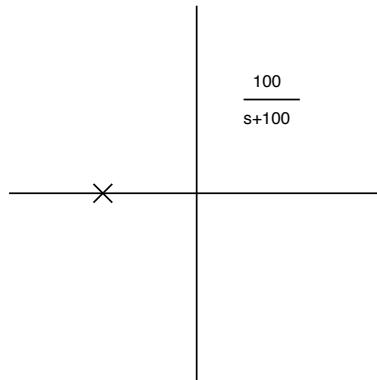


Figure 82: Pole-Zero Plot.

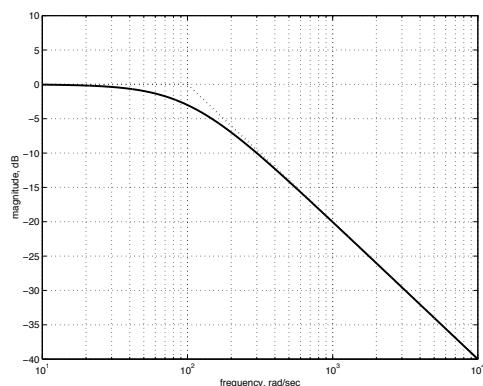


Figure 83: Magnitude Plot.

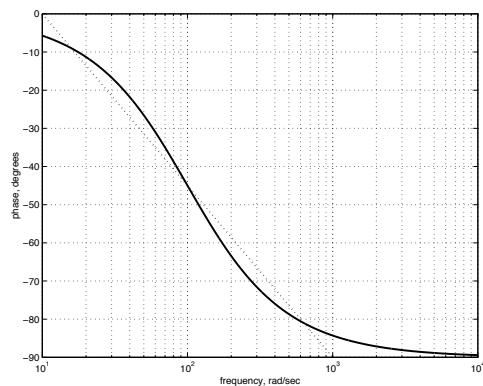


Figure 84: Phase Plot.

Bode Plot Example #2

$$H(s) = \frac{s}{s + 100}$$

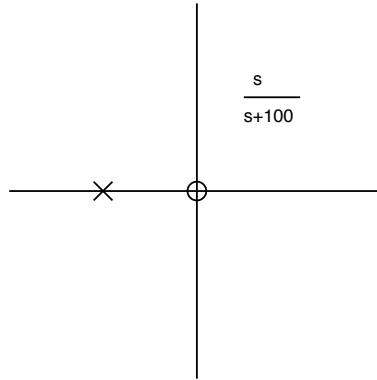


Figure 85: Pole-Zero Plot.

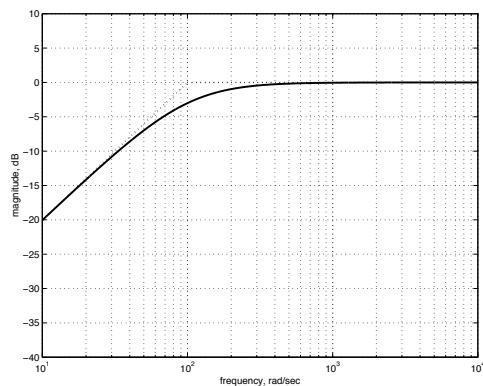


Figure 86: Magnitude Plot.

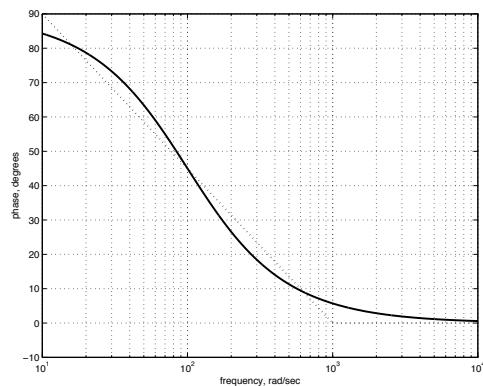
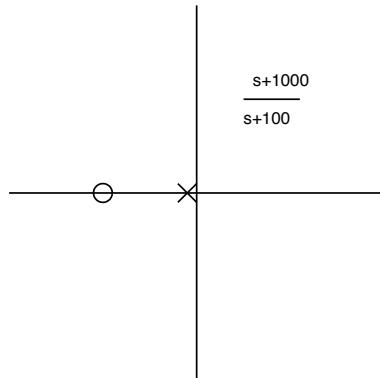
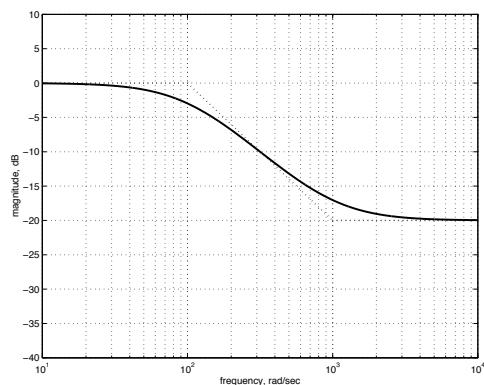
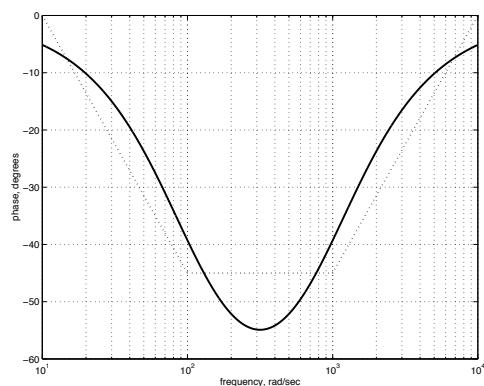


Figure 87: Phase Plot.

Bode Plot: First-Order Shelving Filter

$$H(s) = 0.1 \cdot \frac{s + 1000}{s + 100}$$

Figure 88: *Pole-Zero Plot.*Figure 89: *Magnitude Plot.*Figure 90: *Phase Plot.*

Bode Plot: Second-Order Shelving Filter

$$H(s) = 0.01 \cdot \frac{(s + 1000)^2}{(s + 100)^2}$$

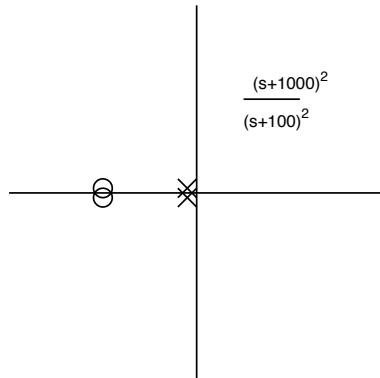


Figure 91: Pole-Zero Plot.

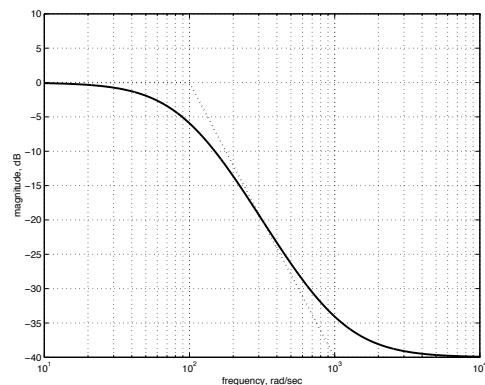


Figure 92: Magnitude Plot.

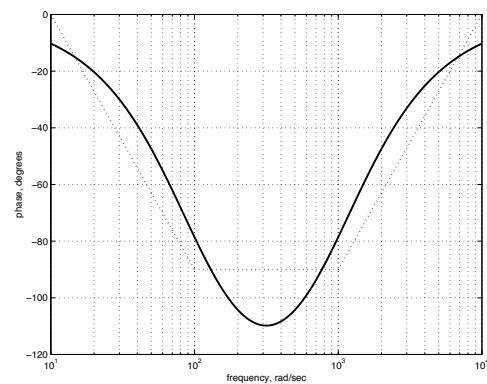


Figure 93: Phase Plot.

Circuit Properties:

- RL or RC circuits are referred to as 'relaxation' circuits. All poles are on the real axis in the left-half-plane.
- LC circuits are passive, lossless, and marginally stable. All poles are on the $j\omega$ axis.

note: 'passive' indicates that the net energy which can be extracted from the system cannot exceed a finite bound, which depends on the initial conditions of the system.

'stable' indicates either that a bounded input always produces a bounded output, or that the effects at the output of small perturbations on the input will remain small, or both.

'lossless' indicates that the system does not absorb energy. The output energy is equal to the input energy.

- for circuits with real-valued elements, all poles and zeros are either real, or occur in complex-conjugate pairs.

$$(s + a + bj)(s + a - bj) = s^2 + 2as + (a^2 + b^2)$$

Biquad Filters:

$$H(s) = \frac{b_2 s^2 + b_1 s + b_0}{a_2 s^2 + a_1 s + a_0}$$

- numerator and denominator polynomials are both quadractic, hence 'biquad'
- system has two poles and two zeros: for real coefficients, these can be complex-conjugate pairs, or real-valued.
- any real-coefficient transfer function $Y(s)$ can be realized by a cascade of real-coefficient biquads, $\prod B_n(s)$.

Example: Resonant peaking filter

$$H(s) = \frac{\frac{s}{\omega_0}}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q} \left(\frac{s}{\omega_0}\right) + 1}$$

- zero at $s = 0$
- for $Q > 0.5$, complex-conjugate poles, with real part at $-\omega_0/2Q$, and radius ω_0 . Stable for positive Q.

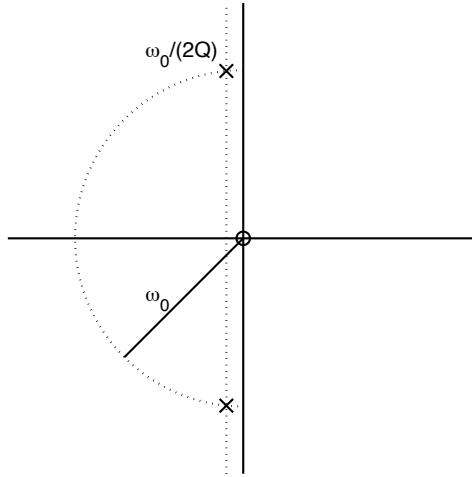


Figure 94: Pole-Zero Plot.

- for $\omega \ll \omega_0$, $H(j\omega) \approx \frac{j\omega}{\omega_0}$.
- for $\omega \gg \omega_0$, $H(j\omega) \approx \frac{\omega_0}{j\omega}$.
- asymptotic gains intersect at $\omega = \omega_0$, at zero dB. Gain of system at $\omega = \omega_0$ is equal to Q.
- in log-log space, transfer function is symmetric about $\omega = \omega_0$.

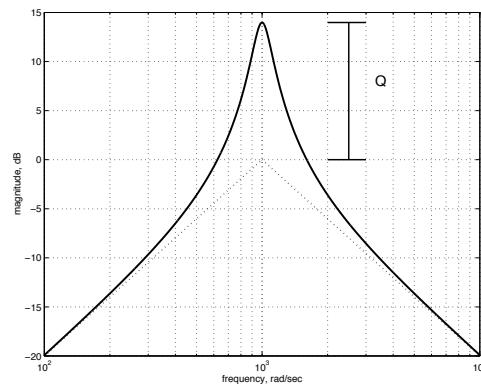


Figure 95: Magnitude Plot.

- asymptotic low-frequency phase is 90 degrees.
- asymptotic high-frequency phase is -90 degrees.

- phase is zero at $\omega = \omega_0$.

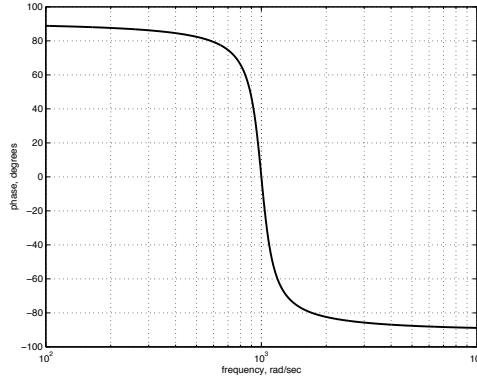


Figure 96: Phase Plot.

- where gain is 3dB below peak,

$$\frac{\omega}{\omega_0} = -\frac{1}{2Q} + \sqrt{\frac{1}{4Q^2} + 1}$$

for $\omega < \omega_0$. Phase here is 45 degrees.

- for large Q , bandwidth between half-power (-3dB) points is ω_0/Q .

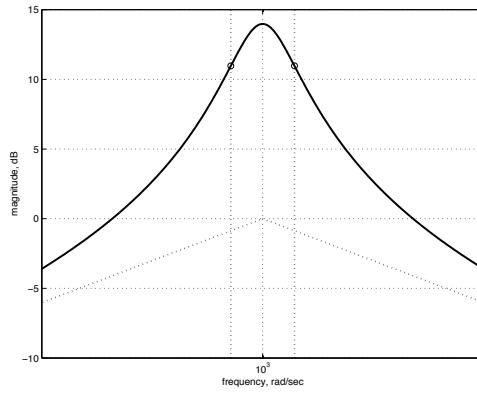


Figure 97: Magnitude Plot.

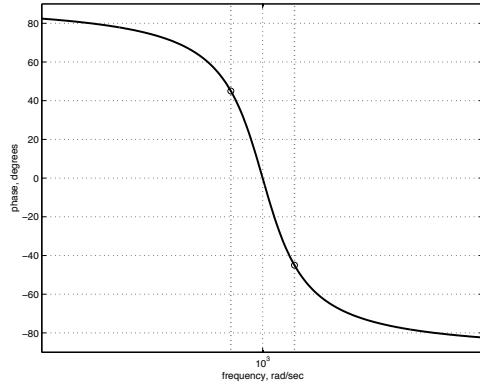


Figure 98: Phase Plot.

Example: Resonant high-cut filter

- similar to peaking filter, but with no zeros at DC

$$H(s) = \frac{1}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q} \left(\frac{s}{\omega_0}\right) + 1}$$

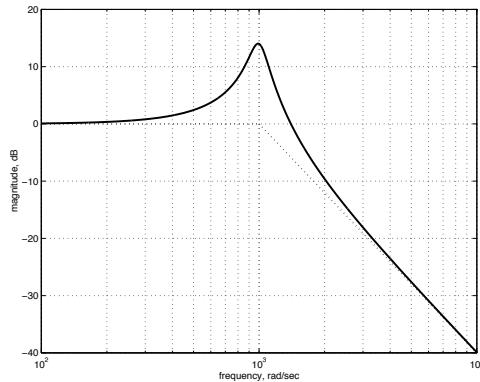
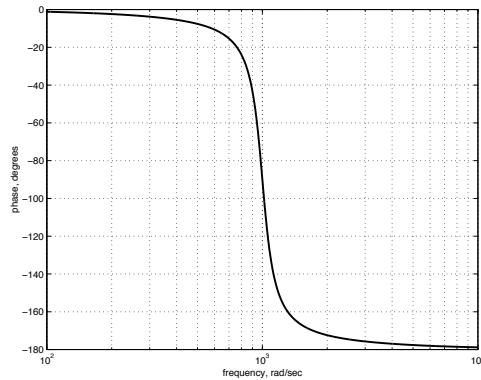
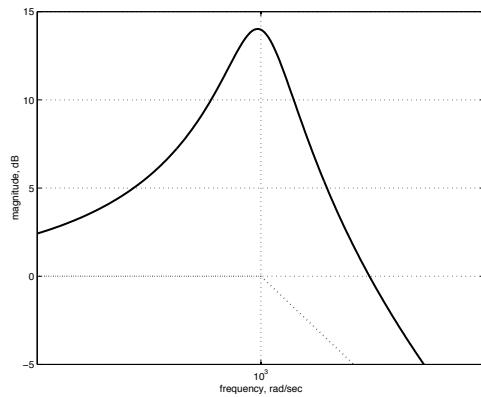
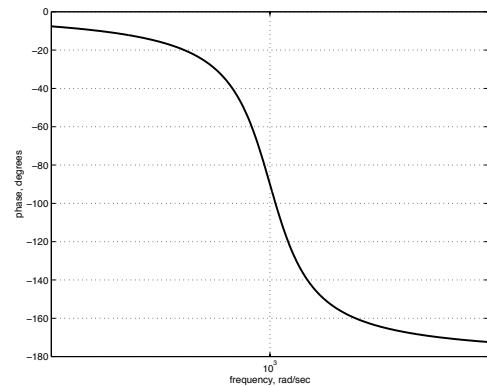


Figure 99: Magnitude Plot.

Figure 100: *Phase Plot.*

- by deleting zero, symmetry of magnitude transfer function is destroyed.
- peak is no longer exactly at ω_0
- phase is exactly 90° at ω_0

Figure 101: *Magnitude Plot.*

Figure 102: *Phase Plot.*

Eigenfunctions of LTI systems:

recall the convolution integral:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)d\tau$$

let $x = e^{st}$:

$$y(t) = \int_{-\infty}^{\infty} h(\tau)e^{s(t-\tau)}d\tau = \left[\int_{-\infty}^{\infty} h(\tau)e^{-s\tau}d\tau \right] e^{st} = H(s)e^{st},$$

since

$$H(s) = \int_{-\infty}^{\infty} h(t)e^{-st}dt$$

so e^{st} is an eigenfunction of the system, with $H(s)$ as the eigenvalue.

19. Parametric Sections and Shelf Filters

Parametric Sections

For “boost” (gain $\gamma > 1$)

$$H(s) = \frac{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q} \cdot \gamma \left(\frac{s}{\omega_0}\right) + 1}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q} \left(\frac{s}{\omega_0}\right) + 1} \quad (1)$$

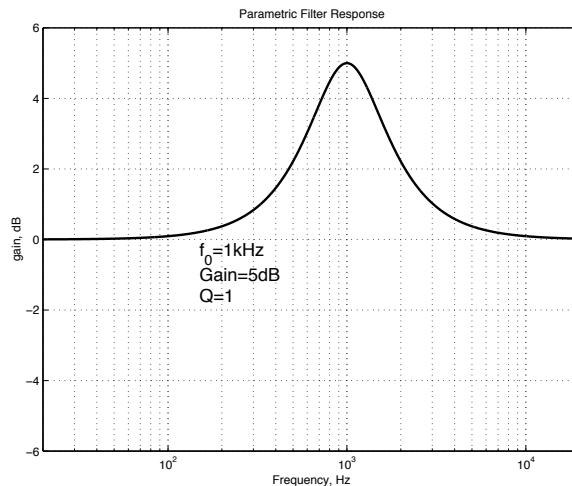


Figure 103: *Parametric Filter*.

1. magnitude symmetric about ω_0
 - (a) both numerator and denominator are symmetric
2. gain = γ at ω_0
3. phase is odd about ω_0

4. phase = 0 at ω_0

For “cut” ($\gamma < 1$), cannot use same transfer function:

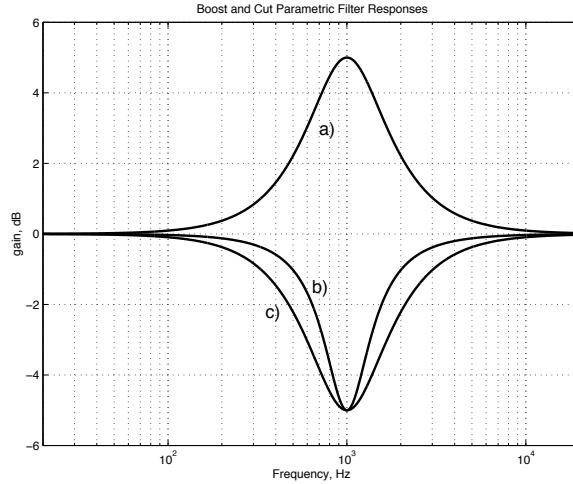


Figure 104: *Parametric Filters.*

use of equation 103 leads to curve b), which has a different bandwidth than a). Curve c) generated by

$$H(s) = \frac{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q} \left(\frac{s}{\omega_0}\right) + 1}{\left(\frac{s}{\omega_0}\right)^2 + \frac{1}{Q \cdot \gamma} \left(\frac{s}{\omega_0}\right) + 1} \quad (2)$$

this is the reciprocal of eq. 103.

1. Q for parametric section defined as higher of numerator and denominator Q

Discrete-Time Signals / Systems

Z-transform for $x(n)$:

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n}$$

for $z = e^{j\omega}$,

$X(z)$ is the Fourier transform of $x(n)$.

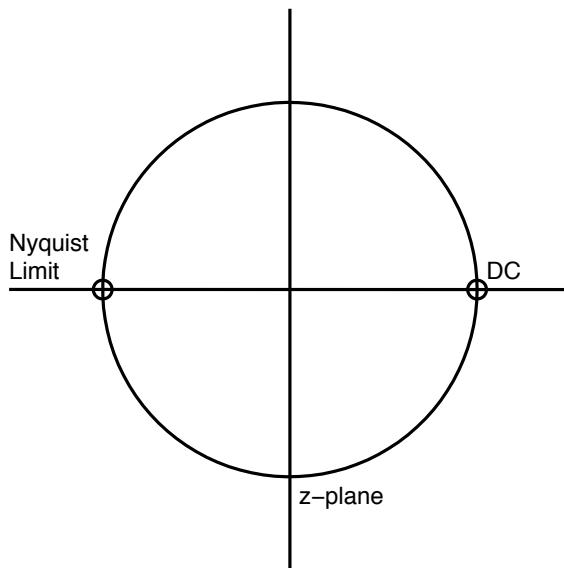


Figure 105: The *z*-plane.

z^{-1} is the unit delay.

Direct-Form Systems

The discrete-time biquad:

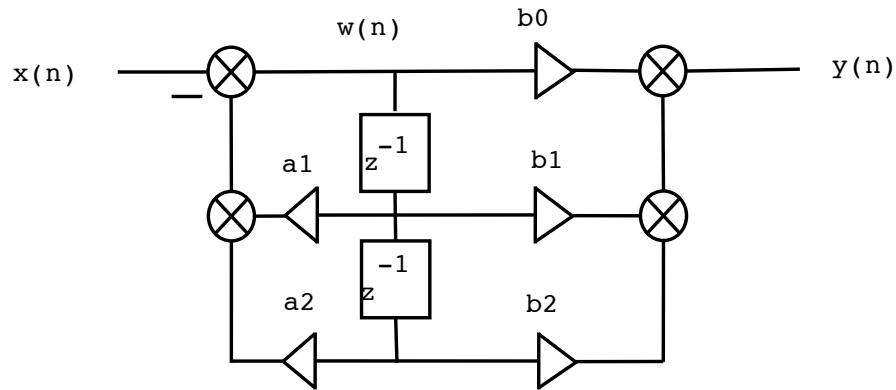


Figure 106: The *z*-plane.

$$w(n) = x(n) - a_1 w(n-1) - a_2 w(n-2)$$

$$y(n) = b_0 w(n) + b_1 w(n-1) + b_2 w(n-2)$$

$$y(n) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} x(n)$$

Filter Discretization

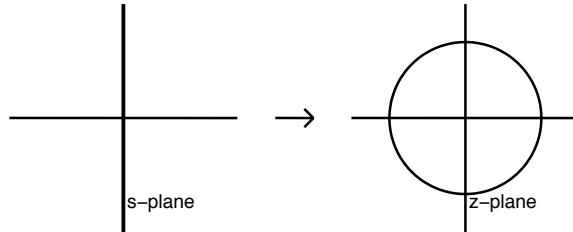


Figure 107: the *s*- and *z*-planes.

1. need mapping from s to z
2. discrete-time model is desired
 - (a) preserve filter order
 - (b) preserve stability
 - (c) retain 'salient features' of filter

Bilinear transform:

$$s = \frac{2}{T_d} \left(\frac{1 - z^{-1}}{1 + z^{-1}} \right)$$

or,

$$z = \frac{1 + \frac{T_d}{2}s}{1 - \frac{T_d}{2}s}$$

for continuous-time frequency Ω and discrete-time frequency ω ,
maps $s = j\Omega$ onto $z = e^{j\omega}$:

$$z = \frac{1 + \frac{T_d}{2}j\omega}{1 - \frac{T_d}{2}j\omega} \rightarrow |z| = 1$$

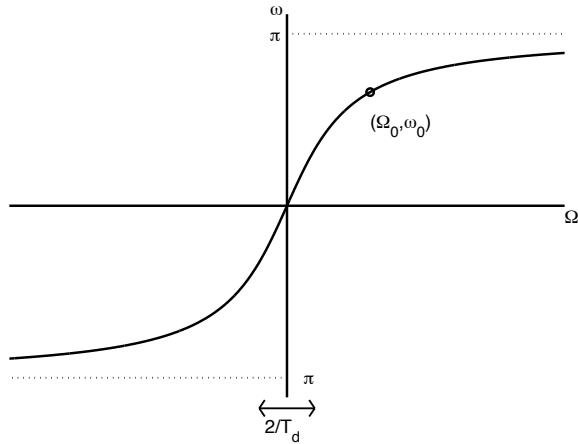


Figure 108: Frequency mapping.

$$\Omega = \frac{2}{T_d} \tan\left(\frac{\omega}{2}\right),$$

or

$$\omega = 2 \tan^{-1}\left(\frac{\Omega T_d}{2}\right)$$

can “match” a given frequency $\Omega_0 \rightarrow \omega_0$

$$\tan\left(\frac{\omega_0}{2}\right) = \Omega_0 \cdot \frac{T_d}{2}$$

$$T_d = \frac{2}{\Omega_0} \tan\left(\frac{\omega_0}{2}\right)$$

1. gain “features” of transfer function preserved
2. frequencies warped e.g. phase linearity not preserved

as $\Omega_0, \omega_0 \rightarrow 0, T_d \rightarrow 1/f_s$ (makes slope = 1 at DC)

Methods to minimize warping

1. pick f_s such that Nyquist limit is well above any features of interest

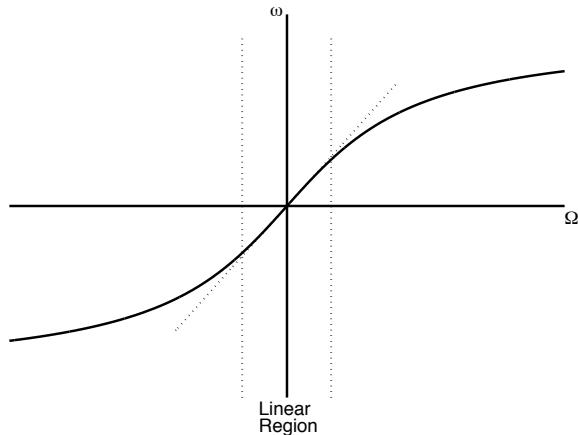


Figure 109: Frequency mapping.

2. prewarp continuous-time transfer function to counter effects of bilinear transform (2 papers)

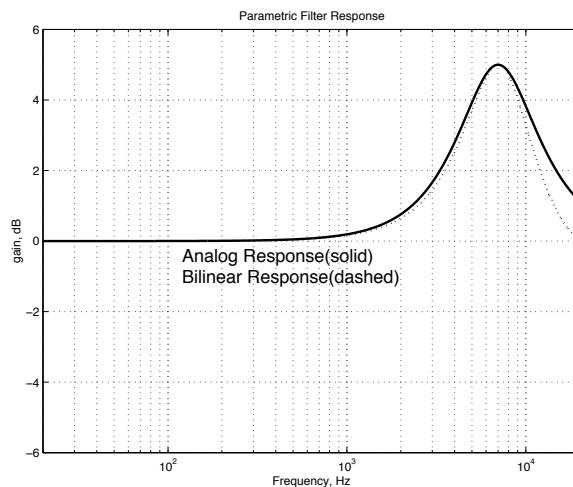


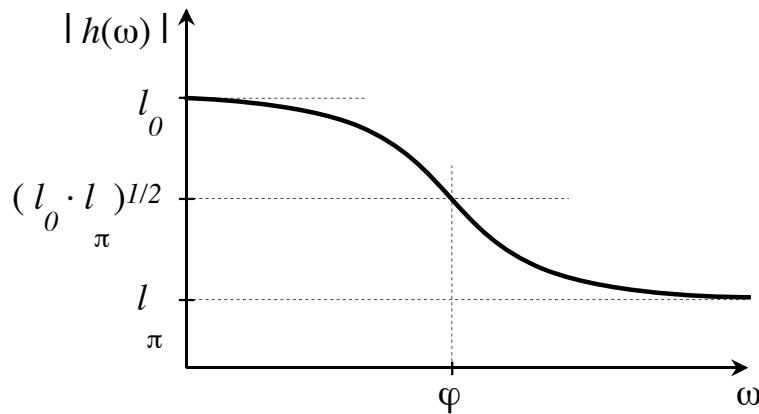
Figure 110: Continuous- and discrete-time parametric filters.

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}$$

1. five coefficients give five degrees of freedom
 - (a) gain at DC
 - (b) gain at Nyquist
 - (c) location of peak
 - (d) gain at peak
 - (e) bandwidth

20. Digital Peaking and Shelving Filters, Graphic Equalizer

Peaking and Shelving Filter Properties



Shelf Filter Specification

- A *shelf filter* $h(\omega)$ takes on a gain of ℓ_0 at DC, a gain of ℓ_π at high frequencies, and a gain of $(\ell_0 \cdot \ell_\pi)^{\frac{1}{2}}$ at a transition frequency φ .
- A *low shelf filter* has the high-frequency gain ℓ_π fixed at one, with ℓ_0 free to vary.
- In a *high shelf filter*, the high-frequency gain ℓ_π is varied with the DC gain ℓ_0 fixed at one.

First-Order Shelf Filter Analog Prototype

- Consider the following first-order analog prototype filter,

$$h(s) = \frac{\ell_\pi \cdot s / \rho + \ell_0}{s / \rho + 1}.$$

- Note that the analog prototype filter takes on the desired DC and high-frequency gains,

$$h(\omega = 0) = \ell_0, \quad h(\omega \rightarrow \infty) = \ell_\pi,$$

with the factor ρ controlling the transition frequency.

- The analog prototype filter is designed to have a transition frequency of one. Setting $|h(j \cdot 1)|^2 = \ell_0 \cdot \ell_\pi$, the factor ρ may be specified.

$$|h(j \cdot 1)|^2 = \ell_0 \cdot \ell_\pi = \frac{\ell_\pi^2 / \rho^2 + \ell_0^2}{1 / \rho^2 + 1},$$

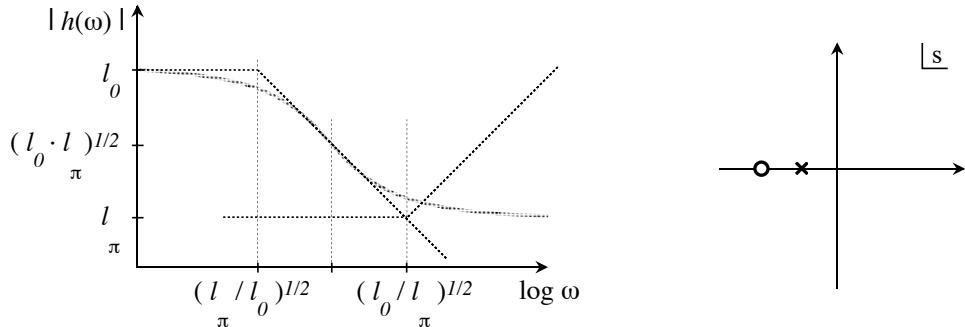
which implies

$$\rho = \left(\frac{\ell_\pi}{\ell_0} \right)^{\frac{1}{2}}.$$

- Substituting into the expression for $h(s)$, we have

$$h(s) = (\ell_0 \ell_\pi)^{\frac{1}{2}} \cdot \frac{s + (\ell_0 / \ell_\pi)^{\frac{1}{2}}}{(\ell_0 / \ell_\pi)^{\frac{1}{2}} s + 1}.$$

First-Order Shelf Filter Analog Prototype Properties



- The analog prototype shelf filter has transfer function

$$h(s) = (\ell_0 \ell_\pi)^{\frac{1}{2}} \cdot \frac{s + (\ell_0/\ell_\pi)^{\frac{1}{2}}}{(\ell_0/\ell_\pi)^{\frac{1}{2}} s + 1}.$$

- It has a real pole and real zero at reciprocal frequencies about the transition frequency $s = j \cdot 1$,

$$s_p = -(\ell_\pi/\ell_0)^{\frac{1}{2}}, \quad s_z = -(\ell_0/\ell_\pi)^{\frac{1}{2}}.$$

- Note that on a dB scale the normalized analog shelf filter is antisymmetric in $\log s$ about the transition frequency,

$$\begin{aligned} h(1/s) &\propto \frac{1/s + (\ell_0/\ell_\pi)^{\frac{1}{2}}}{(\ell_0/\ell_\pi)^{\frac{1}{2}}/s + 1} \\ &= \frac{1 + (\ell_0/\ell_\pi)^{\frac{1}{2}}s}{(\ell_0/\ell_\pi)^{\frac{1}{2}} + s} \propto 1/h(s). \end{aligned}$$

First-Order Digital Shelf Filter via Bilinear Transform

- The digital shelf filter is formed via bilinear transform on the analog prototype,

$$h(z) = h(s(z)), \quad s = \frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}}.$$

We have

$$h(s) = (\ell_0 \ell_\pi)^{\frac{1}{2}} \cdot \frac{s + 1/\rho}{s/\rho + 1}, \quad \rho = (\ell_\pi / \ell_0)^{\frac{1}{2}},$$

and,

$$\begin{aligned} h(z) &= (\ell_0 \ell_\pi)^{\frac{1}{2}} \cdot \frac{\rho(1 - z^{-1}) + \frac{T}{2}(1 + z^{-1})}{(1 - z^{-1}) + \rho \frac{T}{2}(1 + z^{-1})}, \\ &= (\ell_0 \ell_\pi)^{\frac{1}{2}} \cdot \frac{(\frac{T}{2} + \rho) + (\frac{T}{2} - \rho)z^{-1}}{(\rho \frac{T}{2} + 1) + (\rho \frac{T}{2} - 1)z^{-1}}, \\ &= (\ell_0 \ell_\pi)^{\frac{1}{2}} \left[\frac{\rho + \frac{T}{2}}{1 + \rho \frac{T}{2}} \right] \cdot \frac{1 - \left(\frac{1-T/2\rho}{1+T/2\rho} \right) z^{-1}}{1 - \left(\frac{1-\rho T/2}{1+\rho T/2} \right) z^{-1}}. \end{aligned}$$

- Note that the pole at $-\rho$ and zero at $-1/\rho$ have been transformed according to the bilinear transform.

Selecting T

- The "sampling period" T is chosen to put the transition frequency of the analog prototype, $\Omega = 1$, at the desired frequency φ .
- Evaluating the bilinear transform on the unit circle, $z^{-1} = e^{-j\omega}$,

$$\begin{aligned}s &= \frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}} = \frac{2}{T} \cdot \frac{1 - e^{-j\omega}}{1 + e^{-j\omega}}, \\ &= \frac{2}{T} \cdot \frac{e^{j\omega/2} - e^{-j\omega/2}}{e^{j\omega/2} + e^{-j\omega/2}} = \frac{2}{T} \cdot \frac{j \sin(\omega/2)}{\cos(\omega/2)},\end{aligned}$$

we see that the frequency ω on the unit circle corresponds to the frequency

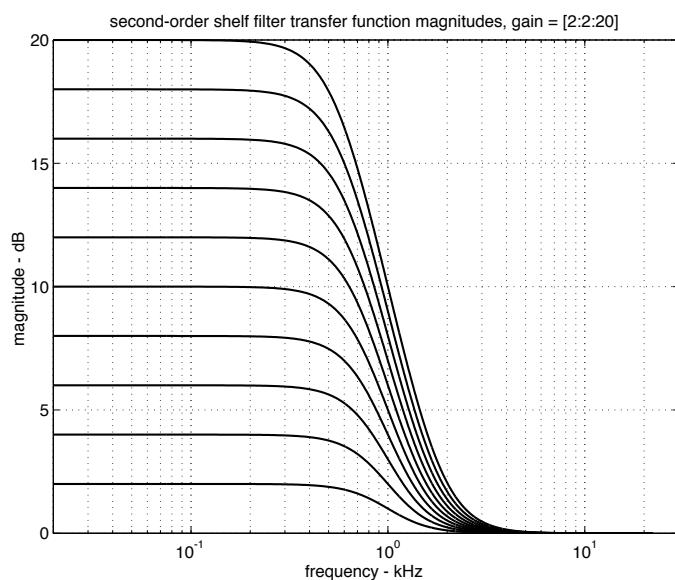
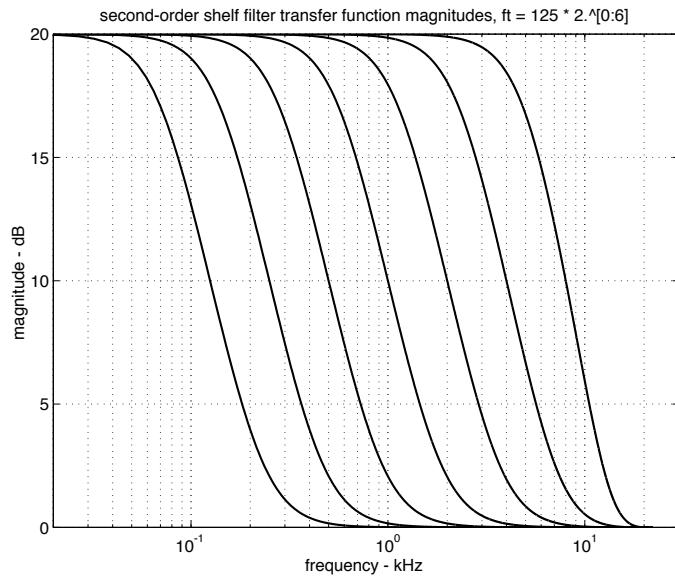
$$\Omega = \frac{2}{T} \tan(\omega/2)$$

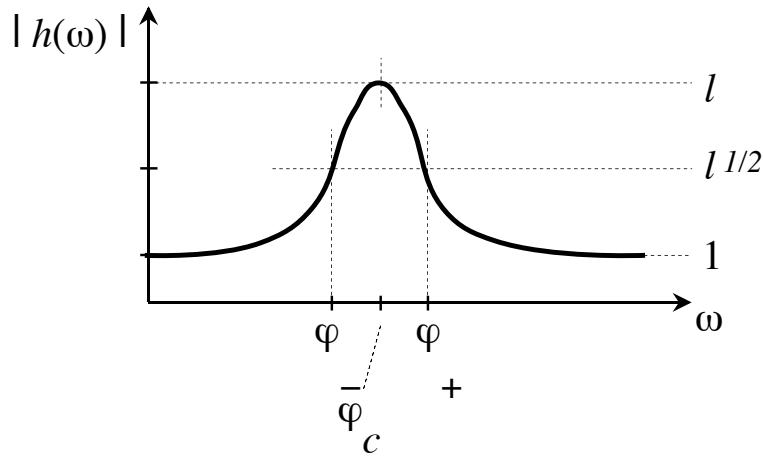
on the s -plane imaginary axis. (Note that for small frequencies $|\omega| \ll 1$, $\tan(\omega/2) \approx \omega/2$, and $\Omega \approx \frac{1}{T}$.

- To map the analog prototype filter transition frequency of 1 to the desired transition frequency φ , we set

$$\frac{T}{2} = \tan(\varphi/2).$$

Example Shelf Filters





Peak Filter Specification

- A *peak (or notch) filter* $p(\omega)$ takes on a gain of 1 at DC and high frequencies, and achieves a maximum (or minimum) gain of l at some point between the transition frequencies φ_{\pm} , at which the gain is \sqrt{l} .
- The frequency of the magnitude extremum is called the *center frequency*, φ_c .

Digital Peak Filter

The second-order digital filter

$$p(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}$$

with coefficients given by

$$\begin{aligned} a_2 &= \frac{2Q - \sin \varphi_c}{2Q + \sin \varphi_c}, \\ a_1 &= b_1 = -(1 + a_2) \cos \varphi_c, \\ b_0 &= \frac{1}{2}(1 + a_2) + \frac{1}{2}(1 - a_2)\ell, \\ b_2 &= \frac{1}{2}(1 + a_2) - \frac{1}{2}(1 - a_2)\ell, \end{aligned}$$

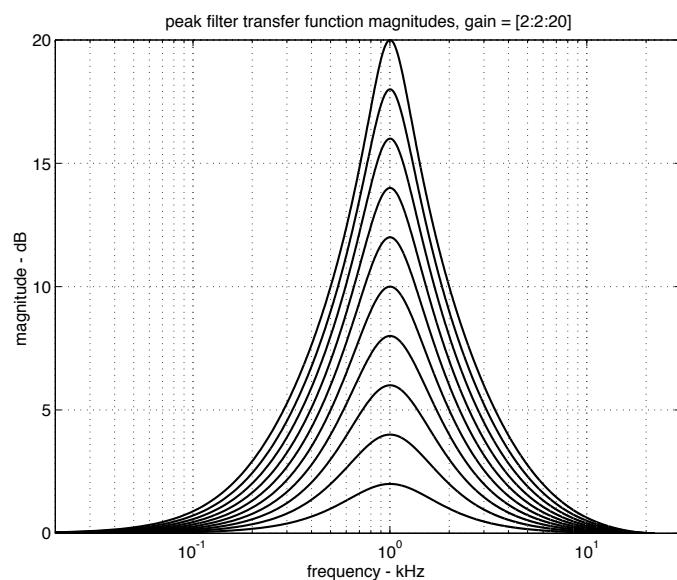
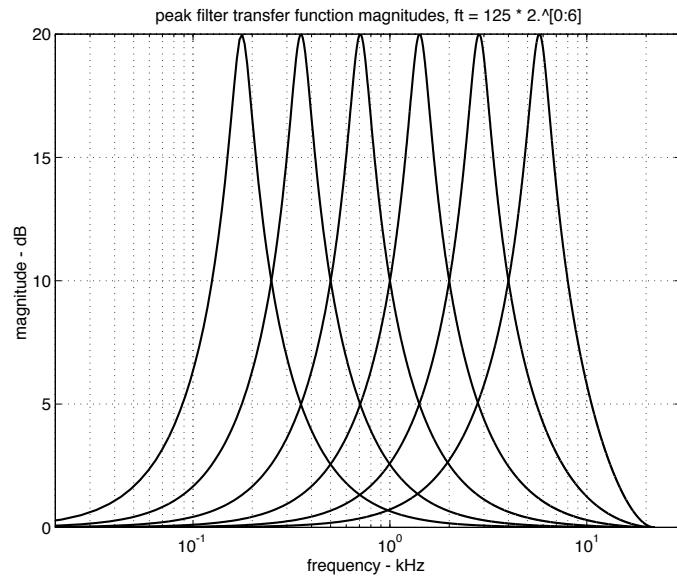
implements a peak (or notch) filter with maximum (or minimum) gain ℓ at a center frequency φ_c between the specified transition frequencies φ_{\pm} , at which the filter takes on magnitude $\sqrt{\ell}$. The center frequency φ_c and the inverse bandwidth Q may be written in terms of the transition frequencies φ_{\pm} and dB peak gain λ ,

$$\begin{aligned} \varphi_c &= \arccos \left\{ \kappa - \text{sign}\{\kappa\} (\kappa^2 - 1)^{\frac{1}{2}} \right\}, \\ \kappa &= \frac{1 + \cos \varphi_- \cos \varphi_+}{\cos \varphi_- + \cos \varphi_+} \\ Q &= \frac{1}{2} \left[\frac{\ell \cdot \sin^2 \varphi_c \cdot (\cos \varphi_- + \cos \varphi_+)}{2 \cos \varphi_c - \cos \varphi_- - \cos \varphi_+} \right]^{\frac{1}{2}}. \end{aligned}$$

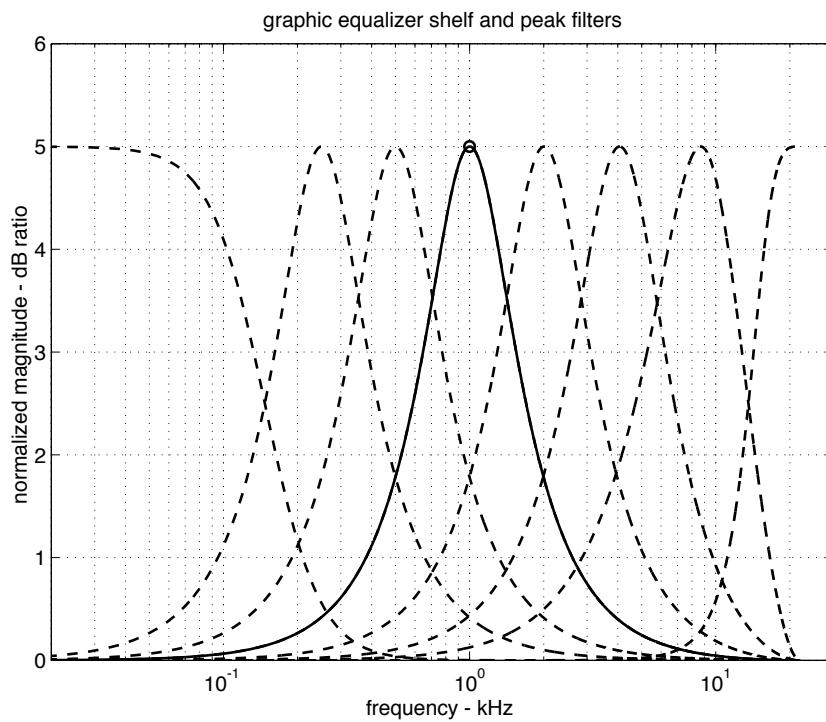
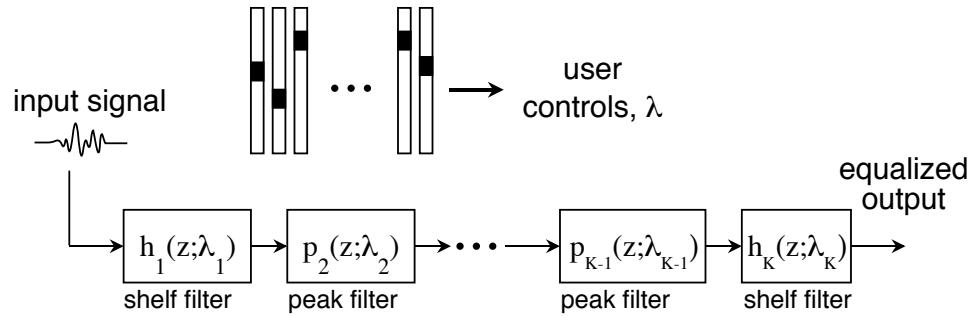
In the case that $\varphi_+ + \varphi_- = \pi$, we have

$$\begin{aligned} \varphi_c &= \pi/2, \\ Q &= \frac{\sqrt{\ell}}{2} |\cot \delta|, \quad \delta = \frac{1}{2}(\varphi_- - \varphi_+). \end{aligned}$$

Example Peak Filters

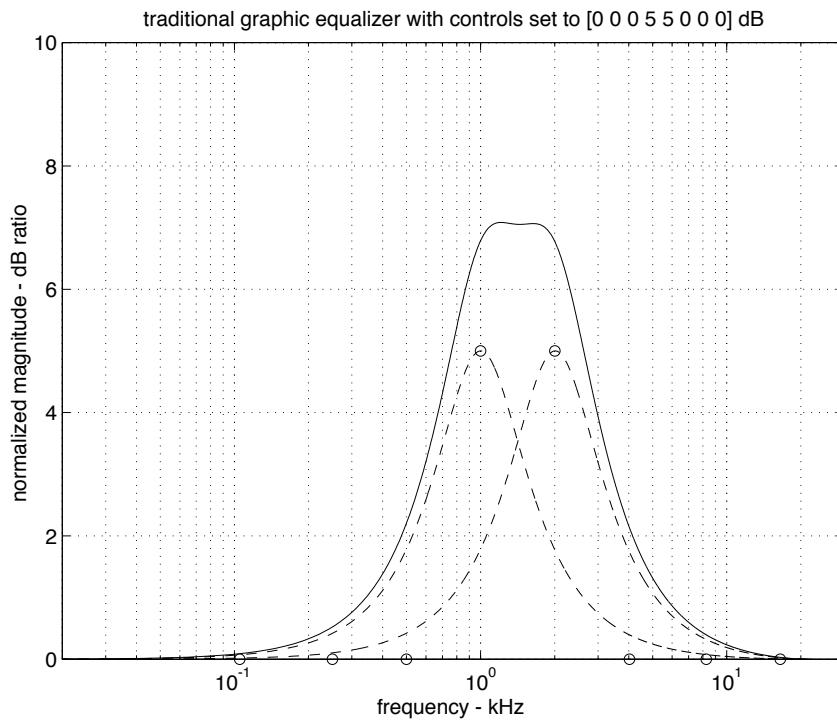


Graphic Equalizer



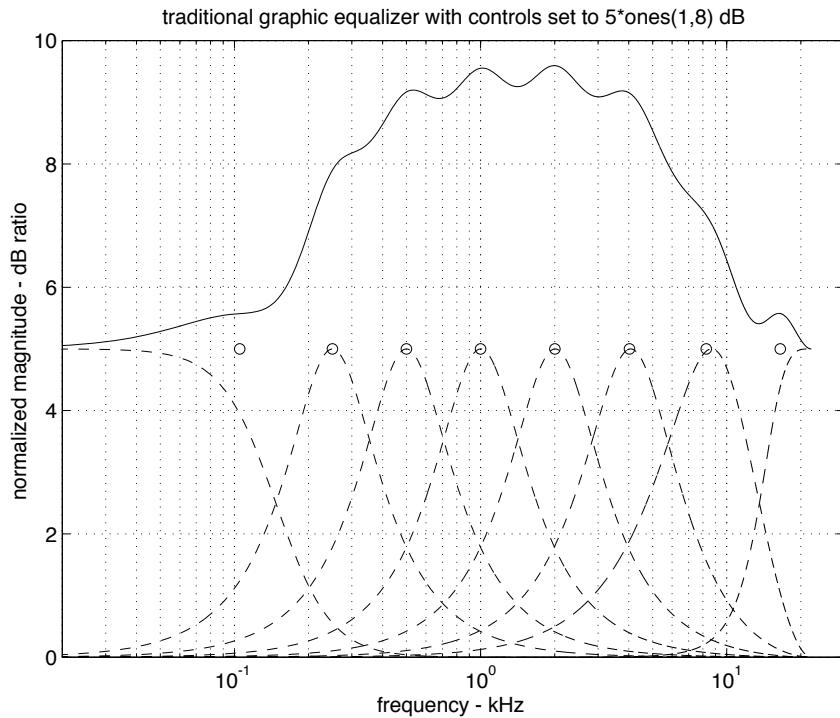
- A graphic equalizer is implemented as a cascade of peak and shelf filters having *a priori* specified bandwidths and user-controlled gains.

Graphic Equalizer Behavior



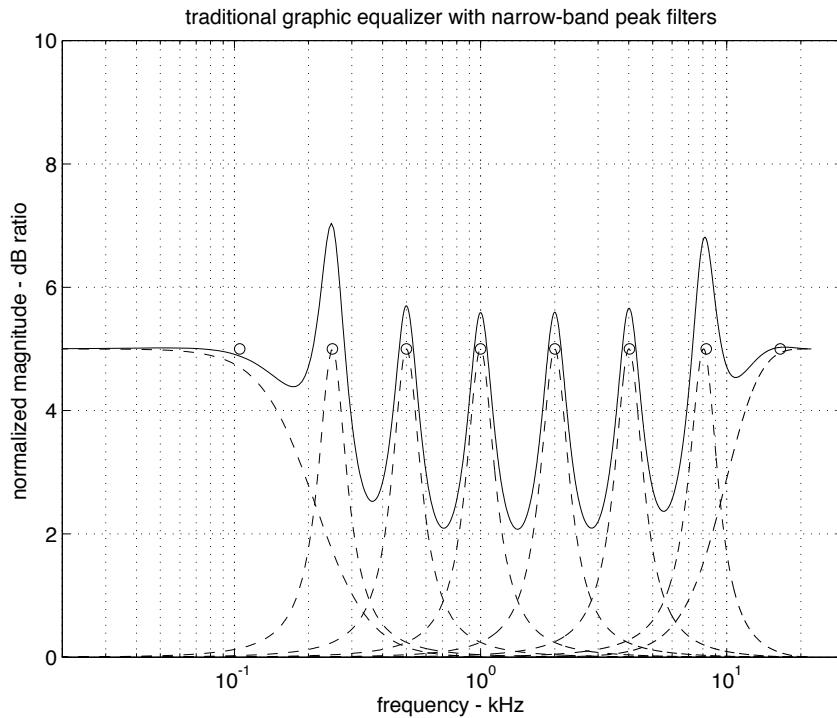
- By adjusting the gains of the peak and shelf filters in the cascade, a wide variety of transfer function magnitudes may be produced.
- Although the idea is that the transfer function magnitude should smoothly interpolate the specified gains, it doesn't always work out that way.

Graphic Equalizer Behavior



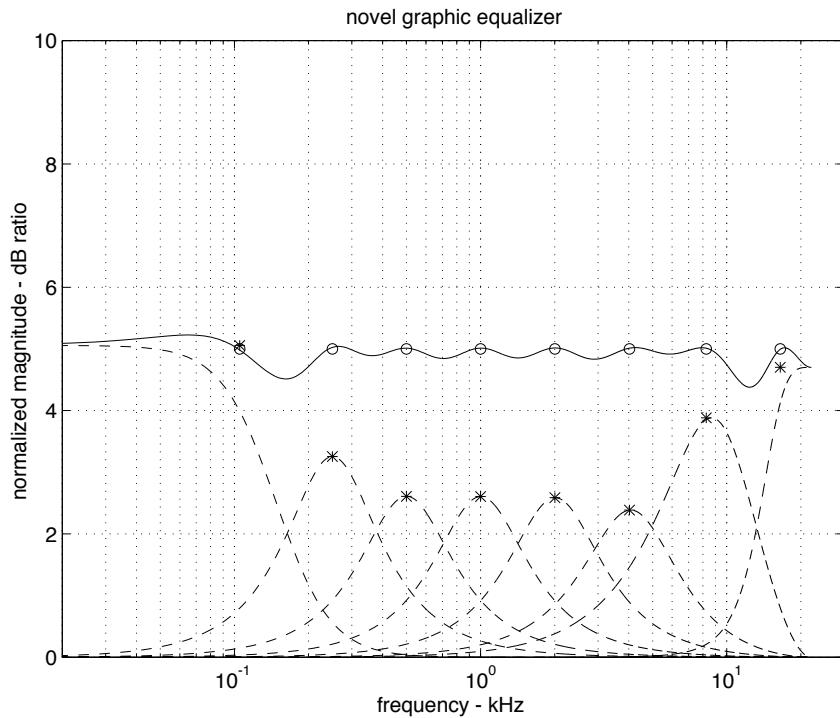
- Contributions from adjacent bands cause the filter to overshoot the desired gains.

Graphic Equalizer Behavior



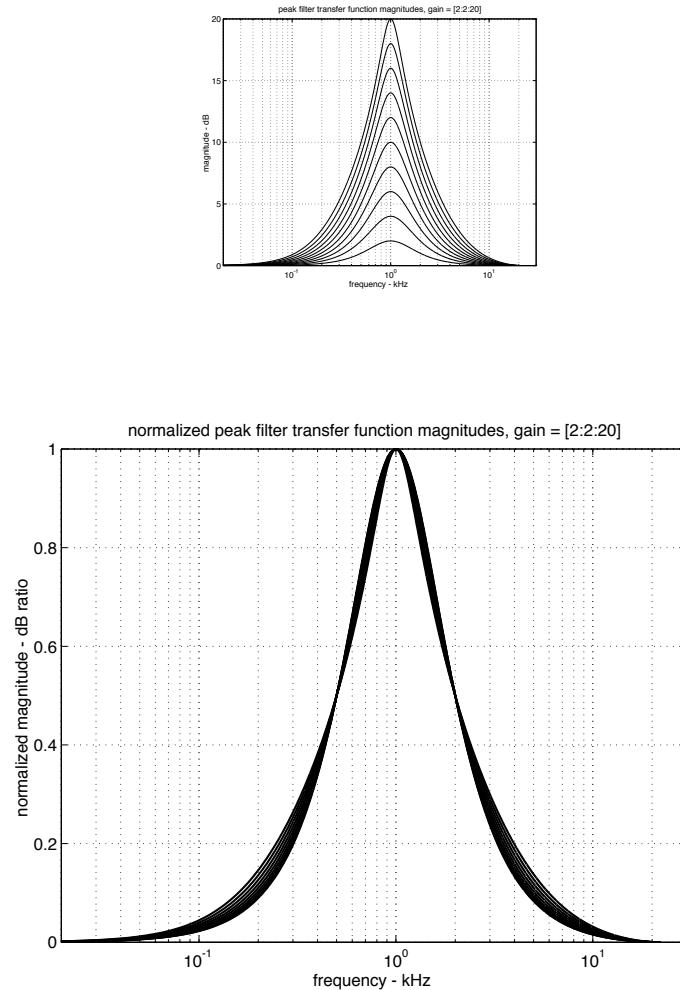
- Traditional graphic equalizers don't smoothly interpolate the given gains.
 - Small filter bandwidths lead to rippled transfer functions.
 - Large filter bandwidths overshoot desired gains.

Alternative Gain Computation



- The idea is to find a set of peak and shelf filter gains which account for the overlap between bands, so that the resulting peak and shelf filter cascade interpolates the specified band gains.

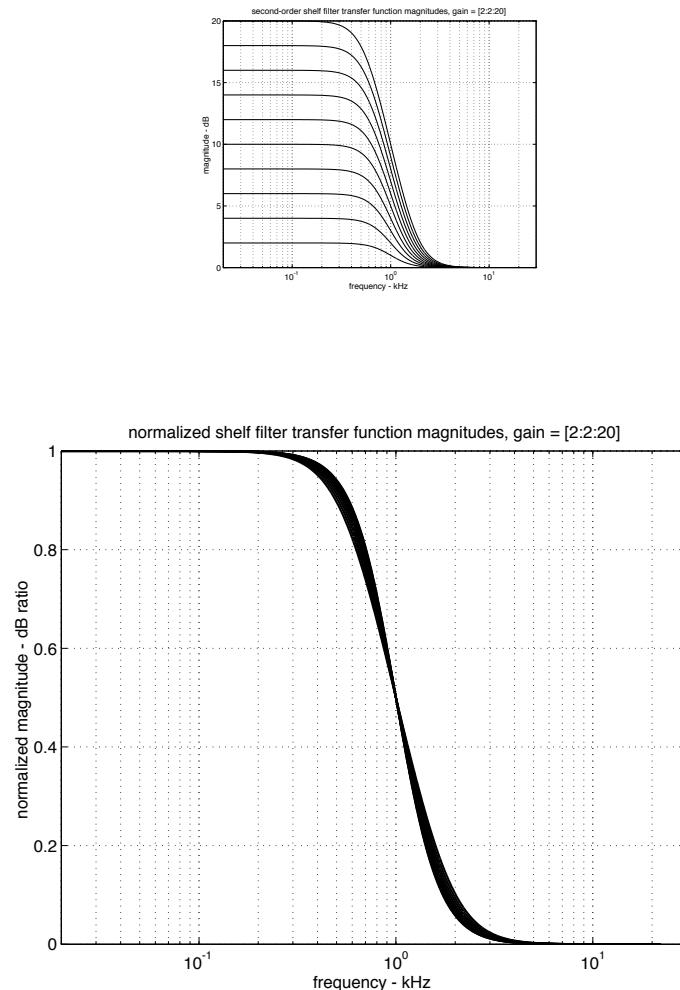
Peak Filter Self Similarity



- Peaking filters parameterized by a maximum dB gain λ , achieved somewhere between transition frequencies φ_- and φ_+ , at which the dB gain is $\lambda/2$, are approximately self similar on a log magnitude scale.

$$\alpha \cdot \log |p(\omega; \lambda, \varphi_{\pm})| \approx \log |p(\omega; \alpha \cdot \lambda, \varphi_{\pm})|.$$

Shelf Filter Self Similarity



- Low shelf filters specified by a low-frequency dB gain λ , a dB gain $\lambda/2$ at a transition frequency φ , and a high-frequency gain of one are also approximately self similar on a log magnitude scale.

$$\alpha \cdot \log |h(\omega; \lambda, \varphi)| \approx \log |h(\omega; \alpha \cdot \lambda, \varphi)|.$$

Filter Design Approach

- Consider a cascade $g(\omega; \boldsymbol{\theta})$ of K peak and shelf filters having gains and $\lambda_k, k = 1, \dots, K$ and transition frequencies $\varphi_k, k = 1, \dots, K - 1$ stacked in the column $\boldsymbol{\theta}$,

$$g(\omega; \boldsymbol{\theta}) = h(\omega; \lambda_1, \varphi_1) \cdot h(\omega; \lambda_K, \varphi_{K-1}) \cdot \prod_{k=2}^{K-1} p(\omega; \lambda_k, \varphi_{k-1}, \varphi_k).$$

- Because of the self similarity property, the dB magnitude of the cascade, denoted by $\gamma(\omega; \boldsymbol{\theta})$,

$$\begin{aligned} \gamma(\omega; \boldsymbol{\theta}) &\stackrel{\text{def}}{=} 20 \log_{10}\{g(\omega; \boldsymbol{\theta})\}, \\ &= \sigma(\omega; \lambda_1, \varphi_1) + \sigma(\omega; \lambda_K, \varphi_{K-1}) + \\ &\quad \sum_{k=2}^{K-1} \pi(\omega; \lambda_k, \varphi_{k-1}, \varphi_k), \end{aligned}$$

is approximately linear in the filter gains,

$$\begin{aligned} \gamma(\omega; \boldsymbol{\theta}) &\approx \lambda_1 \cdot \sigma(\omega; 1.0, \varphi_1) + \lambda_K \cdot \sigma(\omega; 1.0, \varphi_{K-1}) + \\ &\quad \sum_{k=2}^{K-1} \lambda_k \cdot \pi(\omega; 1.0, \varphi_{k-1}, \varphi_k). \end{aligned}$$

Filter Design Approach

- At a particular frequency ω_i , we have

$$\gamma(\omega_i; \boldsymbol{\theta}) \approx [\sigma_1(\omega_i) \ \pi_2(\omega_i) \ \cdots \ \pi_{K-1}(\omega_i) \ \sigma_K(\omega_i)] \boldsymbol{\lambda},$$

where $\boldsymbol{\lambda}$ is the stack of dB band gains,

$$\boldsymbol{\lambda} = [\lambda_1 \ \cdots \ \lambda_K]^\top,$$

and where $\sigma_k(\omega_i)$ and $\pi(\omega_i)$ are the transfer function dB magnitudes of shelf and peak filters with specified transition frequencies and 1.0 dB nominal gains, evaluated at ω_i ,

$$\sigma_k(\omega_i) = 20 \log_{10}\{h(\omega_i; 1.0\text{dB}, \varphi_k)\}$$

and

$$\pi(\omega_i) = 20 \log_{10}\{p(\omega_i; 1.0\text{dB}, \varphi_{k-1}, \varphi_k)\}.$$

- Stacking instances of $\gamma(\omega; \boldsymbol{\theta})$ evaluated at a set of frequencies ω_i to form the column $\boldsymbol{\gamma}$, we have

$$\boldsymbol{\gamma} \approx \mathbf{B}\boldsymbol{\lambda},$$

$$\mathbf{B} = [\boldsymbol{\sigma}_1 \ \boldsymbol{\pi}_2 \ \cdots \ \boldsymbol{\pi}_{K-1} \ \boldsymbol{\sigma}_K],$$

where $\boldsymbol{\sigma}_k$ and $\boldsymbol{\pi}_k$ are columns of dB magnitudes evaluated at ω_i of shelf and peak filters having gains of 1.0 dB and specified transition frequencies.

Graphic Equalizer Design

- The self similarity of second-order peaking and shelving filters leads to an approximate linear relationship between the dB filter gains and the dB gain of the cascade,

$$\gamma \approx \mathbf{B}\lambda,$$

- Therefore, given a set of shelf and peak filters having specified transition frequencies, and positive definite weighting matrix \mathbf{W} , the gains

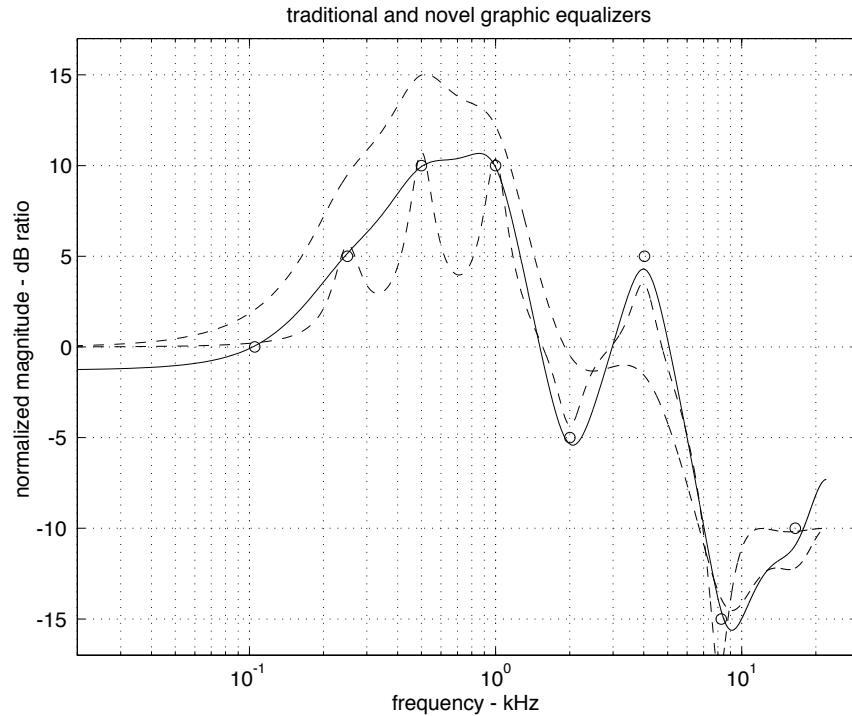
$$\hat{\lambda} = (\mathbf{B}^\top \mathbf{W} \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{W} \gamma$$

will approximately minimize the weighted square difference between a desired dB magnitude response γ and the shelf and peak filter cascade dB magnitude at frequencies ω_i , γ .

- For a graphic equalizer with $K - 1$ fixed band edges, the frequencies ω_i can be chosen as the K band centers, and the gains $\hat{\lambda}$ simply computed as the control gains γ scaled by the basis inverse,

$$\hat{\lambda} = \mathbf{B}^{-1} \gamma.$$

Graphic Equalizer Design Example

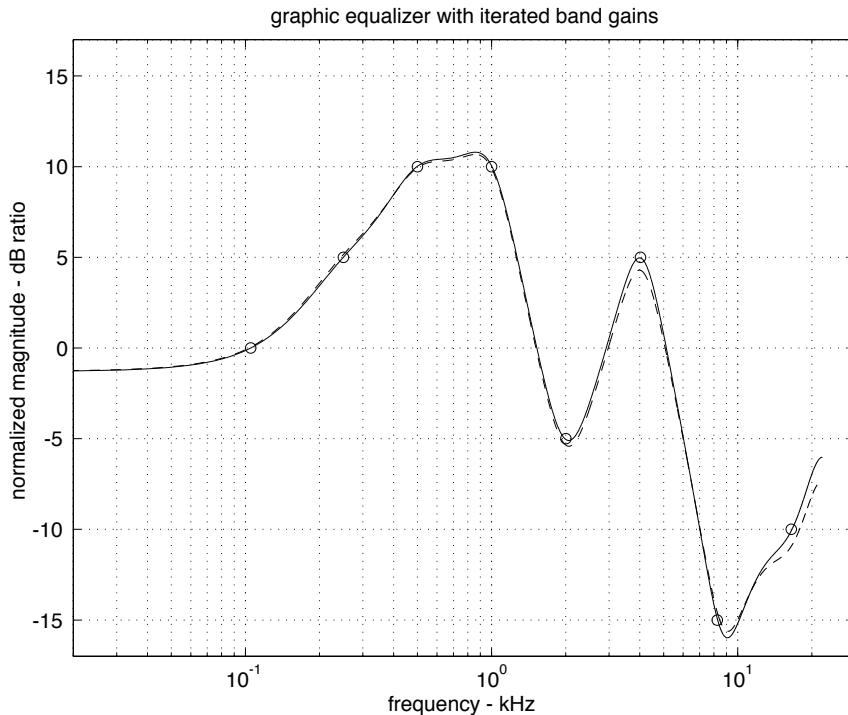


- For a graphic equalizer with $K - 1$ fixed band edges, the frequencies ω_i can be chosen as the K band centers, and the gains $\hat{\lambda}$ simply computed as the control gains γ scaled by the basis inverse,

$$\hat{\lambda} = \mathbf{B}^{-1}\gamma.$$

- Note that for a graphic equalizer with fixed band edges, \mathbf{B}^{-1} may be computed *a priori*.
- To account for discrepancies in the self similarity property, $\hat{\lambda}$ may be computed iteratively, forming \mathbf{B} using the gains from the previous solution.

Iterated Solution

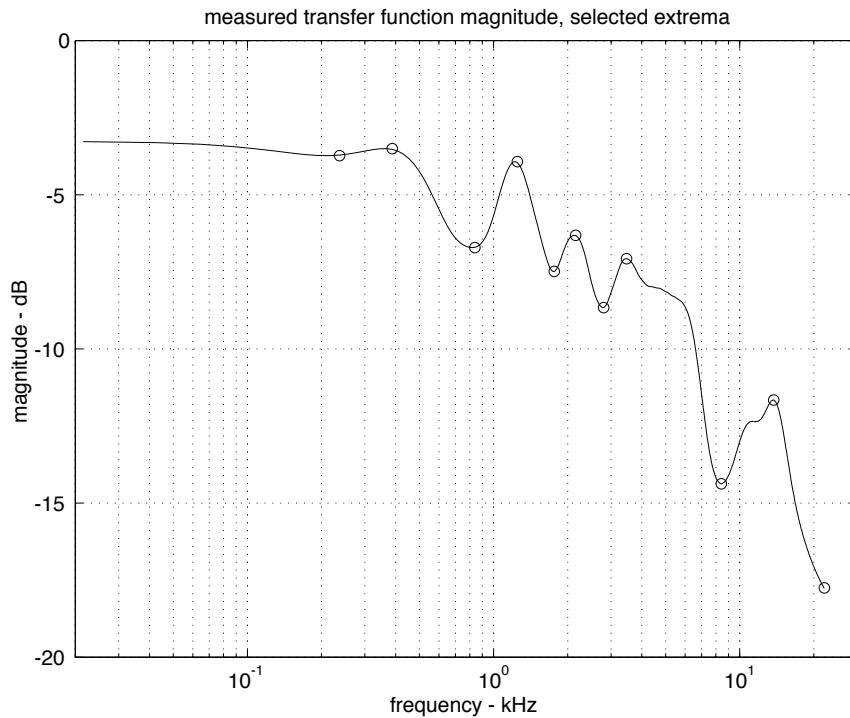


- Iterative solution of the equation for the corrected band gains can correct for discrepancies in filter self similarity,

$$\boldsymbol{\lambda}_{n+1} = \mathbf{B}(\boldsymbol{\lambda}_n)^{-1}\boldsymbol{\gamma}.$$

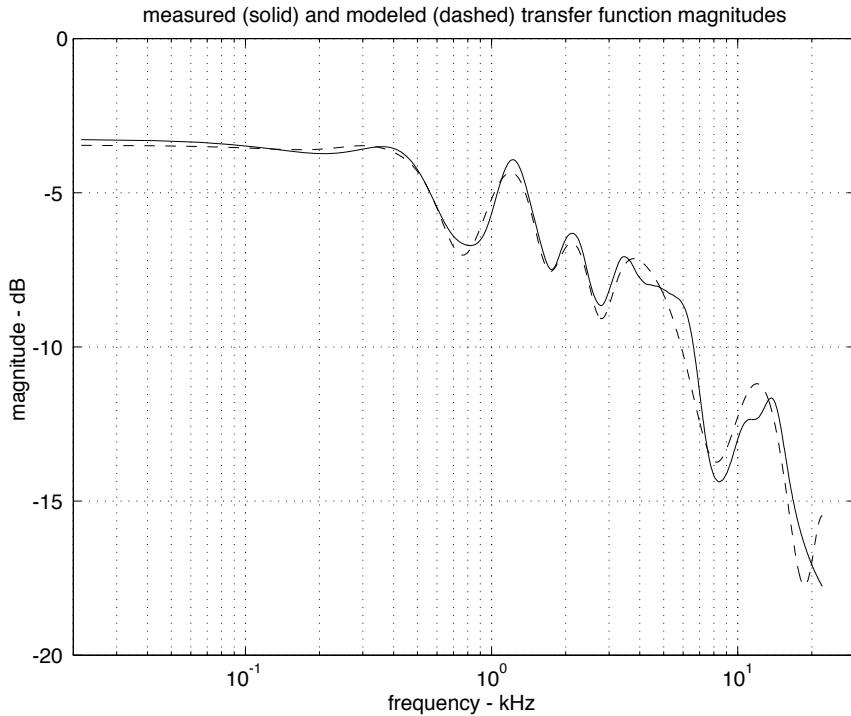
- It seems likely that the iteration converges since the sensitivity of the basis to changes in the gain are small, and changes in the basis \mathbf{B} are monotonic with changes in the gains $\boldsymbol{\lambda}$.

Transfer Function Modeling



- A cascade of peak and shelf filters can be fit to an arbitrary transfer function by selecting a set of transition frequencies, and fitting the gains.
- Transition frequency selection method.
 - Tabulate extrema or inflection point frequencies.
 - Pick transition frequencies as geometric means of significant extrema or at significant inflection points.

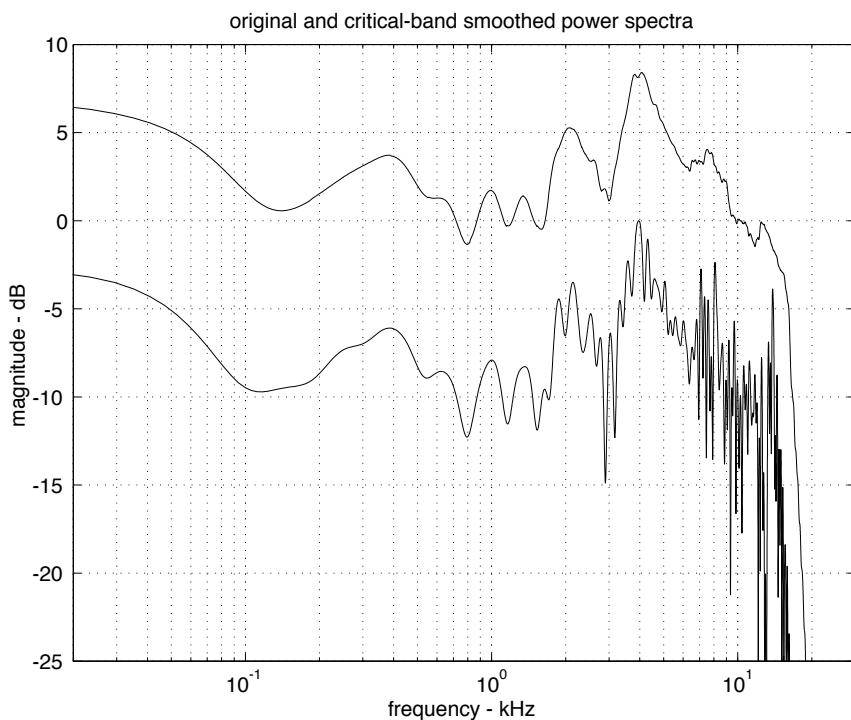
Transfer Function Modeling



- Band gain selection method.
 - Pick a dense sampling of frequencies ω_i , say Bark or ERB spaced.
 - Form basis, and compute gains,
$$\hat{\boldsymbol{\lambda}} = (\mathbf{B}^\top \mathbf{W} \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{W} \boldsymbol{\gamma}.$$
- The resulting transfer function will not interpolate the specified gains $\boldsymbol{\gamma}$, but rather approximate the desired magnitude, minimizing the weighted mean square dB difference.

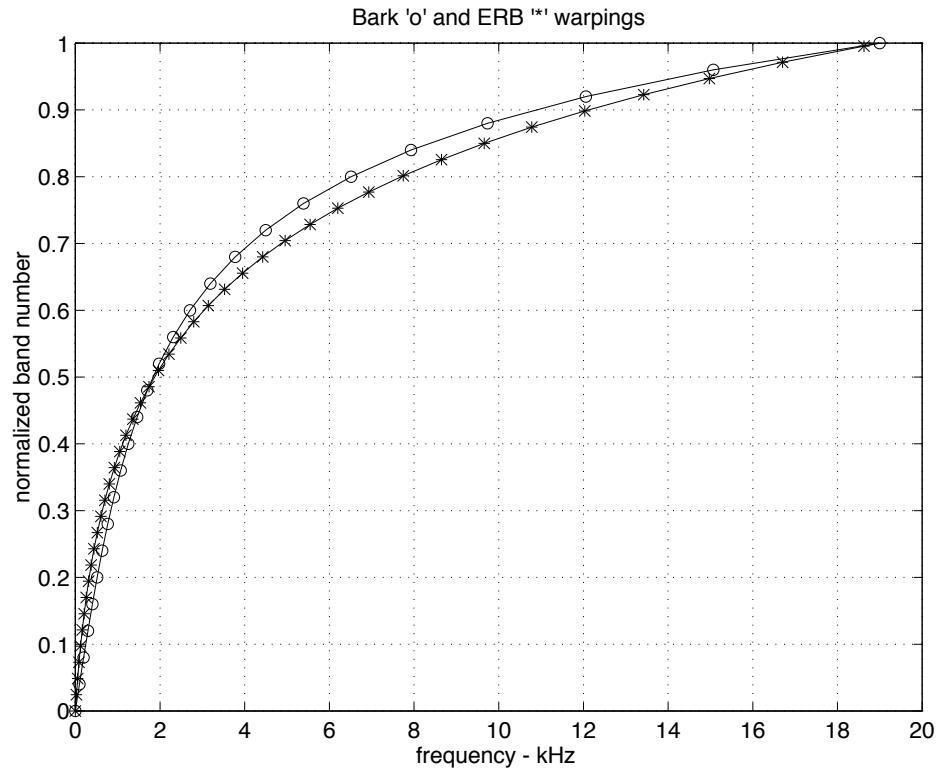
21. Critical Band Smoothing

Critical Bandwidth and Critical Band Smoothing



- When applied to broad-band signals, human hearing is only sensitive to filter features wider in frequency than a *critical bandwidth*.
- Accordingly, filters used to equalize audio signals may be “blurred” to remove features smaller than a critical bandwidth, reducing their complexity without changing their perceived character.

Critical Bandwidth

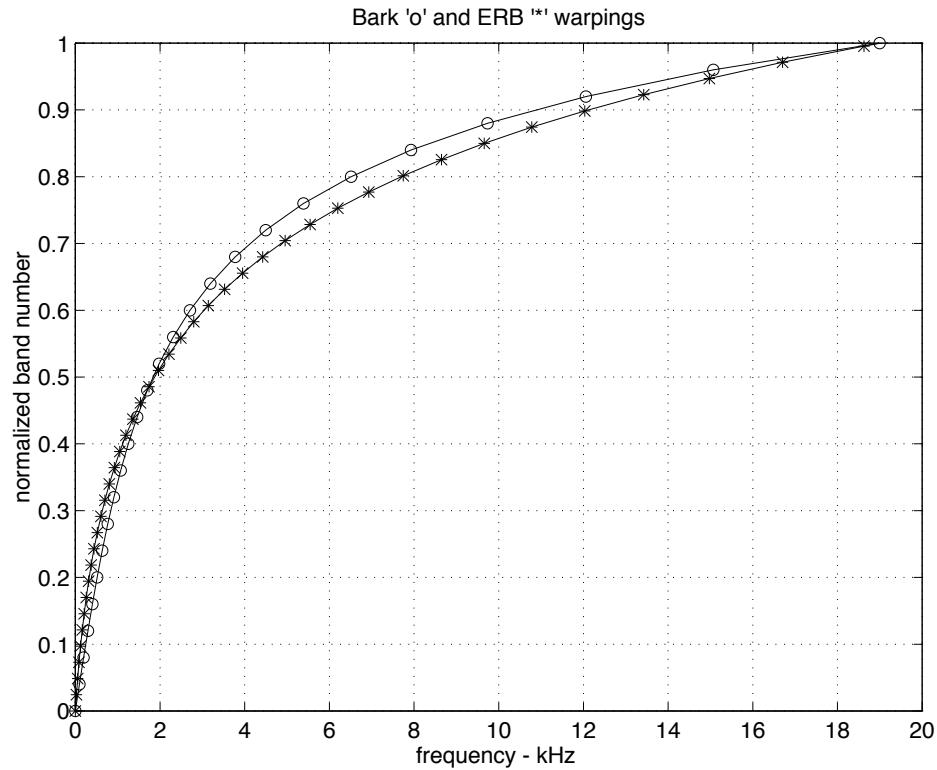


- The Bark frequency scale divides the human hearing range into about 25 critical bands.
- Frequency f in kHz and critical-band rate b_B in Bark are related via the following expressions,

$$b_B = 10.3 [\log(1 + f^2)]^{\frac{1}{2}},$$

$$f = [\exp \{(b_B/10.3)^2\} - 1]^{\frac{1}{2}}.$$

Critical Bandwidth

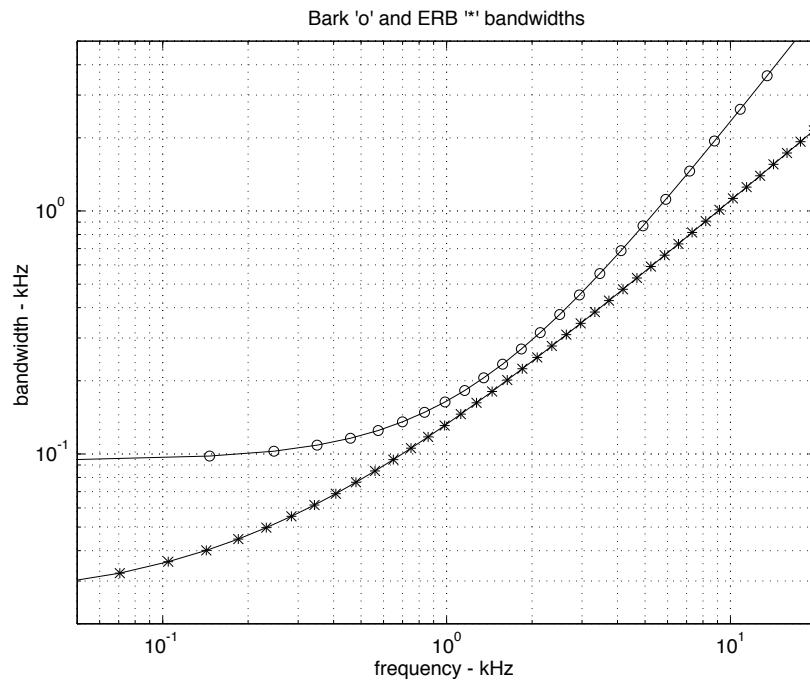


- Relatively recent experiments have revised Zwicker's scale: The ERB (equivalent rectangular bandwidth) frequency scale divides the human hearing range into about fourty ciritcal bands.
- Frequency f in kHz and critical-band rate b_E in ERB are related via the following expressions,

$$b_E = 21.4 \log_{10}(4.37f + 1.0),$$

$$f = \left[10^{b_E/21.4} - 1.0 \right] / 4.37.$$

Critical Bandwidth



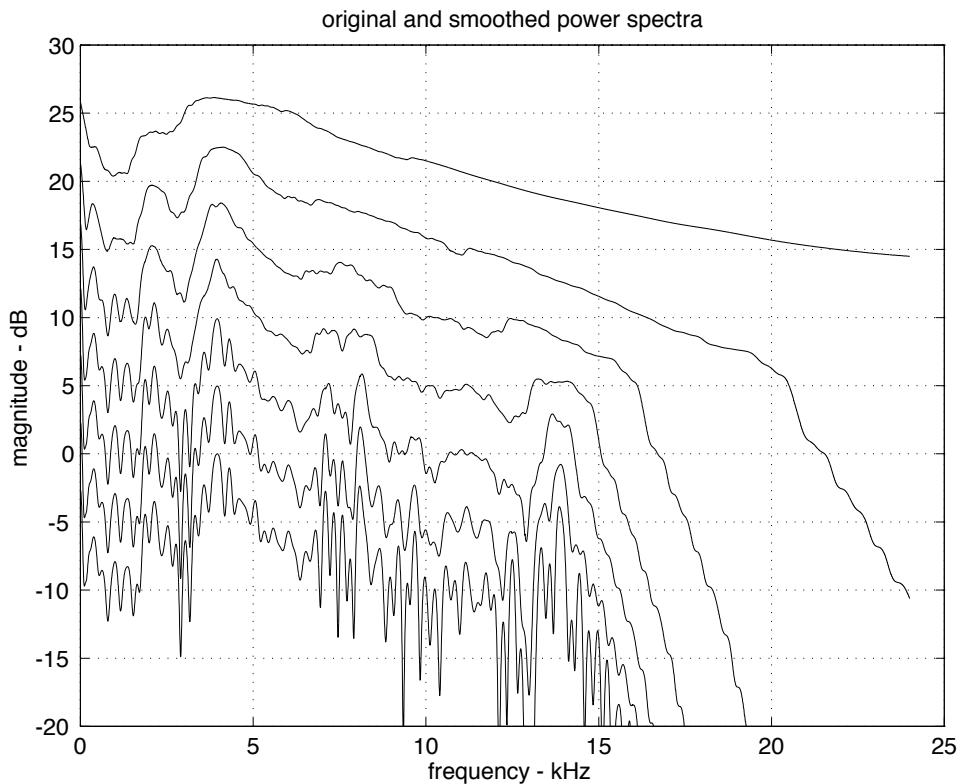
- Bark critical bandwidth as a function of frequency is approximately

$$w_B = 0.094 + 0.071f^{\frac{3}{2}}.$$

- ERB critical bandwidth as a function of frequency is

$$w_E = 0.0247 + 0.1079f.$$

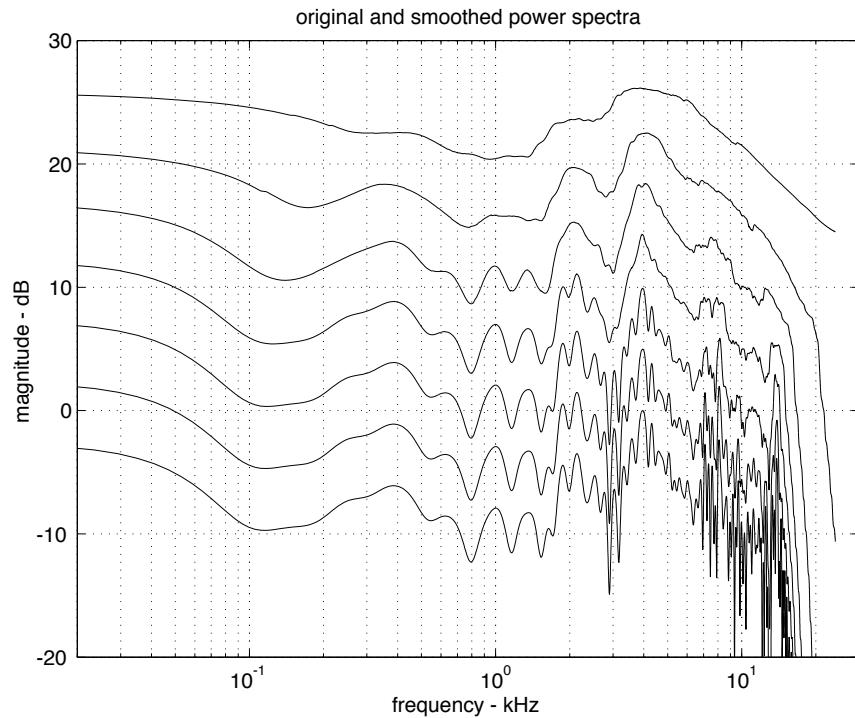
Critical Band Smoothing



- By smoothing the filter power over bandwidths proportional to critical bandwidth, filter complexity is reduced while retaining psychoacoustically relevant cues.
- Critical-band smoothed filter power may be computed via a running mean over frequency,

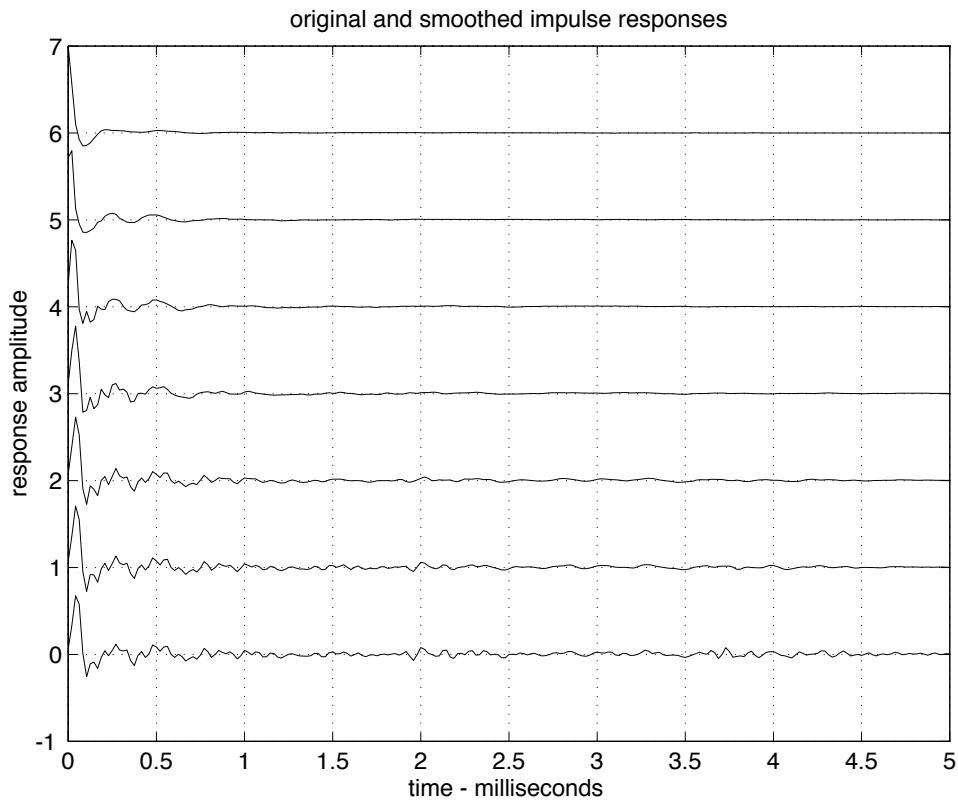
$$P(\omega; \beta) = \sum_{\varphi=f(b(\omega)-\beta/2)}^{f(b(\omega)+\beta/2)} |H(\varphi)|^2.$$

Critical Band Smoothing



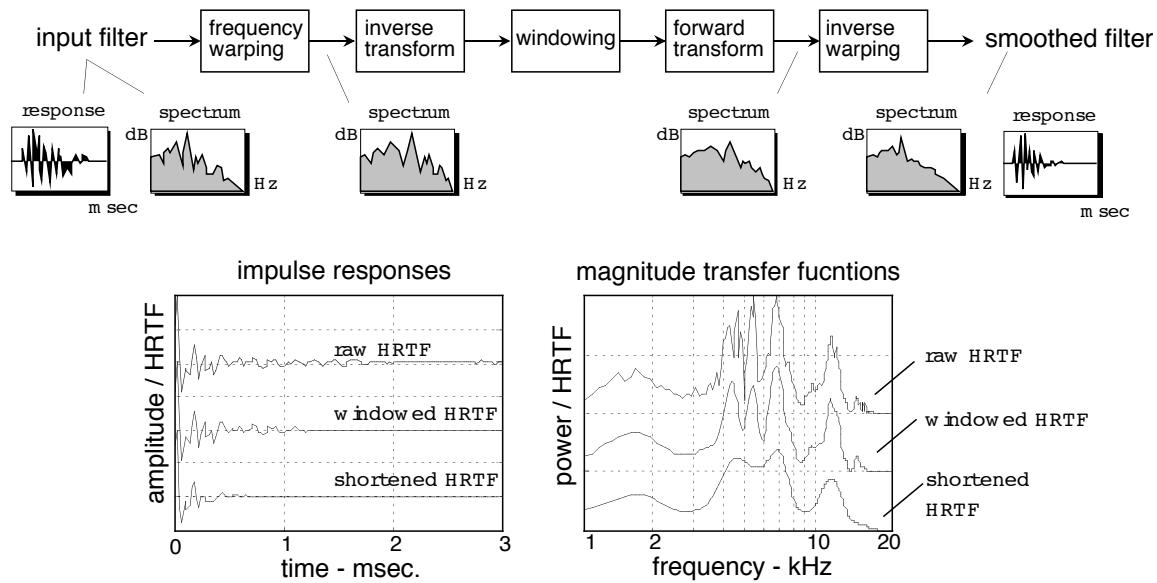
- Critical-band smoothing bandwidth is relatively constant on a log frequency axis.
- Example: filtered drum signal.

Filter Complexity Reduction



- Filter impulse responses are made shorter by critical band smoothing.
- Note that time-domain windowing imposes a constant-bandwidth running mean (convolution with the window transform) on the spectrum. It's therefore not surprising that a filter impulse response is made shorter by critical-band smoothing.

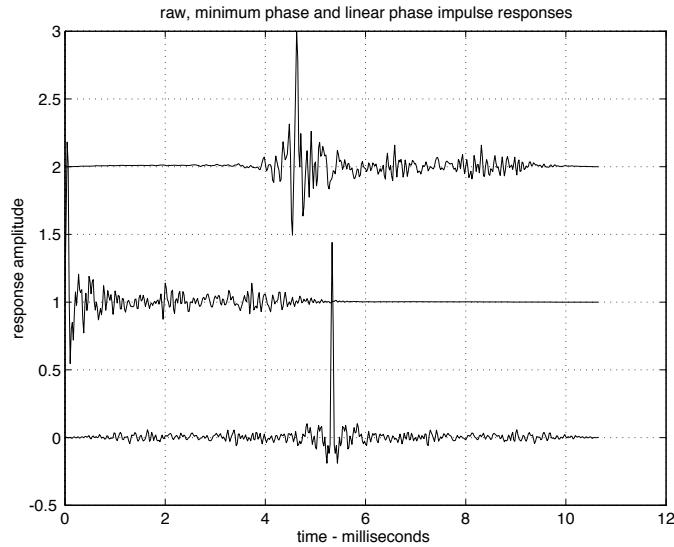
Alternate Critical Band Smoothing Method



- Critical-band smoothing may be done by time-domain windowing the autocorrelation of the frequency-warped filter power.

22. Filter Phase, Linear and Minimum Phase

Filter Phase



- Humans are relatively insensitive to filter phase.
- A *linear phase* filter delays all frequency components the same amount τ ,

$$h_{\text{lin}}(n) = \mathcal{F}^{-1} \left\{ e^{-j\tau\omega} \cdot |H(\omega)| \right\}.$$

- A *minimum phase* filter $h_{\min}(n)$ delays input energy the least of all filters $g(n)$ having the same magnitude response $|G(\omega)| = |H(\omega)|$,

$$\sum_{n=0}^m h_{\min}^2(n) \geq \sum_{n=0}^m g^2(n), \quad \forall m.$$

Impulse Response and Transform Symmetries

- Consider an impulse response $h(n)$ which is
 - causal,
 - real and
 - stable.
 - What can we say about its z -transform $H(z)$?
- ⇒ Stability implies that the unit circle is in its region of convergence.
- ⇒ The impulse response $h(n)$ being real means that the transform $H(z)$ has Hermitian symmetry. Denoting by $H_R(z)$ and $H_I(z)$ the real and imaginary parts of the transform $H(z)$, $h(n)$ real implies

$$H_R(z) = H_R(z^*), \quad H_I(z) = -H_I(z^*).$$

- ⇒ That the impulse response $h(n)$ is causal implies a *Hilbert transform* relationship between the real and imaginary components of the transform $H(z)$.

Impulse Response Causality

- If the impulse response $h(n)$ is causal, multiplying by the unit step $u(n)$ leaves it unchanged,

$$h(n) = u(n) \cdot h(n).$$

- The transform $H(\omega)$ is then equal to its convolution with the transform of the unit step $U(\omega)$,

$$H(\omega) = U(\omega) * H(\omega).$$

- The unit step transform $U(\omega)$ may be found as the limiting case of the transform of a decaying exponential $r^{-n}u(n)$, $r > 1$,

$$U(\omega) = \lim_{r \rightarrow 1^+} \sum_{n=0}^{\infty} r^{-n} e^{-j\omega n} = \lim_{r \rightarrow 1^+} \frac{1}{1 - re^{-j\omega}},$$

although this approach requires extreme care in taking the limit for small ω .

- An alternative approach is to decompose the unit step $u(n)$ into even (constant and impulse) and odd (signum function) terms,

$$\begin{aligned} u(n) &= \frac{1}{2} \cdot [1 + \delta(n) + \text{sign}(n)], \\ \text{sign}(n) &= \begin{cases} 1, & n > 0, \\ 0, & n = 0, \\ -1, & n < 0. \end{cases} \end{aligned}$$

Impulse Response Causality

- The unit step transform $U(\omega)$ is then

$$U(\omega) = \frac{1}{2} \cdot [\delta(\omega) + 1 + \mathcal{F}\{\text{sign}(n)\}],$$

where the transform of the signum function involves a more straightforward limit,

$$\begin{aligned} \mathcal{F}\{\text{sign}(n)\} &= \lim_{r \rightarrow 1^+} \sum_{n=1}^{\infty} r^{-n} (e^{-j\omega n} - e^{j\omega n}) \\ &= \lim_{r \rightarrow 1^+} \frac{r^{-1} e^{-j\omega}}{1 - r^{-1} e^{-j\omega}} - \frac{r^{-1} e^{j\omega}}{1 - r^{-1} e^{j\omega}} \\ &= -j \cdot \frac{\sin \omega}{1 - \cos \omega} = -j \cdot \cot(\omega/2). \end{aligned}$$

- We have

$$U(\omega) = \frac{1}{2} \cdot [\delta(\omega) + 1 - j \cdot \cot(\omega/2)].$$

- The transform $H(\omega)$ therefore satisfies

$$H(\omega) = \frac{1}{2} \cdot [\delta * H(\omega) + 1 * H(\omega) - j \cot(\omega/2) * H(\omega)],$$

which reduces to

$$H(\omega) = 1 * H(\omega) - j \cot(\omega/2) * H(\omega).$$

Impulse Response Causality

- We have

$$H(\omega) = 1 * H(\omega) - j \cot(\omega/2) * H(\omega).$$

- Substituting

$$1 * H(\omega) = h(0)$$

and

$$H(\omega) = H_R(\omega) + jH_I(\omega),$$

and equating real and imaginary parts, we obtain

$$H_I(\omega) = -\cot(\omega/2) * H_R(\omega).$$

$$H_R(\omega) = h(0) + \cot(\omega/2) * H_I(\omega).$$

- If an impulse response $h(n)$ is causal, therefore, its transform $H(\omega)$ may be recovered from only its real part $H_R(\omega)$ or from only its imaginary part $H_I(\omega)$.
- The operation of convolving with $-\cot(\omega/2)$ on the unit circle in the z -plane, or with $-2/\omega$ on the imaginary axis in the s -plane is referred to as the *Hilbert transform*, and is equivalent to multiplying the time signal by $j\text{sign}(t)$.
- Applied in the time domain, the Hilbert transform shifts the signal phase by $\pm\pi/2$, turning $\sin\omega$ into $\cos\omega$, for instance.

Minimum Phase Relationship

- Consider the transform $H(z)$ specified in polar form,

$$H(z) = |H(z)|e^{j\arg[H(z)]},$$

and denote by $\check{H}(z)$ the log of $H(z)$,

$$\check{H}(z) = \log[H(z)] = \log|H(z)| + j\arg[H(z)],$$

and by $\check{h}(n)$ the sequence having $\check{H}(z)$ as its z -transform.

- The sequence $\check{h}(n)$ will be causal, real, and stable if and only if $\log|H(z)|$ and $\arg[H(z)]$ form a Hilbert transform pair.
- If the sequence $\check{h}(n)$ is causal, real, and stable, then the system $H(z)$ is said to be *minimum phase*, and has the following properties.
 - $H(z)$ has no poles or zeros outside the unit circle,
 - $H(z)$ has a causal, real, stable inverse system $H^{-1}(z)$. $H(z) = 1$.
 - The impulse response $h(n)$ concentrates energy,

$$\sum_{n=0}^m h^2(n) \geq \sum_{n=0}^m g^2(n), \quad \forall m,$$

for all $g(n)$ such that $|G(\omega)| = |H(\omega)|$.

Minimum Phase and Rational Systems

- Consider the log transform

$$\check{X}(z) = \log(1 - \alpha \cdot z^{-1})$$

with the singularity inside the unit circle, $|\alpha| < 1$.

- Expanding the log in a Taylor series about $\alpha z^{-1} = 0$, we have

$$\check{X}(z) = - \sum_{n=1}^{\infty} \frac{\alpha^n z^{-n}}{n},$$

which implies

$$\check{x}(n) = \begin{cases} -\frac{\alpha^n}{n}, & n \geq 1, \\ 0, & n \leq 0, \end{cases}$$

a *causal*, stable sequence.

- Now consider a log transform of a system with a singularity outside the unit circle, $|\alpha| > 1$,

$$\check{X}(z) = \log(1 - z/\alpha).$$

We now have

$$\check{x}(n) = \begin{cases} 0, & n \geq 0, \\ -\frac{\alpha^n}{n}, & n \leq -1, \end{cases}$$

an *anticausal*, stable sequence.

Minimum Phase and Rational Systems

- Consider a rational system $H(z)$ of the form

$$H(z) = \gamma z^r \cdot \frac{\prod_n (1 - \beta_n z^{-1}) \prod_m (1 - z/b_m)}{\prod_n (1 - \alpha_n z^{-1}) \prod_m (1 - z/a_m)},$$

with $|\alpha_n|, |\beta_n| < 1$ and $|a_m|, |b_m| > 1$.

- Its log transform $\check{H}(z) = \log H(z)$ is the sum of the log pole and zero terms,

$$\begin{aligned} \check{H}(z) = & \sum_n \log(1 - \beta_n z^{-1}) - \sum_n \log(1 - \alpha_n z^{-1}) \\ & + \sum_m \log(1 - z/b_m) - \sum_m \log(1 - z/a_m) \\ & + \log \gamma + r \log z. \end{aligned}$$

- Ignoring the $\log z^r = j\omega r$ term (which only contains information regarding the time origin of $h(n)$), the complex cepstrum of $h(n)$ is

$$\check{h}(n) = \begin{cases} \sum_p \frac{\alpha_p^n}{n} - \sum_p \frac{\beta_p^n}{n}, & n \geq 1, \\ \sum_q \frac{a_q^n}{n} - \sum_q \frac{b_q^n}{n}, & n \leq -1, \\ \log \gamma, & n = 0. \end{cases}$$

Minimum Phase and Rational Systems

- The following statements are equivalent, as can be seen by analyzing the complex cepstrum.
 - The complex cepstrum $\check{h}(n)$ is causal and stable.
 - The system $H(z)$ has no poles or zeros outside the unit circle.
 - A Hilbert transform relationship exists between the log magnitude $\log |H(\omega)|$ and phase $\arg H(\omega)$ of the transfer function $H(\omega)$.
- Finally, note that any minimum phase system will have a minimum phase inverse system (exchange the poles and the zeros).

Minimum Phase Interpretation

- Any rational system can be written as the product of a minimum phase system and an allpass system.

$$H(z) = H_{\min}(z) \cdot H_{\text{ap}}(z).$$

- This can be seen by considering a zero outside the unit circle at z_0 . We have

$$H(z) = H_1(z) \cdot (z^{-1} - z_0),$$

which may be written as

$$H(z) = H_1(z) \cdot (1 - z_0^* z^{-1}) \cdot \frac{z^{-1} - z_0}{1 - z_0^* z^{-1}}.$$

- A quick Bode plot will show that an allpass system has phase lag, and therefore, a minimum phase system is the one having *minimum phase lag* for a given transfer function magnitude.

Minimum Phase Energy Concentration

- Minimum phase sequences concentrate energy in the beginning of the sequence.

$$\sum_{n=0}^m h_{\min}^2(n) \geq \sum_{n=0}^m h^2(n), \quad \forall m,$$

for all sequences $h(n)$ having transfer function magnitude $|H(\omega)|$.

- To see that this is true, consider a minimum phase system

$$H_{\min}(z) = Q(z) \cdot (1 - z_k z^{-1}),$$

and the same system with a nonminimum phase zero

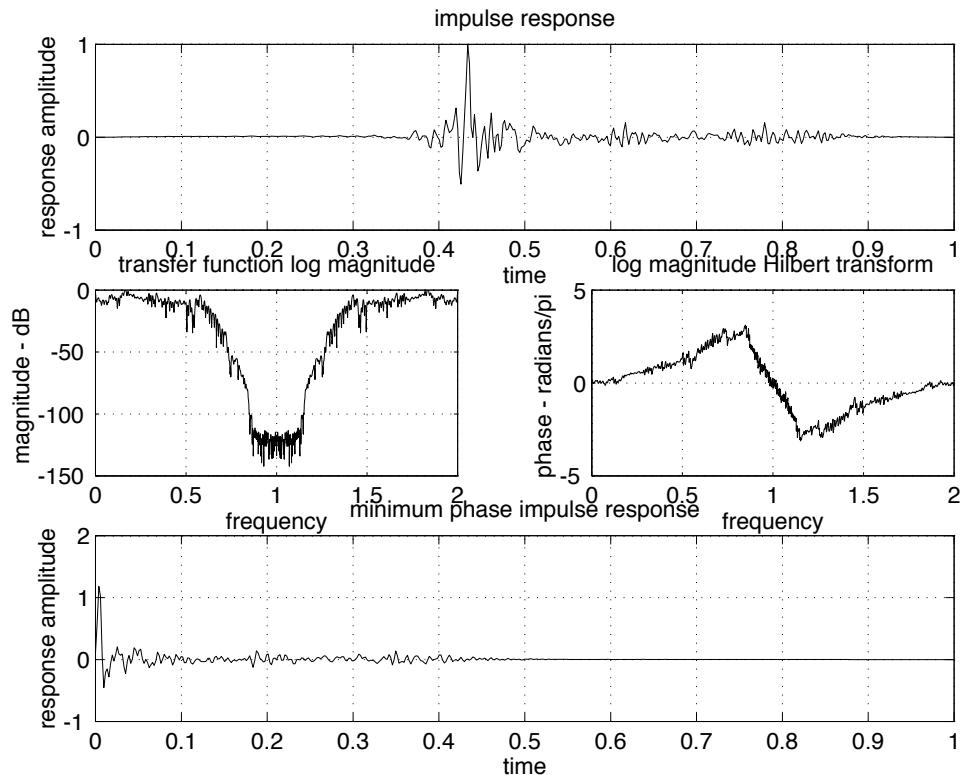
$$H(z) = Q(z) \cdot (z^{-1} - z_k^*).$$

- The difference in energy concentration is given by

$$\sum_{n=0}^m [h_{\min}^2(n) - h^2(n)] = (1 - |z_k|^2) \cdot q(m)^2,$$

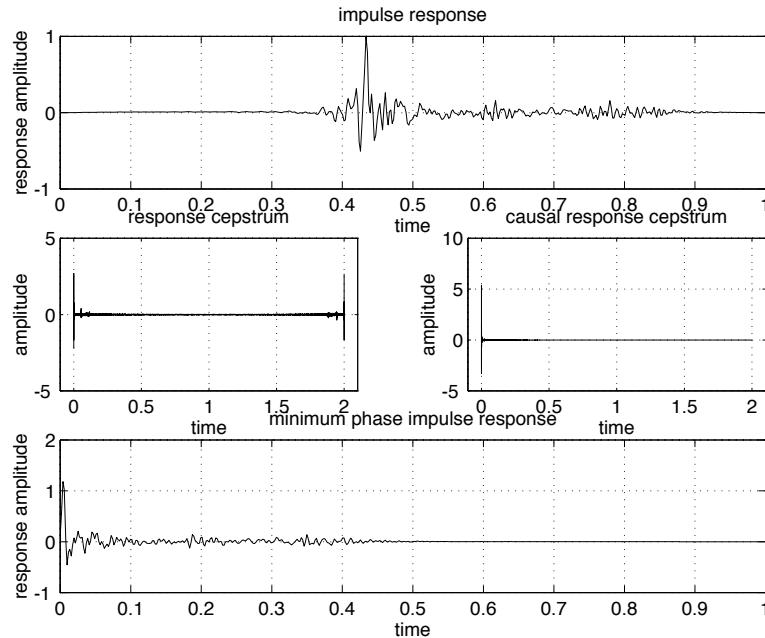
which is greater than or equal to zero for all $|z_k| \leq 1$.

Computing Minimum Phase Impulse Responses



- To convert $h(n)$ to minimum phase:
 - Form the log magnitude transform,
$$\log |H(\omega)| = \log |\mathcal{F}\{h(n)\}|.$$
 - Generate the transform phase by applying a Hilbert transform to the log magnitude,
$$\arg H(w) = -\mathcal{H}\{\log |H(\omega)|\}.$$
 - Exponentiate and inverse transform,
$$h_{\min} = \mathcal{F}^{-1} \left\{ \exp [\log |H(\omega)| - j\mathcal{H}\{\log |H(\omega)|\}] \right\}.$$

Computing Minimum Phase Impulse Responses



- To convert $h(n)$ to minimum phase:

- Form the complex cepstrum,

$$\check{h}(n) = \mathcal{F}^{-1} \{ \log |\mathcal{F}\{h(n)\}| \} .$$

- Time flip the noncausal cepstrum components,

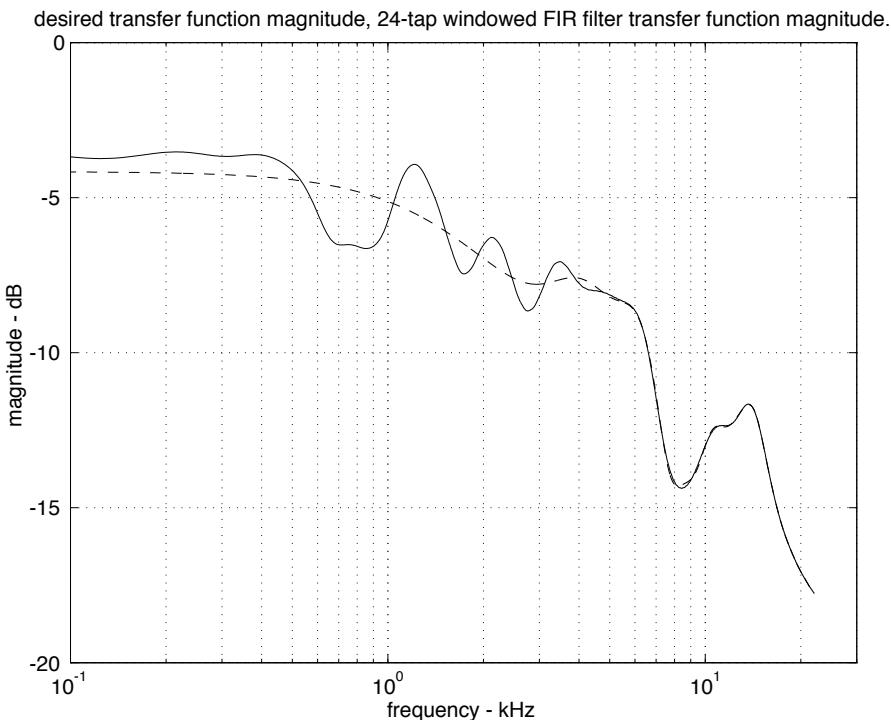
$$\check{h}_c(n) = \begin{cases} \check{h}(0), & n = 0, \\ \check{h}(n) + \check{h}(-n), & n \geq 1, \\ 0, & n \leq -1. \end{cases}$$

- Transform, exponentiate and inverse transform,

$$h_{\min} = \mathcal{F}^{-1} \{ \exp [\mathcal{F}\{\check{h}_c(n)\}] \} .$$

23. Frequency Warping; Warped FIR Filter Design

Windowed Impulse Response Filter Design

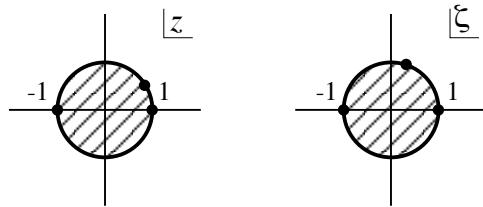


- Set the taps of an FIR filter $g(n)$ to a windowed version of the desired impulse response $h(n)$,

$$g(n) = w(n) \cdot h(n).$$

- The time-domain windowing imposes a uniform bandwidth smearing on the transfer function $H(\omega)$.

First-Order Conformal Map



- Consider the first-order *allpass transformation* (or *conformal map*) with parameter ρ , defined by the substitution

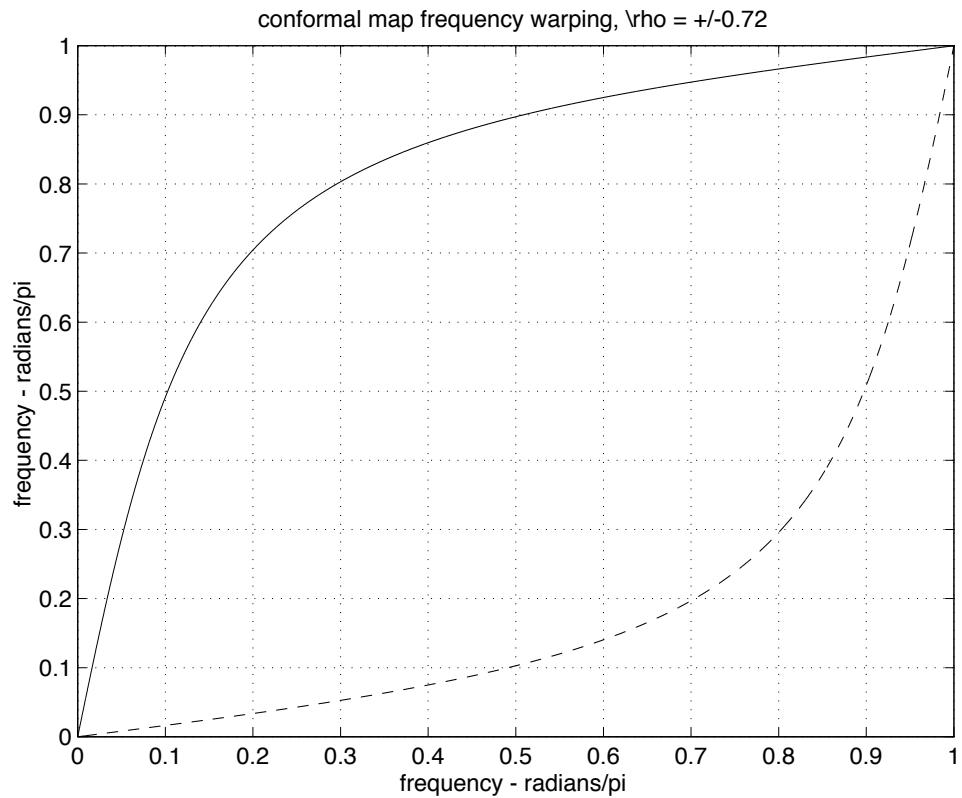
$$\zeta^{-1}(z) = \frac{\rho + z^{-1}}{1 + \rho z^{-1}}.$$

- Note that the transformation maps the unit circle to itself, preserving the locations of DC and the band edge, $z = \pm 1$,

$$|\zeta^{-1}(e^{j\omega})|^2 = \frac{(\rho + z^{-1})(\rho + z)}{(1 + \rho z^{-1})(1 + \rho z)} = 1.$$

- Points inside (outside) the unit circle on the z plane, map to points inside (outside) the unit circle on the ζ plane, $|\zeta| < 1$.

First-Order Conformal Map Frequency Warping



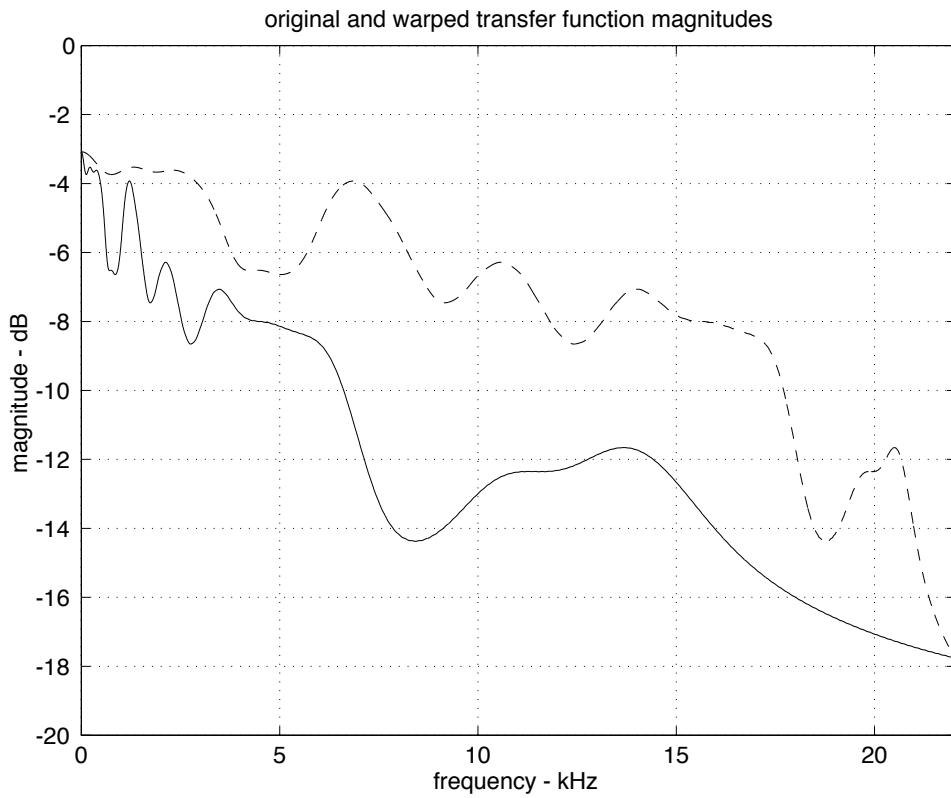
- Points $z^{-1} = e^{-j\omega}$ on the z -plane unit circle are warped to points $\zeta^{-1} = e^{-j\alpha}$ on the ζ -plane unit circle according to ρ ,

$$\rho = \frac{\sin[(\omega - \alpha)/2]}{\sin[(\omega + \alpha)/2]}.$$

- The mapping defined by $-\rho$ is the inverse of the mapping defined by ρ ,

$$z^{-1} = \zeta_{-\rho}^{-1}(\zeta_\rho(z^{-1})).$$

Warped FIR Filter



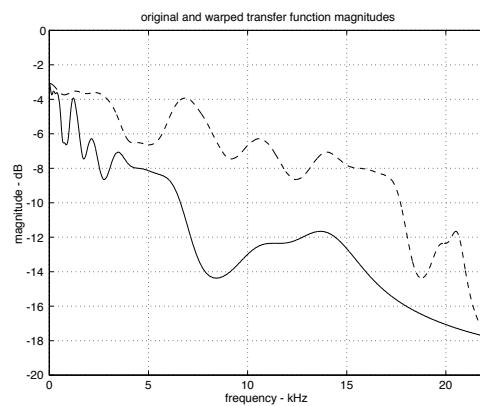
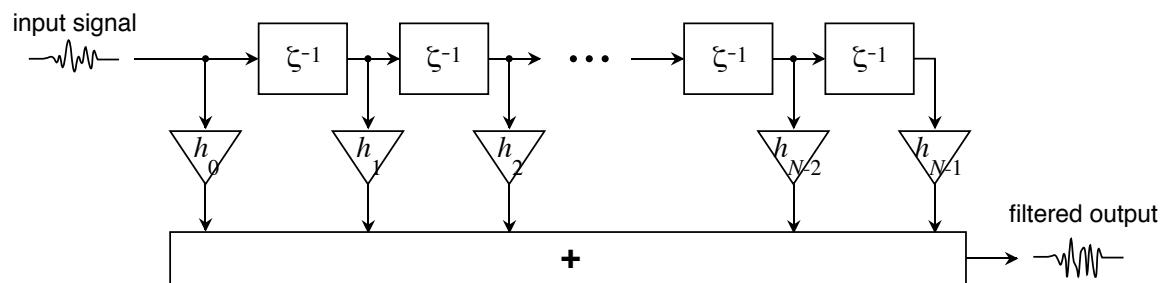
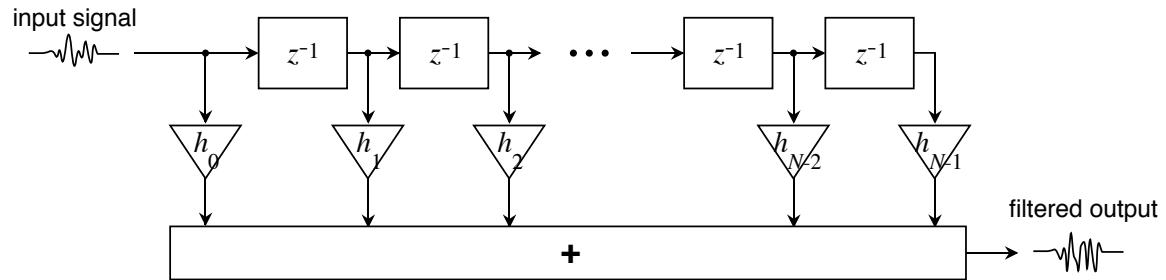
- An FIR filter with impulse response $g(n)$ has z -transform

$$G(z) = \sum_{n=0}^{N-1} g(n) \cdot z^{-n}.$$

- By replacing the unit delay z^{-1} with the warped unit delay $\zeta_\rho^{-1}(z)$, the frequency axis of the original FIR filter is warped.

$$\Gamma_\rho(z) = G(\zeta_\rho(z)) = \sum_{n=0}^{N-1} g(n) \cdot \zeta_\rho^{-n}.$$

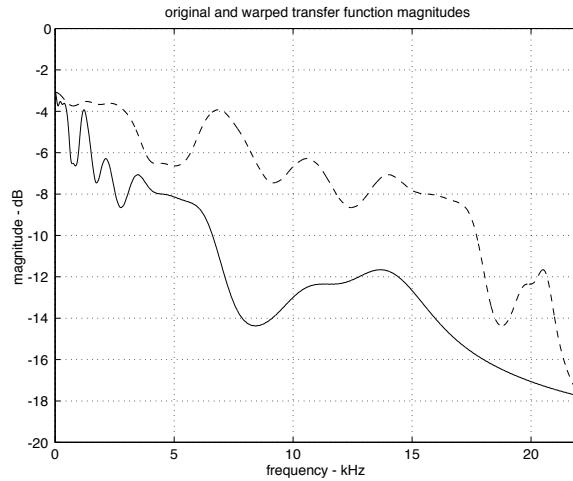
Warped FIR Filter Implementation



- A warped FIR filter may be implemented using a warped tapped delay line.

$$G(z) = \sum_{n=0}^{N-1} g(n)z^{-n}, \quad \Gamma_\rho(z) = \sum_{n=0}^{N-1} g(n)\zeta_\rho^{-n}.$$

Warped FIR Filter Design



- Find ρ such that important spectral features are given large bandwidth when warped through $\zeta_\rho^{-1}(z)$.
- Form $\eta(n)$, the warped impulse response,

$$\eta(n) = \sum_{m=0}^{M-1} h(m) \zeta_\rho^{-m}(z)$$

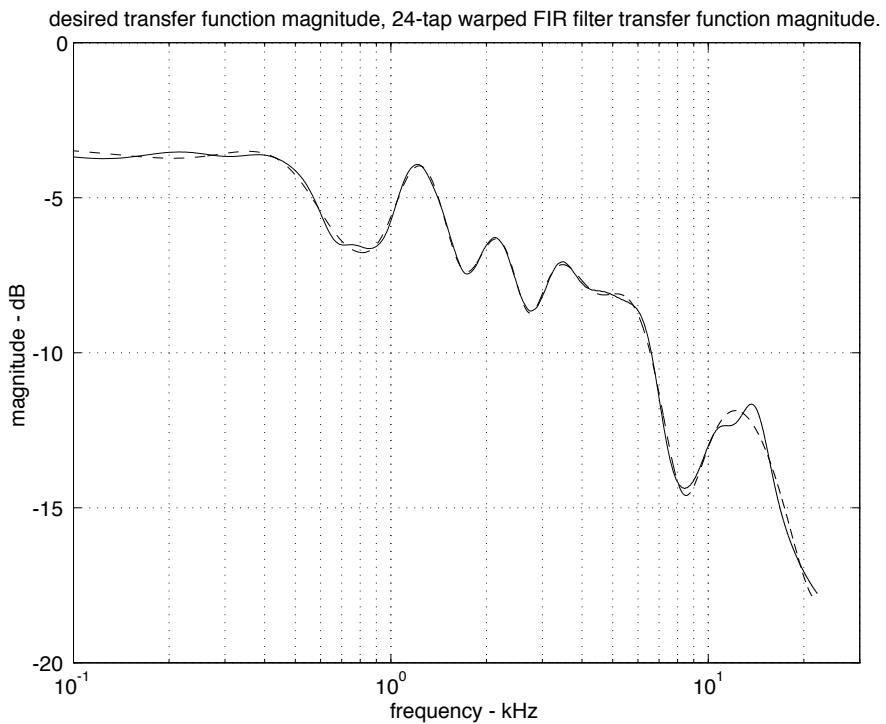
- Window the desired warped impulse response $\eta(n)$ to form $\gamma(n)$,

$$\gamma(n) = w(n) \cdot \eta(n).$$

- Implement the warped FIR filter with allpass parameter $-\rho$ using taps $\gamma(n)$,

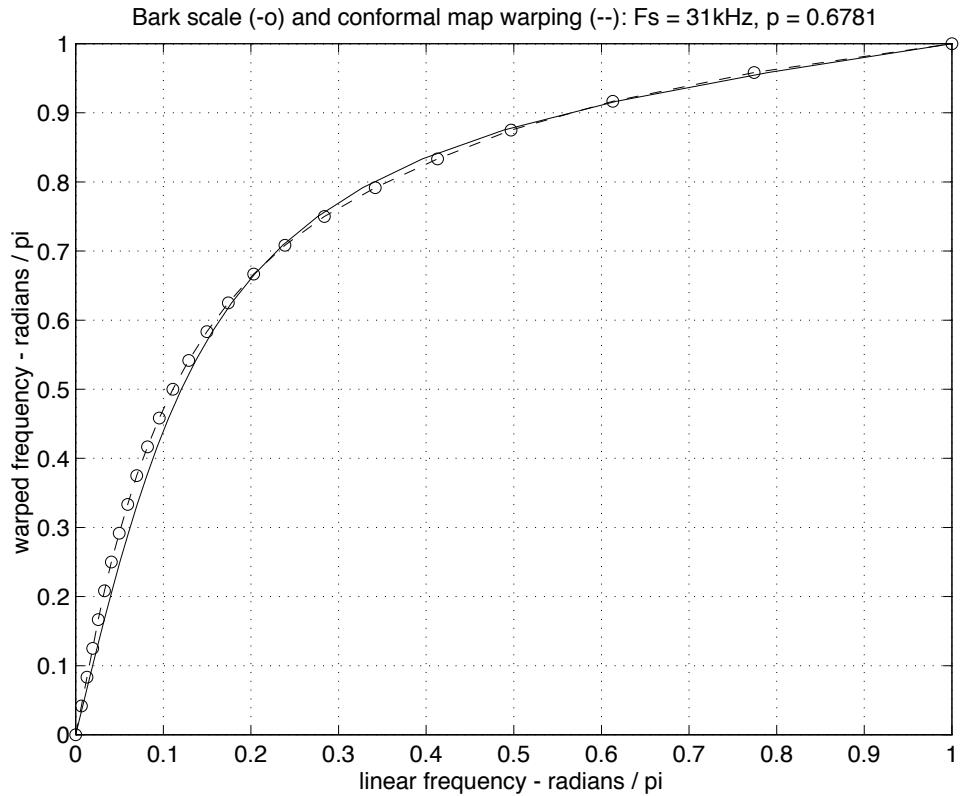
$$\sum_{n=0}^{N-1} \gamma(n) \zeta_{-\rho}^{-n}.$$

Warped FIR Filter Design Example



- In the example shown, the allpass parameter ρ was selected so that the low-frequency features would be accurately modeled.

Allpass Parameter ρ Selection



- For the proper choice of allpass parameter ρ , the first-order conformal map $\zeta_\rho^{-1}(z)$ imposes a frequency warping very much like either the Bark or ERB frequency scale.

$$\rho_{\text{Bark}}^*(f_s) = 1.07 \cdot \left[\frac{2}{\pi} \cdot \arctan(0.066f_s) \right]^{\frac{1}{2}} - 0.191$$

$$\rho_{\text{ERB}}^*(f_s) = 1.05 \cdot \left[\frac{2}{\pi} \cdot \arctan(0.072f_s) \right]^{\frac{1}{2}} - 0.196$$

24. IIR Filter Design and Prony's Method

IIR Filter Design and Prony's Method

- Given a rational filter

$$G(z) = \frac{B(z)}{A(z)}$$

with impulse response

$$g(n) = \begin{cases} -\sum_{k=1}^p a_k \cdot g(n-k) + \\ \quad \sum_{k=0}^q b_k \cdot \delta(n-k), & n \geq 0 \\ 0, & n < 0, \end{cases}$$

the idea is to find filter coefficients a_k and b_k such that $g(n)$ approximates a desired impulse response $h(n)$.

- One approach is to minimize the norm of the *output error*,

$$J = \sum_n \epsilon_O(n)^2, \quad \epsilon_O(n) = h(n) - g(n).$$

Output Error Minimization

- Consider a one-pole system

$$G(z) = \frac{b_0}{1 - a_1 z^{-1}}$$

with impulse response

$$g(n) = \begin{cases} b_0 a_1^n & n \geq 0 \\ 0, & n < 0. \end{cases}$$

- Note that the output error $\epsilon_O(n)$ is highly nonlinear in the unknown pole coefficient a_1 ,

$$\epsilon_O(n) = h(n) - b_0 a_1^n,$$

and, as a result, the coefficients minimizing the sum of square output errors J are difficult to compute directly.

- In fact, the sum of square output errors J will likely have multiple local minima as a function of the unknown coefficients, making its optimization difficult.

Prony's Method for IIR Filter Design

- In Prony's method, desired impulse response samples $h(n)$ are substituted for hypothesized ones $g(n)$ in the recursion for the hypothesized impulse response,

$$h(n) - \epsilon_E(n) = - \sum_{k=1}^p a_k \cdot h(n-k) + \sum_{k=0}^q b_k \cdot \delta(n-k).$$

- Note that an *equation error* ϵ_E has been introduced to account for discrepancies in the recursion resulting from using $h(n)$ in place of $g(n)$.
- Stacking the first $q+1$ samples of the equation error,

$$\epsilon_{Eq} = \mathbf{H}_q \mathbf{a} - \hat{\mathbf{b}} + \mathbf{h}_q,$$

where ϵ_{Eq} is the column of equation errors, \mathbf{h}_q the first $q+1$ samples of $h(n)$, and

$$\mathbf{H}_q = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ h(0) & 0 & \cdots & 0 \\ h(1) & h(0) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ h(q-1) & h(q-2) & \cdots & h(q-p) \end{bmatrix}.$$

- Regardless of the value of $\hat{\mathbf{a}}$, the first $q+1$ samples of the equation error can be made zero by setting

$$\hat{\mathbf{b}} = \mathbf{h}_q + \mathbf{H}_q \hat{\mathbf{a}}.$$

Prony's Method for IIR Filter Design

- Stacking the remaining instances of the equation error, we have

$$\epsilon_E = \mathbf{H}\mathbf{a} + \mathbf{h},$$

where

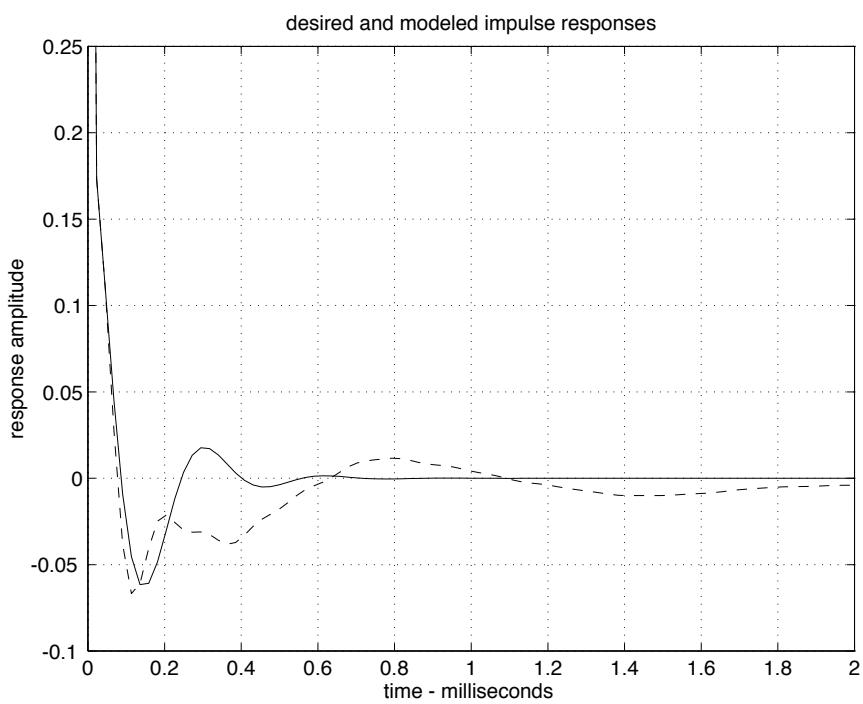
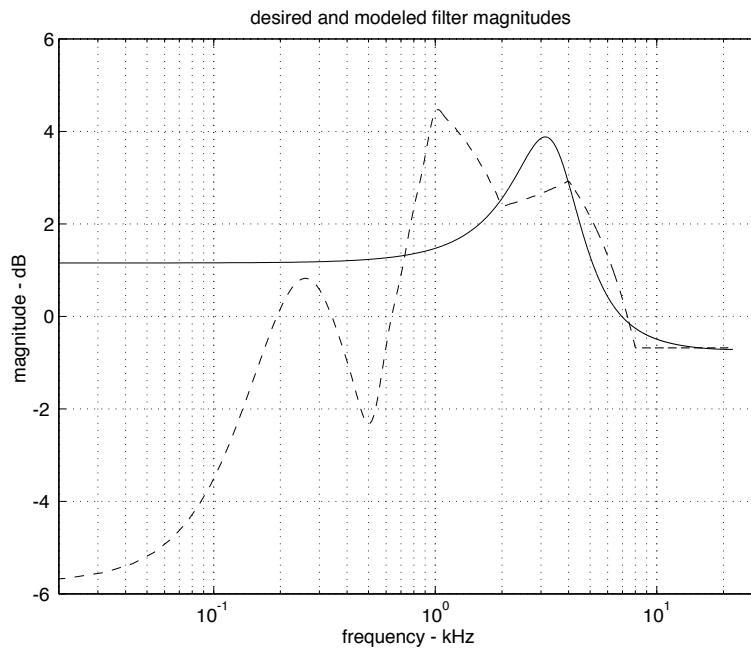
$$\begin{aligned} \mathbf{h} &= [h(q+1) \ h(q+2) \ \cdots \ h(N-1)]^\top, \\ \mathbf{H} &= \begin{bmatrix} h(q) & h(q-1) & \cdots & h(q-p+1) \\ h(q+1) & h(q) & \cdots & h(q-p+2) \\ \vdots & \vdots & \vdots & \vdots \\ h(N-1) & h(N-2) & \cdots & h(N-1-p) \end{bmatrix}. \end{aligned}$$

- The equation error ϵ_E has the benefit of being linear in the unknown filter coefficients, and its norm is easily minimized using least squares techniques.

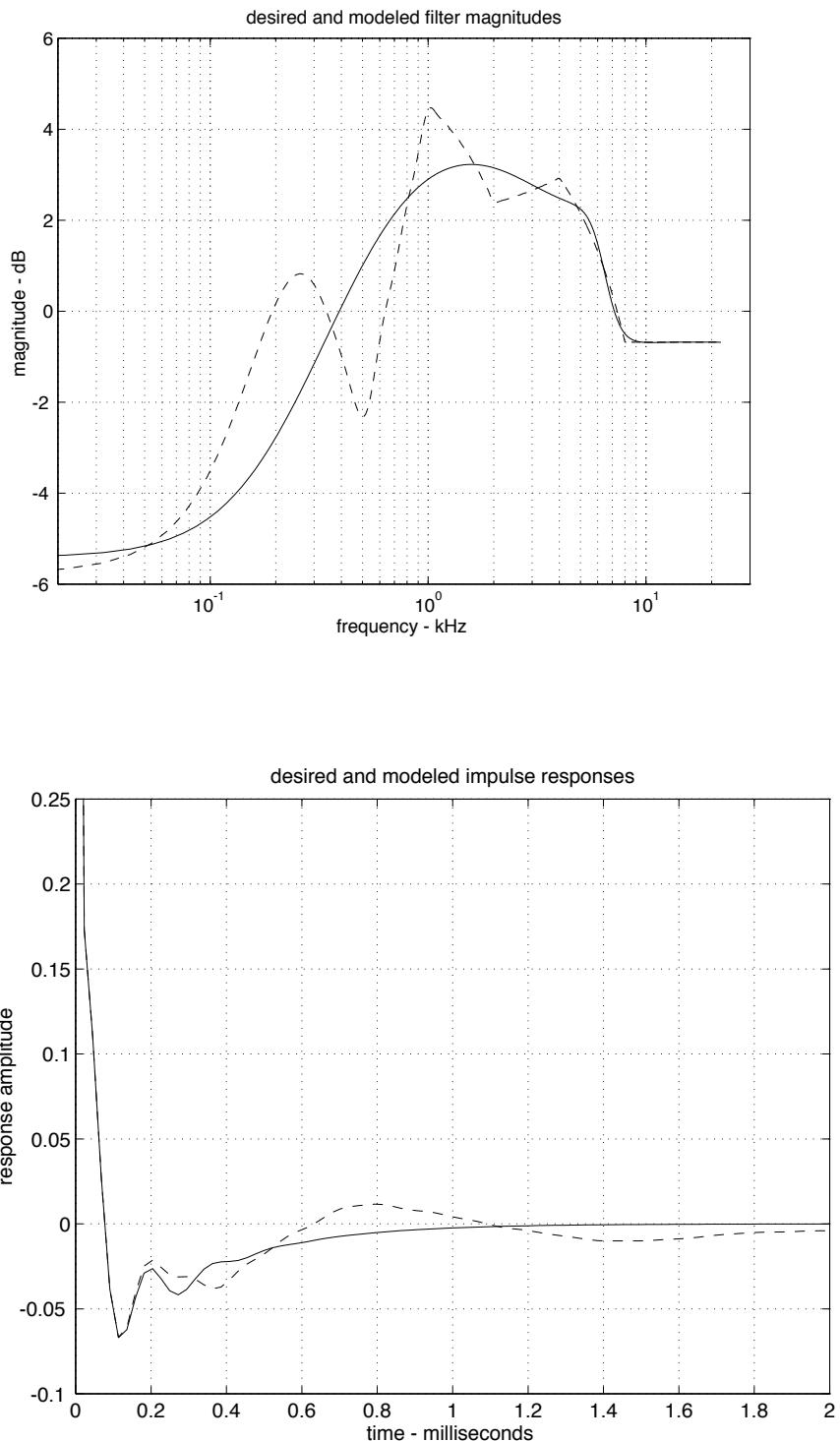
$$\hat{\mathbf{a}} = (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top \mathbf{h}.$$

- Note that the matrix $\mathbf{H}^\top \mathbf{H}$ and the column $\mathbf{H}^\top \mathbf{h}$ are composed of impulse response sample autocorrelations.

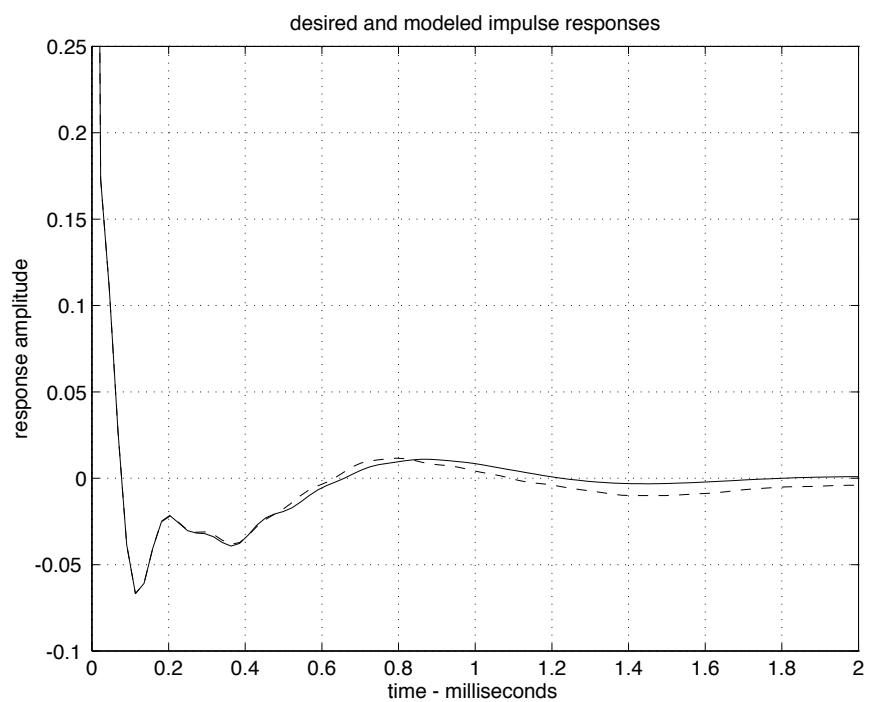
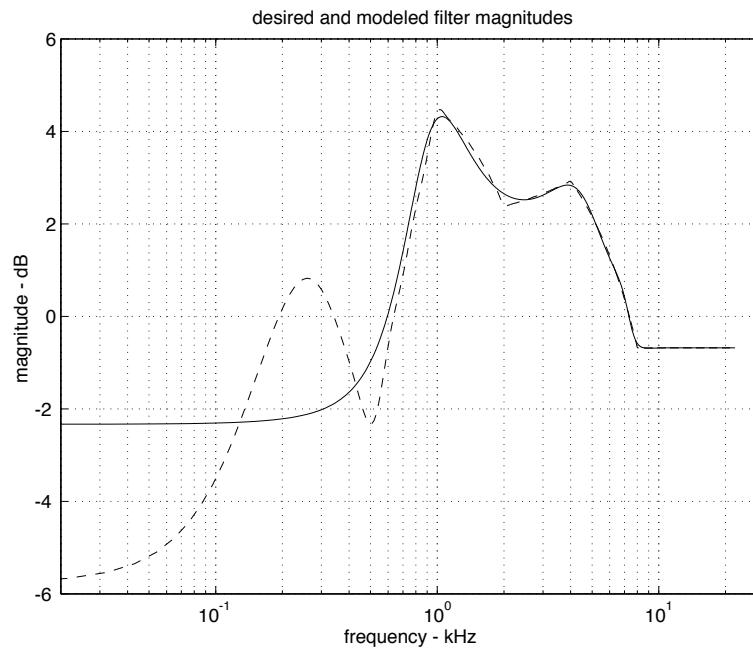
Prony's Method Example



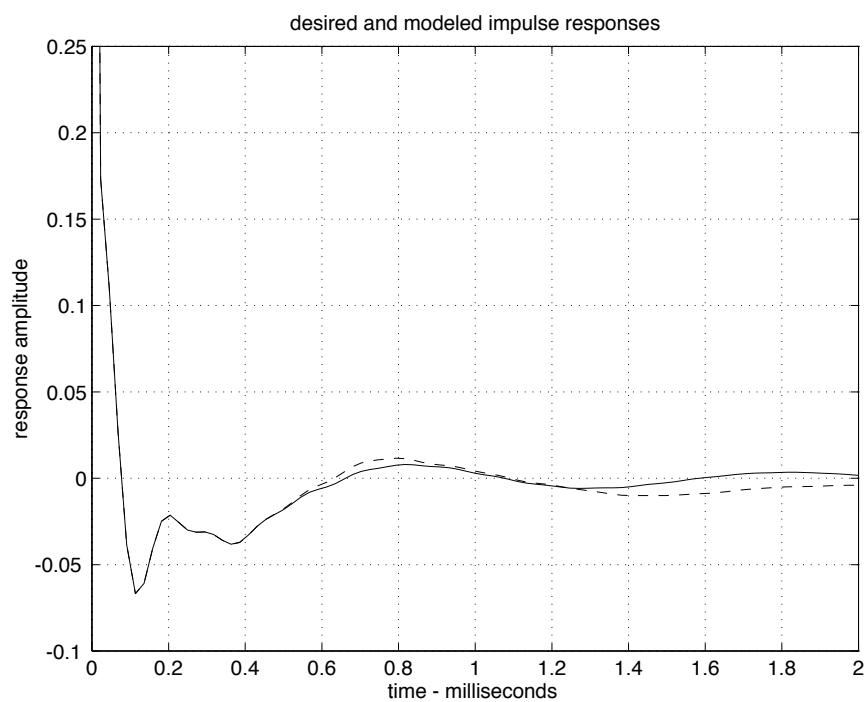
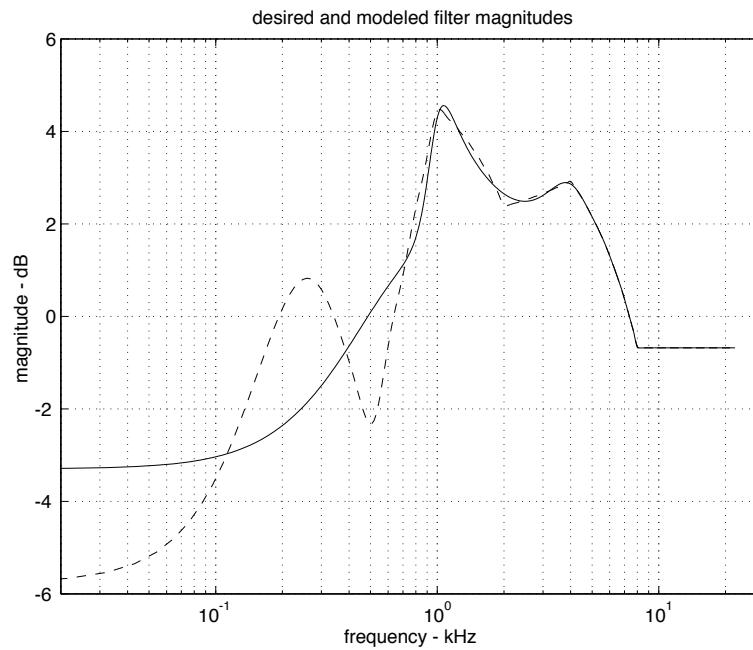
Prony's Method Example



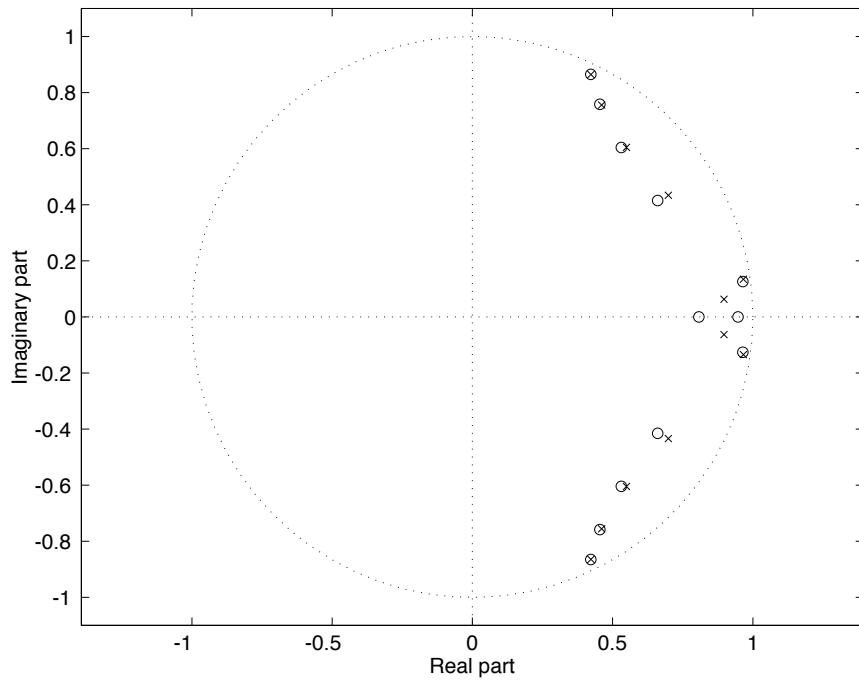
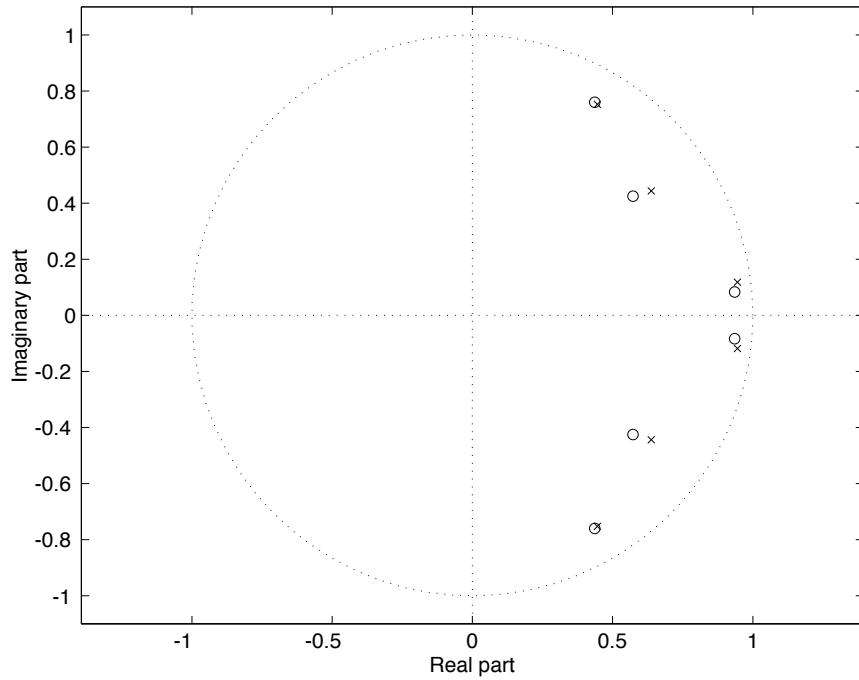
Prony's Method Example



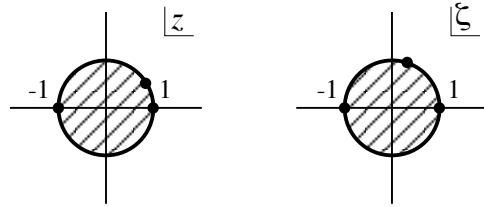
Prony's Method Example



Prony's Method Example



First-Order Conformal Map



- Consider the first-order *allpass transformation* (or *conformal map*) with parameter ρ , defined by the substitution

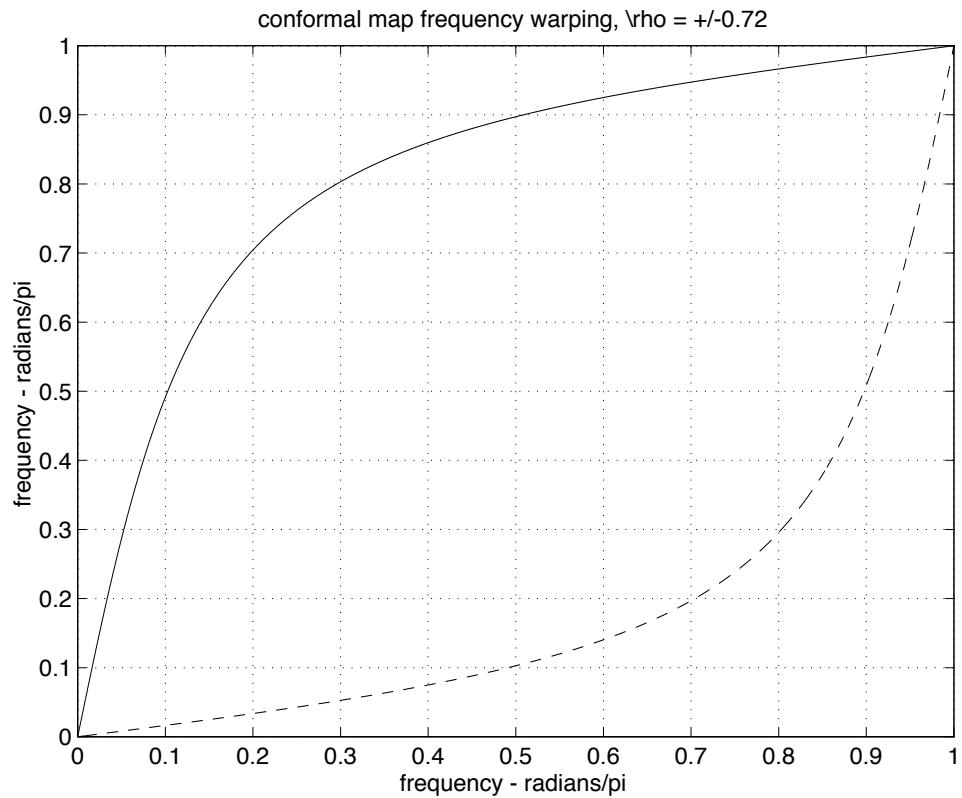
$$\zeta^{-1}(z) = \frac{\rho + z^{-1}}{1 + \rho z^{-1}}.$$

- Note that the transformation maps the unit circle to itself, preserving the locations of DC and the band edge, $z = \pm 1$,

$$|\zeta^{-1}(e^{j\omega})|^2 = \frac{(\rho + z^{-1})(\rho + z)}{(1 + \rho z^{-1})(1 + \rho z)} = 1.$$

- Points inside (outside) the unit circle on the z plane, map to points inside (outside) the unit circle on the ζ plane, $|\zeta| < 1$.

First-Order Conformal Map Frequency Warping



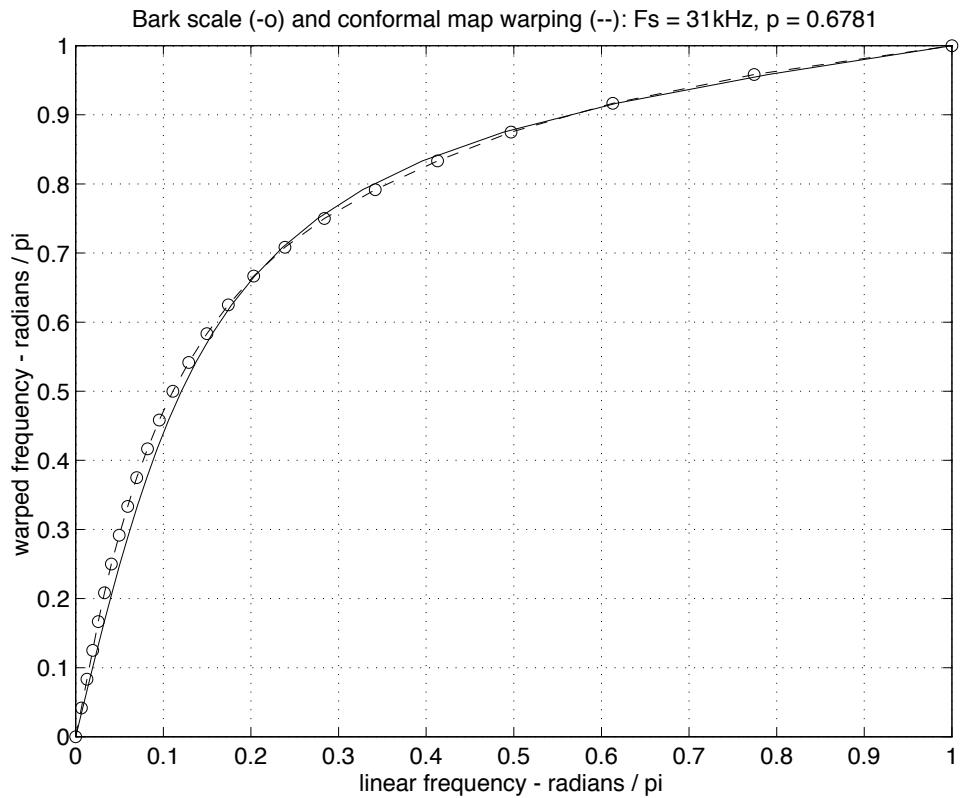
- Points $z^{-1} = e^{-j\omega}$ on the z -plane unit circle are warped to points $\zeta^{-1} = e^{-j\alpha}$ on the ζ -plane unit circle according to ρ ,

$$\rho = \frac{\sin[(\omega - \alpha)/2]}{\sin[(\omega + \alpha)/2]}.$$

- The mapping defined by $-\rho$ is the inverse of the mapping defined by ρ ,

$$z^{-1} = \zeta_{-\rho}^{-1}(\zeta_\rho(z^{-1})).$$

Allpass Parameter ρ Selection

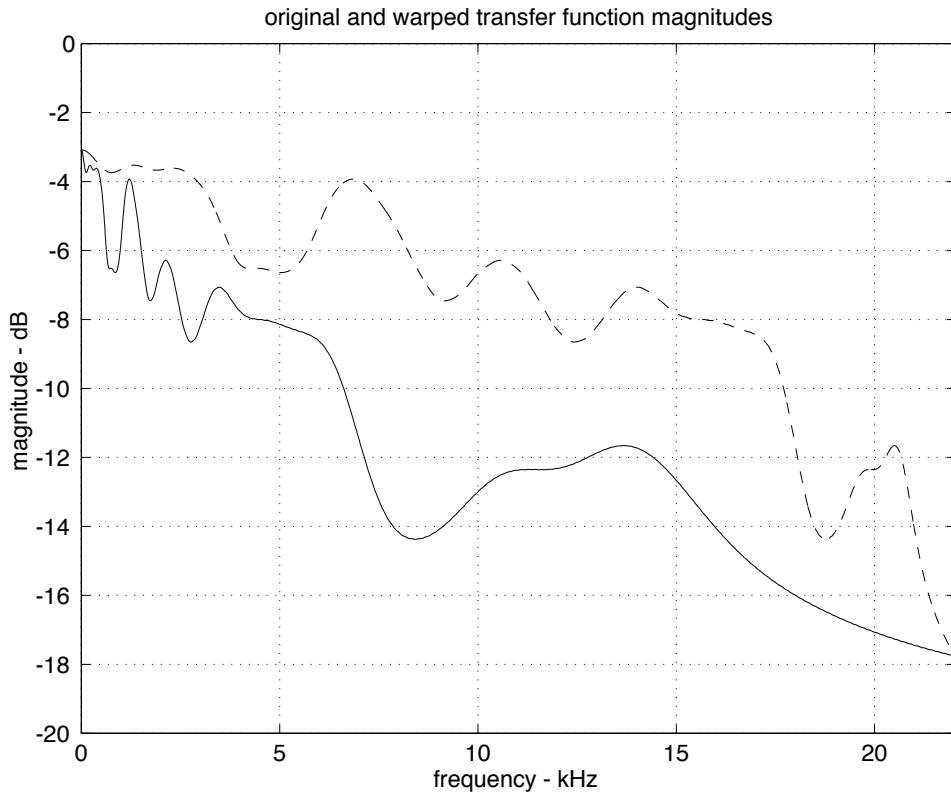


- For the proper choice of allpass parameter ρ , the first-order conformal map $\zeta_\rho^{-1}(z)$ imposes a frequency warping very much like either the Bark or ERB frequency scale.

$$\rho_{\text{Bark}}^*(f_s) = 1.07 \cdot \left[\frac{2}{\pi} \cdot \arctan(0.066f_s) \right]^{\frac{1}{2}} - 0.191$$

$$\rho_{\text{ERB}}^*(f_s) = 1.05 \cdot \left[\frac{2}{\pi} \cdot \arctan(0.072f_s) \right]^{\frac{1}{2}} - 0.196$$

Warped FIR Filter



- An FIR filter with impulse response $g(n)$ has z -transform

$$G(z) = \sum_{n=0}^{N-1} g(n) \cdot z^{-n}.$$

- By replacing the unit delay z^{-1} with the warped unit delay $\zeta_\rho^{-1}(z)$, the frequency axis of the original FIR filter is warped.

$$\Gamma_\rho(z) = G(\zeta_\rho(z)) = \sum_{n=0}^{N-1} g(n) \cdot \zeta_\rho^{-n}.$$

Warped IIR Filter

- A warped IIR filter has its pole and zero locations transformed.
- Consider a system with a single pole and single zero,

$$G(z) = \frac{1 + bz^{-1}}{1 + az^{-1}}.$$

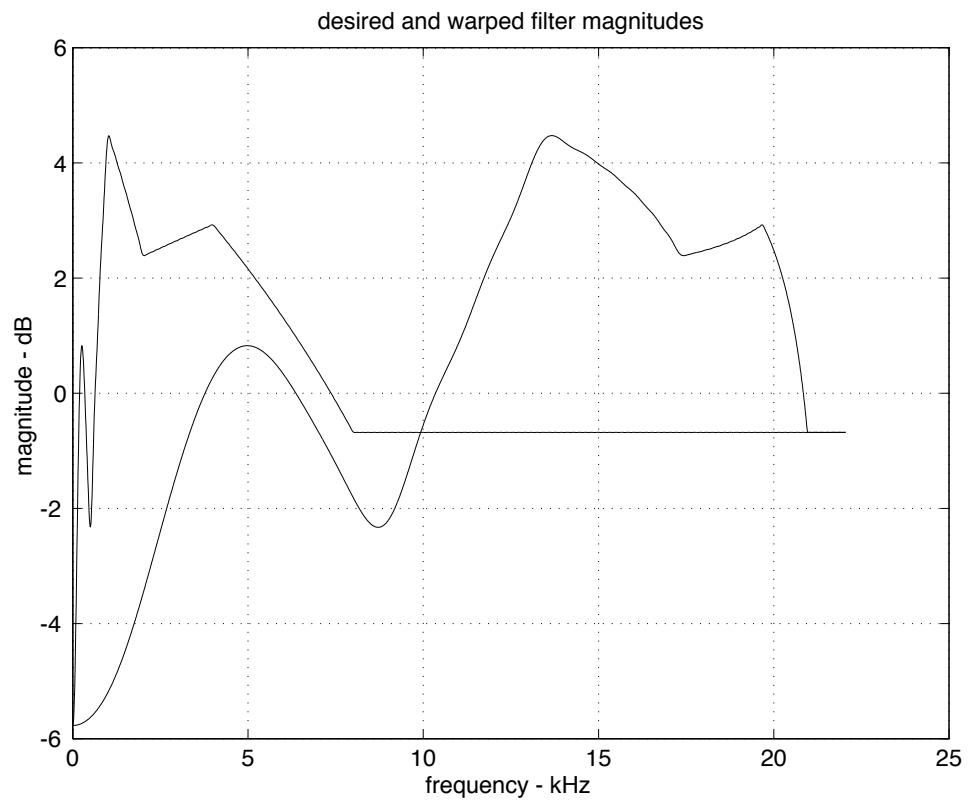
- Substituting $\zeta^{-1} = (\rho + z^{-1})/(1 + \rho z^{-1})$, we have

$$\begin{aligned}\Gamma(\zeta^{-1}(z)) &= \frac{1 + b\zeta^{-1}}{1 + a\zeta^{-1}} = \frac{1 + b(\rho + z^{-1})/(1 + \rho z^{-1})}{1 + a(\rho + z^{-1})/(1 + \rho z^{-1})}, \\ &= \frac{(1 + \rho z^{-1}) + b(\rho + z^{-1})}{(1 + \rho z^{-1}) + a(\rho + z^{-1})}, \\ &= \frac{1 + b\rho}{1 + a\rho} \cdot \frac{1 + [(\rho + a)/(1 + b\rho)]z^{-1}}{1 + [(\rho + a)/(1 + a\rho)]z^{-1}}.\end{aligned}$$

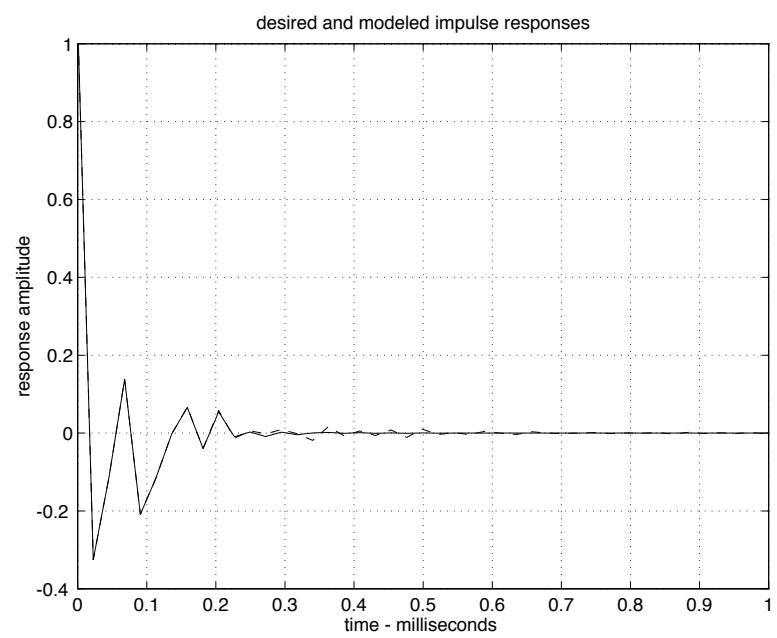
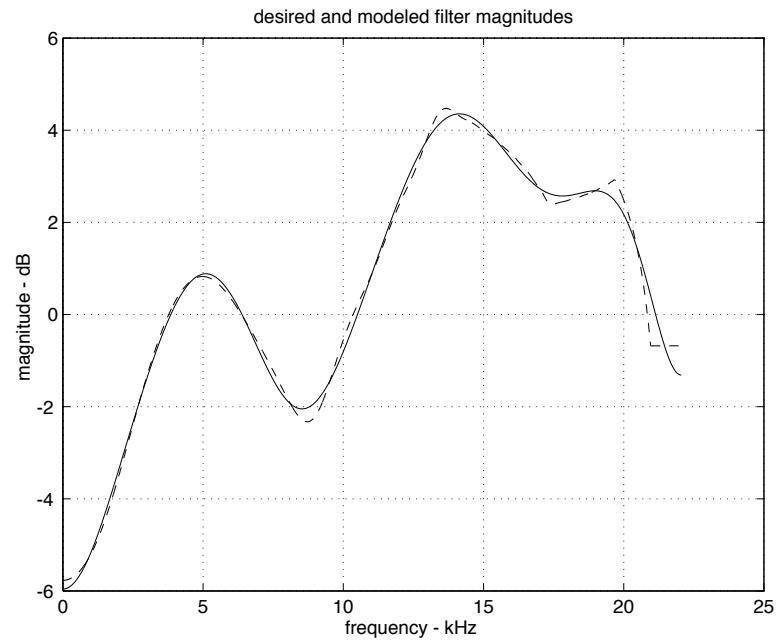
Warped Prony IIR Filter Design

- Frequency bands of interest may be given more weight by frequency warping the impulse response prior to fitting.
- Warped Prony filter design may be done as follows.
 - Warp the input impulse response using an allpass transformation with allpass parameter ρ set so as to assign a good amount of bandwidth to spectral features of importance.
 - Fit the system $\beta(z)/\alpha(z)$ to the warped impulse response.
 - Form $a(z)/b(z)$ by unwarping the fit to the warped impulse response via an allpass transformation using the allpass parameter $-\rho$.

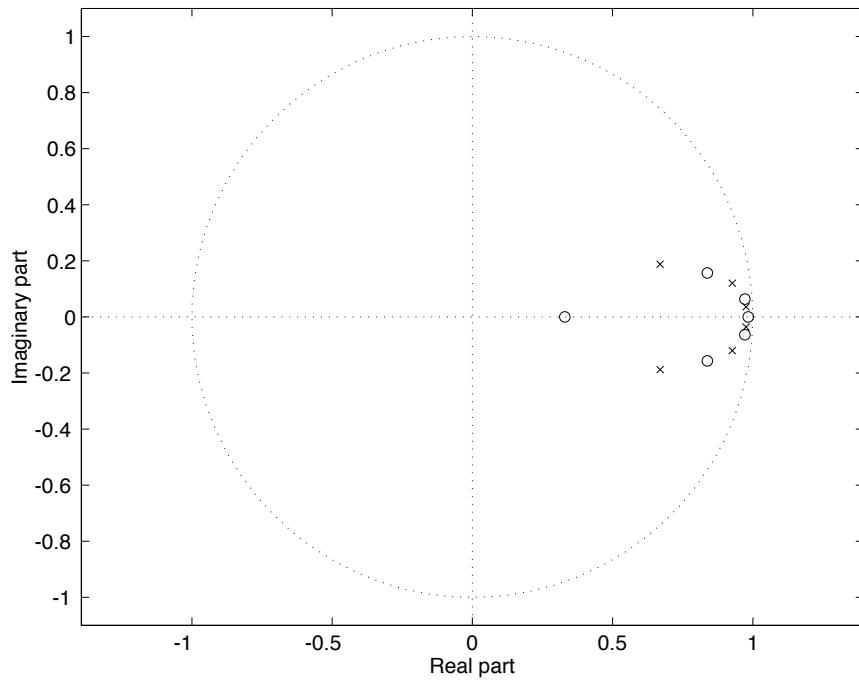
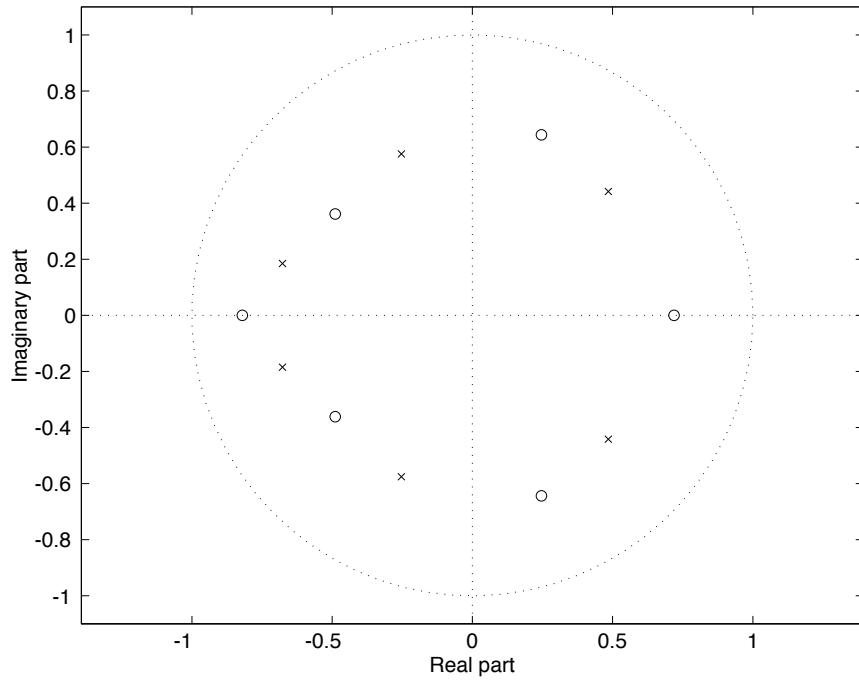
Warped Prony's Method Example



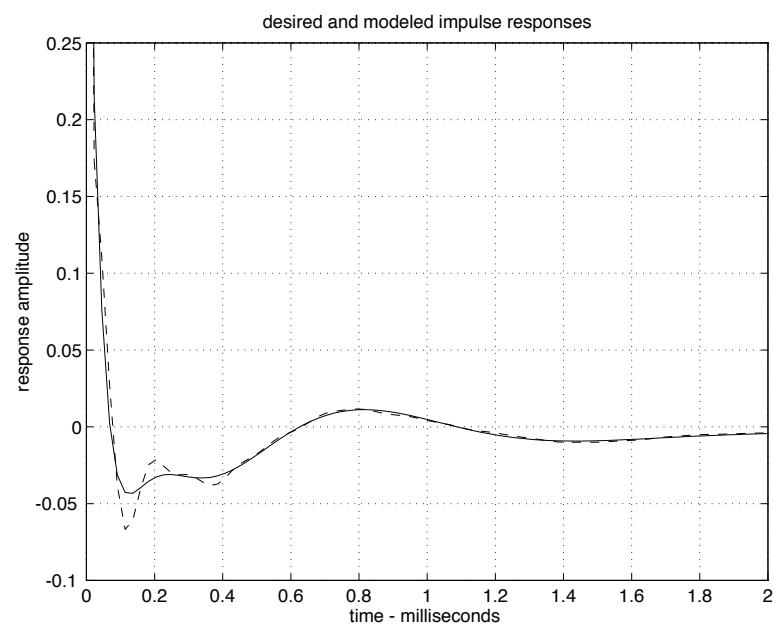
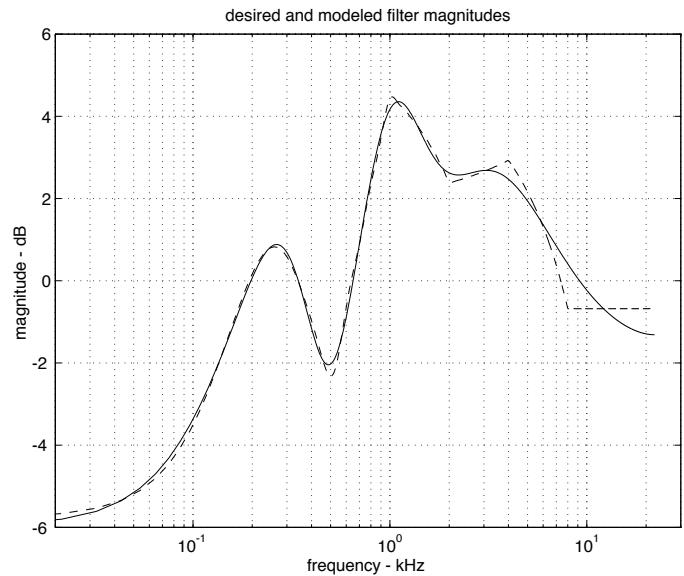
Prony's Method Example



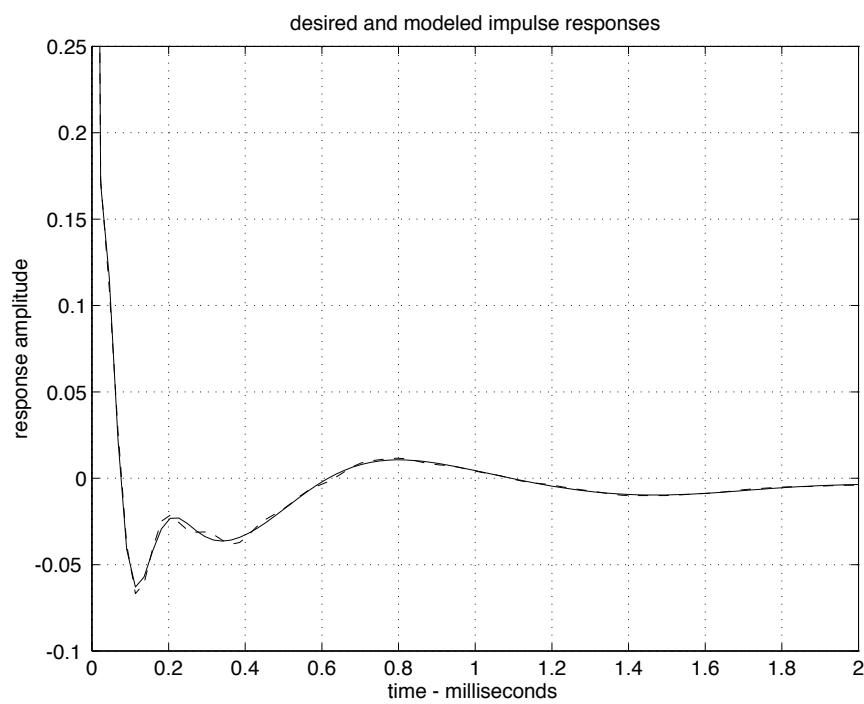
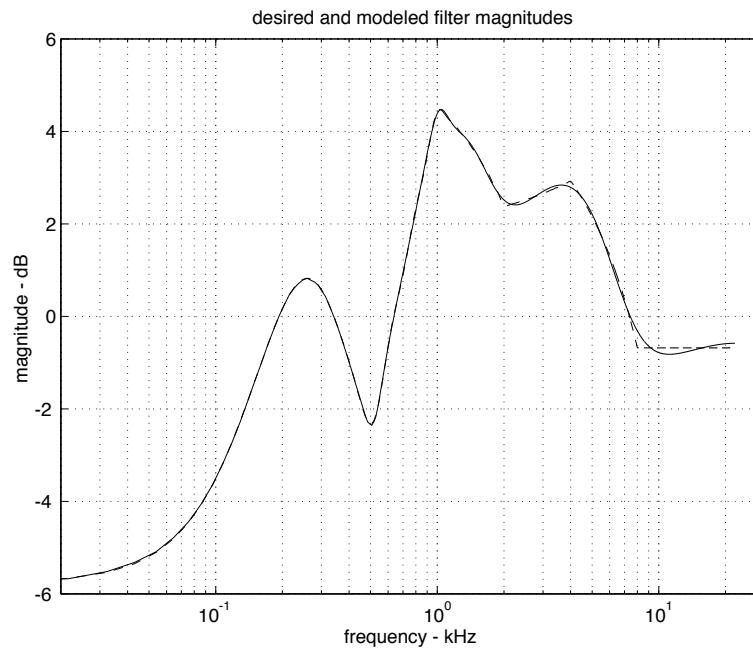
Warped Prony's Method Example



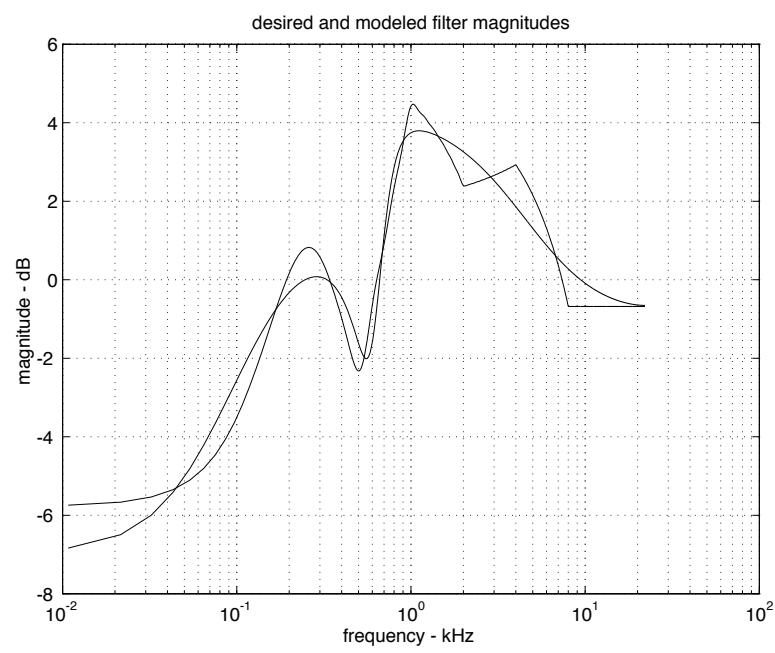
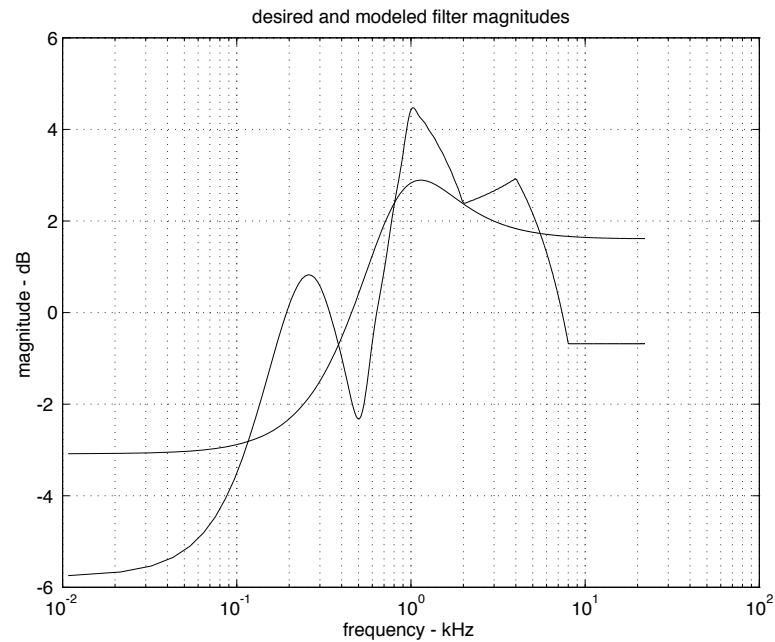
Prony's Method Example



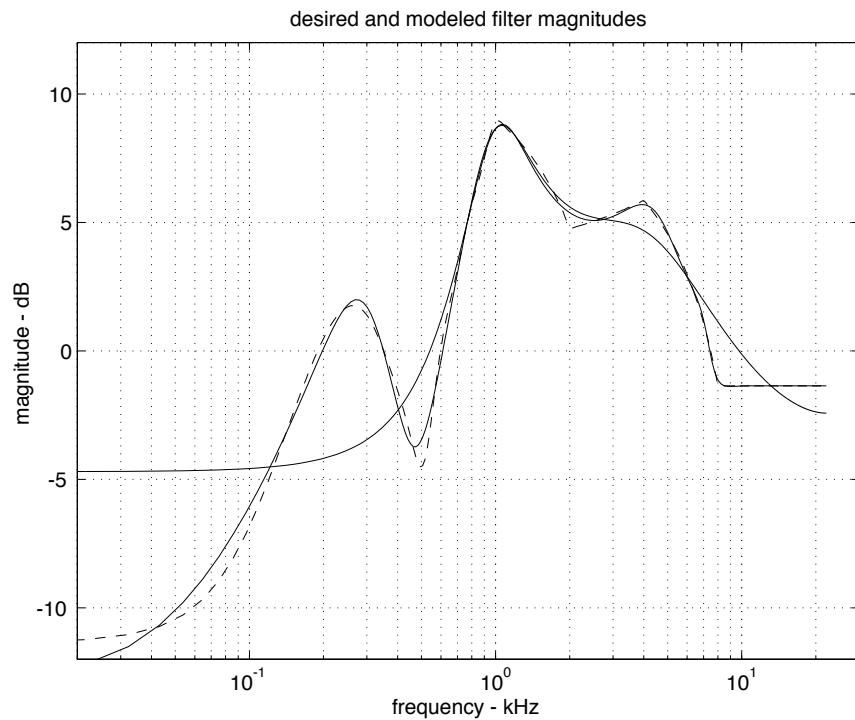
Warped Prony's Method Example



Warped Prony's Method Example



Warped Prony's Method Sound Example



25. Optimal Filter Design

Optimal Filtering

Lagrange Interpolation

- FIR filter
- Maximally flat at DC
 - match $e^{-j\delta\omega}$ (fractional delay δ) to as many derivatives as possible.
 - for n-tap filter, usually have $n/2 < \delta < n/2 + 1$ (n usually even).

Design:

$$h(n) = \sum_{n=0}^N b_n z^{-n} = \sum_{n=0}^N b_n e^{-j\omega n}$$

make $H(e^{j\omega})$ and its first N derivatives match $e^{-j\delta\omega}$ at $\omega = 0$.

$$\begin{aligned} b_0 + b_1 + \cdots + b_N &= 1 \\ -jb_1 - 2jb_2 - \cdots - Njb_N &= -j\delta \\ -b_1 - 4b_2 - 9b_3 - \cdots - N^2b_N &= -\delta^2 \\ &\vdots \\ (-j)^N b_1 + 2^N(-j)^N b_2 + \cdots + N^N(-j)^N b_N &= (-j)^N \delta^N \end{aligned}$$

Example: N=3

$$\begin{aligned} b_0 + b_1 + b_2 + b_3 &= 1 \\ -jb_1 - 2jb_2 - 3jb_3 &= -j\delta \\ -b_1 - 4b_2 - 9b_3 &= -\delta^2 \\ jb_1 + 8jb_2 + 27jb_3 &= j\delta^3 \end{aligned}$$

OR,

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \\ 0 & 1 & 8 & 27 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 1 \\ \delta \\ \delta^2 \\ \delta^3 \end{bmatrix}$$

$$\mathbf{V}\mathbf{b} = \begin{bmatrix} 1 \\ \delta \\ \delta^2 \\ \delta^3 \end{bmatrix}$$

where \mathbf{V} is the *Vandermonde* matrix and $\mathbf{b} = [b_0 \ b_1 \ b_2 \ b_3]^T$.

$$\mathbf{b} = \mathbf{V}^{-1} \begin{bmatrix} 1 \\ \delta \\ \delta^2 \\ \delta^3 \end{bmatrix}$$

To find our filter taps b_n , we need to invert \mathbf{V} . We can then use the elements a_{jk} of \mathbf{V}^{-1} to find

$$b_j = \sum_{k=0}^3 a_{jk} \delta^k$$

What are the elements of \mathbf{V}^{-1} ?

- find roots of polynomials in δ defined by a_{jk} .
 - for integer values of δ , we can find trivially the solutions for b :

$\delta = 0$:

$$b_0 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} = 1$$

$$\mathbf{b} = [1 \ 0 \ 0 \ 0]^T$$

$\delta = 1$:

$$b_0 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} = z^{-1}$$

$$\mathbf{b} = [0 \ 1 \ 0 \ 0]^T$$

$\delta = 2$:

$$\mathbf{b} = [0 \ 0 \ 1 \ 0]^T$$

$\delta = 3$:

$$\mathbf{b} = [0 \ 0 \ 0 \ 1]^T$$

$$b_0(\delta) = 0, \quad \delta = 1, 2, 3$$

$$b_1(\delta) = 0, \quad \delta = 0, 2, 3$$

$$b_2(\delta) = 0, \quad \delta = 0, 1, 3$$

$$b_3(\delta) = 0, \quad \delta = 0, 1, 2$$

But: Each of $b_n(\delta)$ is a cubic polynomial in δ . Therefore, each has exactly three zeros, which we have just located. So,

$$b_0 = (\delta - 1)(\delta - 2)(\delta - 3) \cdot k_0$$

$$b_1 = (\delta)(\delta - 2)(\delta - 3) \cdot k_1$$

$$b_2 = (\delta)(\delta - 1)(\delta - 3) \cdot k_2$$

$$b_3 = (\delta)(\delta - 1)(\delta - 1) \cdot k_3$$

where k_n are normalization coefficients chosen so that b_n is equal to one when δ is equal to n

Result:

We have found the elements of \mathbf{V}^{-1}

- rows of \mathbf{V}^{-1} are coefficients for power-series representations of b_n in δ :

$$\mathbf{V}_{ij}^{-1} = \frac{-1^{(3-i)}}{(3-i)!(i-1)!} \left[\text{coef of } x^j \text{ in } \frac{x(x-1)(x-2)(x-3)}{x-j} \right]$$

Again, normalization is picked s.t. $b_0(0) = b_1(1) = b_2(2) = b_3(3) = 1$.

- Polynomials b_i are Lagrange basis polynomials
 - have zeros at integers $i \neq j$
 - have value 1 at $i = j$

as $N \rightarrow$ large, these begin to look like sinc functions.

- sinc functions perfectly bandlimited
- Lagrange polynomials are minimum-order polynomials able to interpolate $[0, 0, 0, \dots, 0, 1, 0, \dots, 0, 0, 0]$.

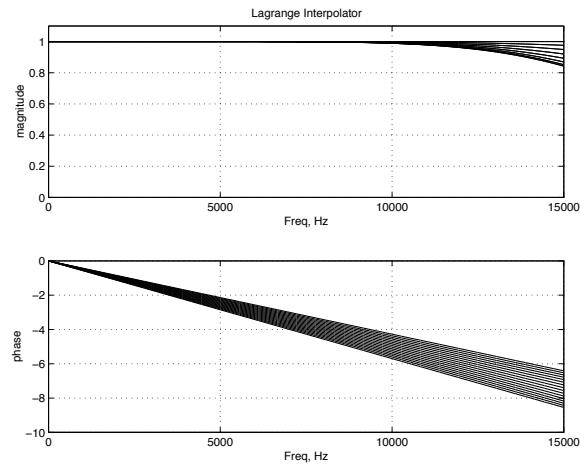


Figure 111: *Lagrange Filters.*

Butterworth filter:

- maximally flat IIR lowpass filter
- all-pole
- continuous-time

$$H(s) = \frac{1}{\sum_{n=0}^N a_n s^n}$$

- asymptotic N-pole rolloff at very high frequency
- has $N - 1$ derivatives equal to zero at DC

if $f(x)$ has $N - 1$ derivatives equal to zero, then $1/f(x)$ also has $N - 1$ derivatives equal to zero:

Chain Rule:

$$\left[\frac{1}{f(x)} \right]' = -\frac{f'(x)}{f^2(x)}$$

Pick

$$\left| H(s) \right|^2 = \frac{1}{1 + \left(\frac{\omega^2}{\omega_n^2} \right)^N}$$

if $|H(s)|^2$ has $2N - 1$ derivatives equal to zero, then $H(s)$ will have $N - 1$ derivatives equal to zero.

$$H(s)H(-s) = \frac{1}{1 + \left(\frac{-s^2}{\omega_n^2} \right)^N}$$

- poles at N^{th} roots of $-\omega_n$

- pick poles in left-half plane for stability
- cutoff frequency is ω_0 (-3dB point)

Discretization:

$$s \rightarrow \frac{2}{T_d} \left(\frac{1 - z^{-1}}{1 + z^{-1}} \right)$$

- flatness at DC preserved
- get N zeros at $z = -1$
- gives N degrees of flatness at Nyquist limit
- can move zeros elsewhere to give up flatness at Nyquist and get more flatness at DC

Time-varying filters:

- go from $H_1 \rightarrow H_2$ smoothly over time T

Idea:

let each coefficient evolve smoothly from $a_1 \rightarrow a_2$, $b_1 \rightarrow b_2$

question:

where do poles and zeros go during transition?

- need to preserve stability for intermediate filters

For biquad:

Complex roots:

$1 + 2r \cos(\theta)z^{-1} + r^2z^{-2} \rightarrow$ roots at radius r

Real roots:

$1 - (a + b)z^{-1} + abz^{-2} \rightarrow$ roots at a, b

so, if $|a - 2| < 1$ and $|a_1| < |1 + a_2|$ we will have roots inside the unit circle.

This coefficient space is *convex*:

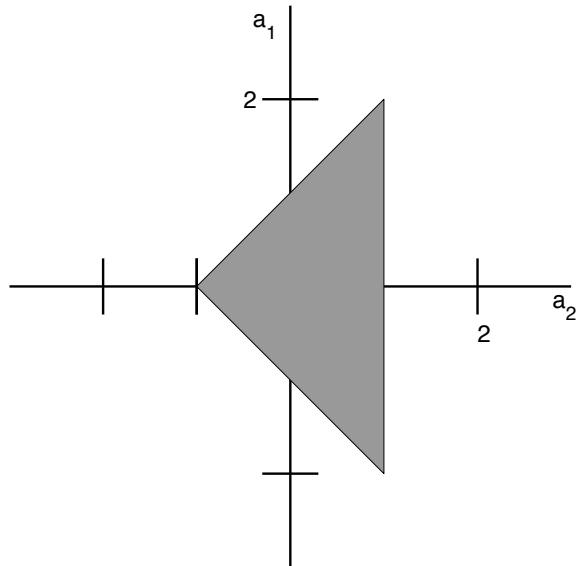


Figure 112: Region of stability for coefficients.

- if starting and ending filters are stable, intermediate filters will also be stable, regardless of coefficient trajectories, as long as all coefficients interpolate starting and ending values using the same function (e.g. linear, exponential, etc).
- for roots inside the unit circle, coefficients are limited to the following values: $|a_1| < 2, |a_2| < 1$.

FIR smoothing:

$$\sum a_n z^{-n} \rightarrow \sum b_n z^{-n}$$

if all coefficients are smoothed together, it is equivalent to running both filters in parallel and crossfading outputs.

- used for HRTF models

Delay and Filtering

26. Parametrized Filters and Discretization

Parametrized Filters

- First-order shelf,

$$H(s) = \frac{s + \beta}{s + \alpha}$$

- First-order allpass,

$$H(s) = \frac{s - \alpha}{s + \alpha}$$

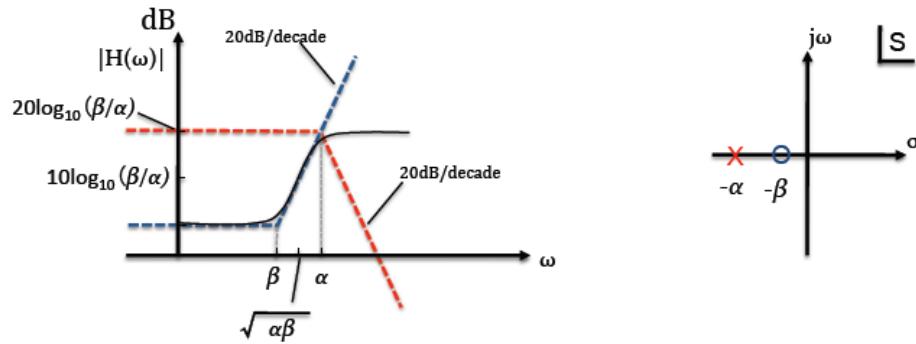
- Resonant low-pass, high-pass, peak

$$H(s) = \frac{1}{s^2 + \frac{1}{Q}s + 1}, \frac{s^2}{s^2 + \frac{1}{Q}s + 1}, \frac{s}{s^2 + \frac{1}{Q}s + 1}$$

- Parametric section,

$$H(s) = \frac{s^2 + \frac{\gamma}{Q}s + 1}{s^2 + \frac{1}{Q}s + 1}$$

First-Order Shelf Filter



- Low shelf,

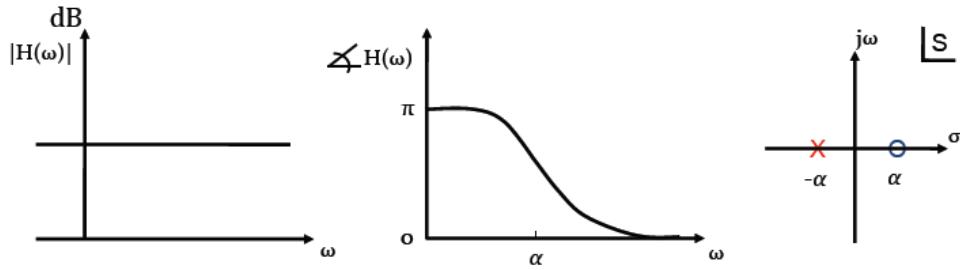
$$H(s) = \frac{s + \beta}{s + \alpha}$$

- High shelf,

$$H(s) = \frac{\beta s + 1}{\alpha s + 1}$$

- DC and high-frequency gains specified.
- At the geometric mean of the pole and zero frequencies, the filter gain is the geometric mean of the low- and high-frequency gains.

First-order Allpass Filter

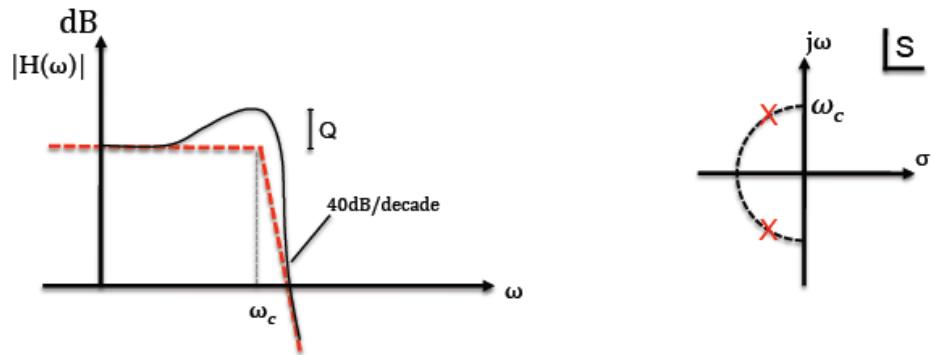


- The first-order allpass filter maintains unit gain across the band,

$$H(s) = \frac{s - \alpha}{s + \alpha}$$

- Group delay concentrated about α .

Resonant Low-Pass Filter

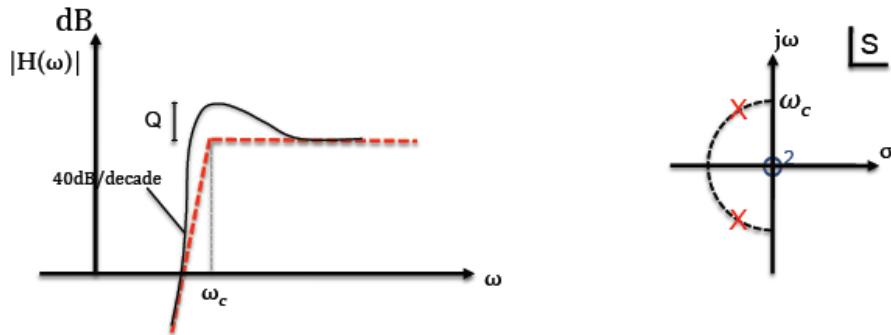


- Resonant low-pass,

$$H(s) = \frac{1}{(s/\omega_c)^2 + \frac{1}{Q}(s/\omega_c) + 1}$$

- ω_c is the cutoff frequency.
- Q determines the height of the resonant peak.
- 12dB/octave roll off.

Resonant High-Pass Filter

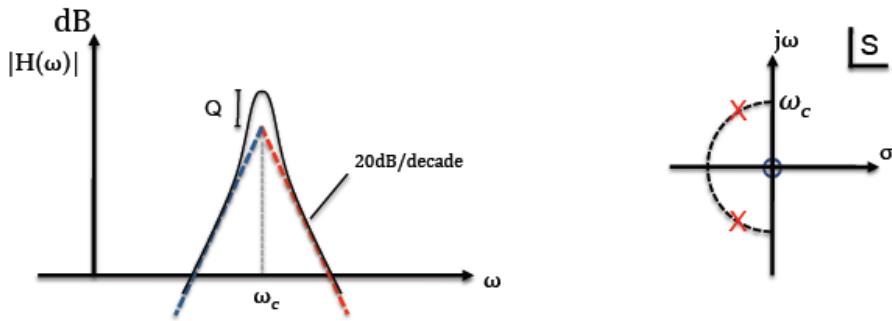


- Resonant high-pass,

$$H(s) = \frac{(s/\omega_c)^2}{(s/\omega_c)^2 + \frac{1}{Q}(s/\omega_c) + 1}$$

- ω_c is the cutoff frequency.
- Q is the filter gain at ω_c .
- 12dB/octave roll off.

Resonant Peak Filter

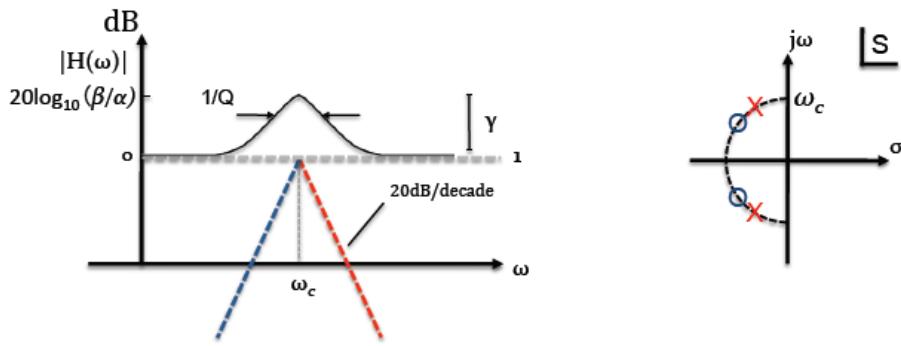


- Resonant peak,

$$H(s) = \frac{(s/\omega_c)}{(s/\omega_c)^2 + \frac{1}{Q}(s/\omega_c) + 1}$$

- ω_c is the cutoff frequency.
- Q is the peak gain.
- 6dB/octave rolloff away from ω_c .

Parametric Section



- Parametric section,

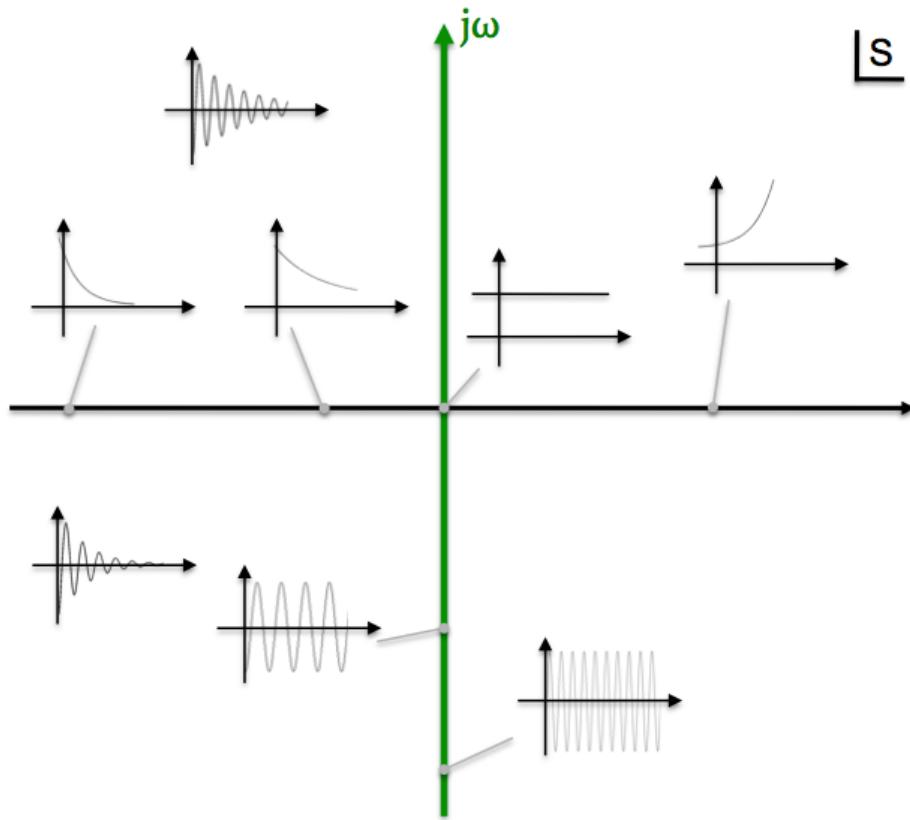
$$H(s) = \frac{(s/\omega_c)^2 + \frac{\gamma}{Q}(s/\omega_c) + 1}{(s/\omega_c)^2 + \frac{1}{Q}(s/\omega_c) + 1}$$

- Provides boost/cut of γ in a region about ω_c , the center frequency and a gain approaching 1 for frequencies away from ω_c .
- Q determines the width of the boost/cut region.
- Note: To make a cut filter that complements a boost filter, move the gain γ to the denominator.

Filter Discretization

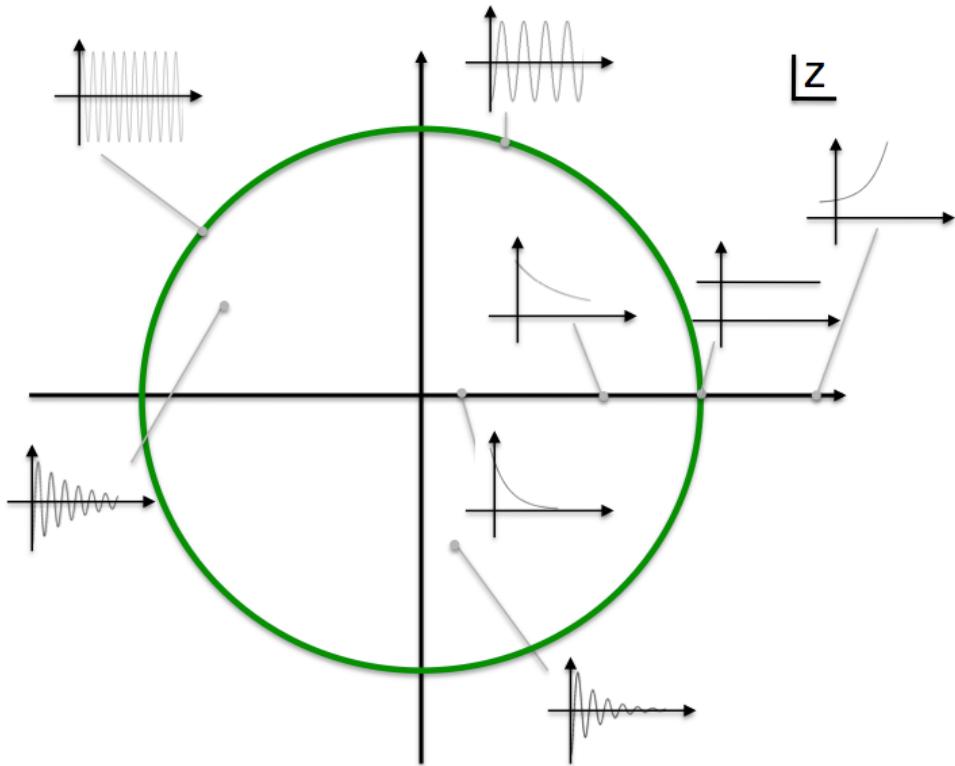
- s-plane structure.
- z-plane structure.
- Bilinear transform
 - definition
 - interpretation
 - properties

s-Plane Structure



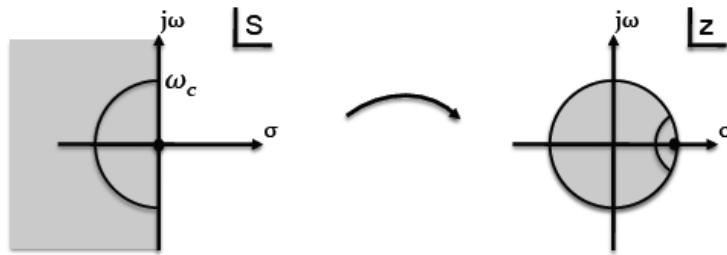
- The basic operation is the derivative, s .
- Defined by exponentials e^{st} , which when scaled by s form the exponential derivative, $s \cdot e^{st}$.

z-Plane Structure



- The basic operation is the delay, z^{-1} .
- Defined by exponentials z^n , that when scaled by z^{-1} delay the signal one sample, z^{n-1} .

Bilinear Transform



- Bilinear transform,

$$s = \frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}}$$

- Trapezoid rule for integration.
- $j\Omega$ s -plane axis mapped to the unit circle $e^{j\omega}$.
- Left-half plane mapped to inside the unit circle.
- Preserves model order.

Discretization Example

- First-order low-pass,

$$H(s) = \frac{1}{s + \alpha}$$

- Substitute

$$\begin{aligned} s &= \frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}} \\ H(z) &= \frac{1}{\frac{\frac{2}{T} \frac{1-z^{-1}}{1+z^{-1}}}{1+z^{-1}} + \alpha} \\ &= \frac{1 + z^{-1}}{1 + \frac{\alpha - \frac{T}{2}}{\alpha + \frac{T}{2}} z^{-1}} \cdot \frac{1}{\alpha + \frac{T}{2}} \end{aligned}$$

- Model order preserved.
- Poles, zeros warped according to the bilinear transform.

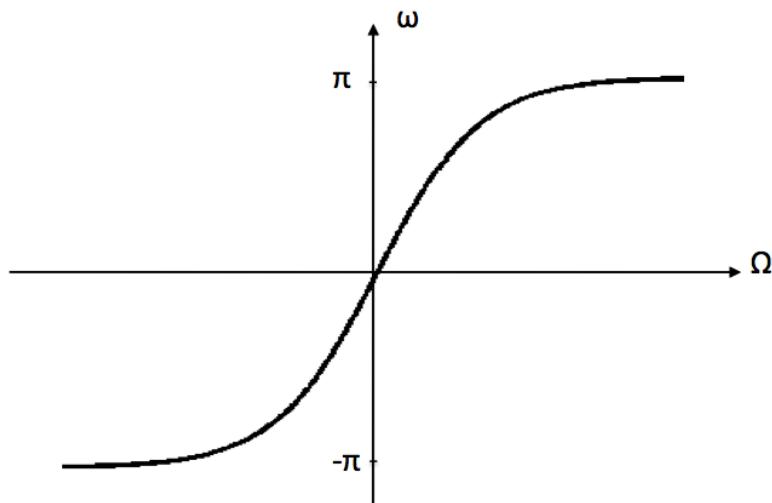
Frequency Warping

- Bilinear transform

$$s = \frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}}$$

- Evaluate the bilinear transform on the unit circle, $z^{-1} = e^{-j\omega}$

$$\longrightarrow \Omega = \frac{2}{T} \cdot \tan(\omega/2)$$

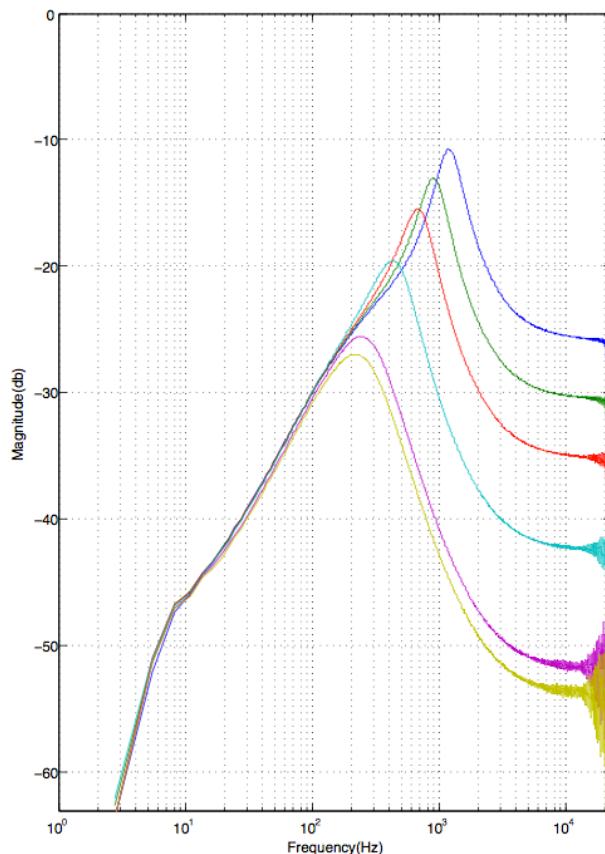


- Analog and digital frequencies which are small compared to the sampling rate match,

$$\Omega \approx \frac{\omega}{T}, \quad |\omega| \ll 1$$

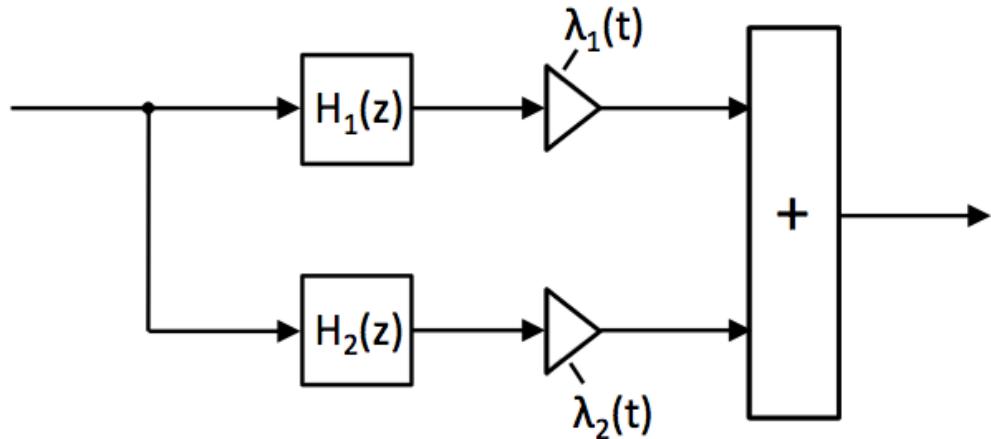
27. Time-Varying Filtering

Time-Varying Filter Example - Wah Pedal



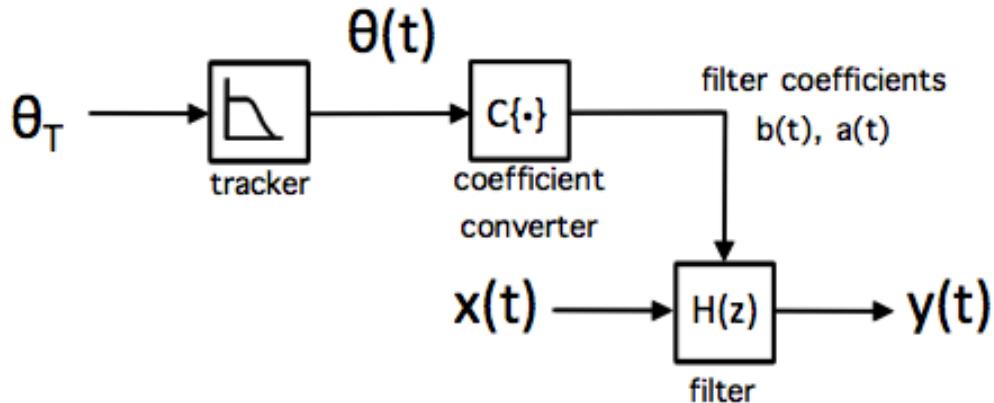
- Dunlop 535Q CryBaby transfer function magnitudes, various pedal positions.

Filter Crossfade



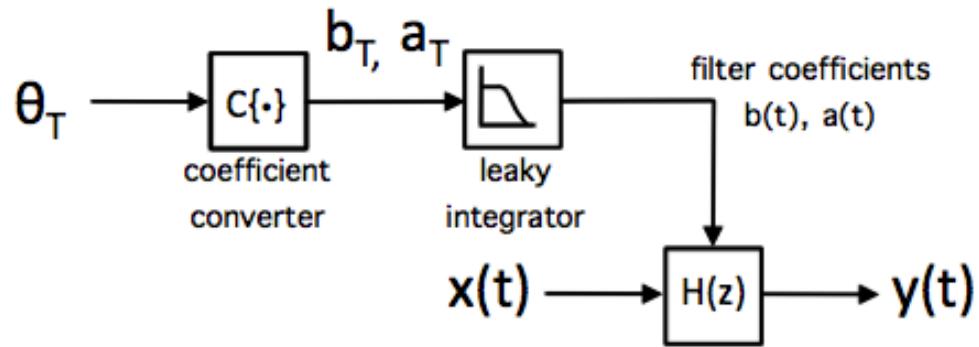
- Crossfade between successive filters.
- Use linear crossfade between filters with similar phase.
- Start crossfade after filter transient.

Smoothed Filter Parameters



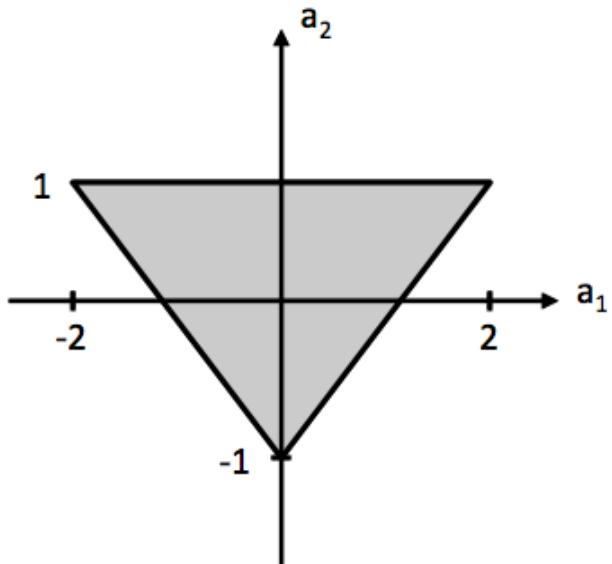
- Generate slowly changing filter coefficients by converting smooth parameter trajectories.

Smoothed Filter Coefficients



- Generate time varying filter coefficients by tracking coefficient targets.

Stability Region



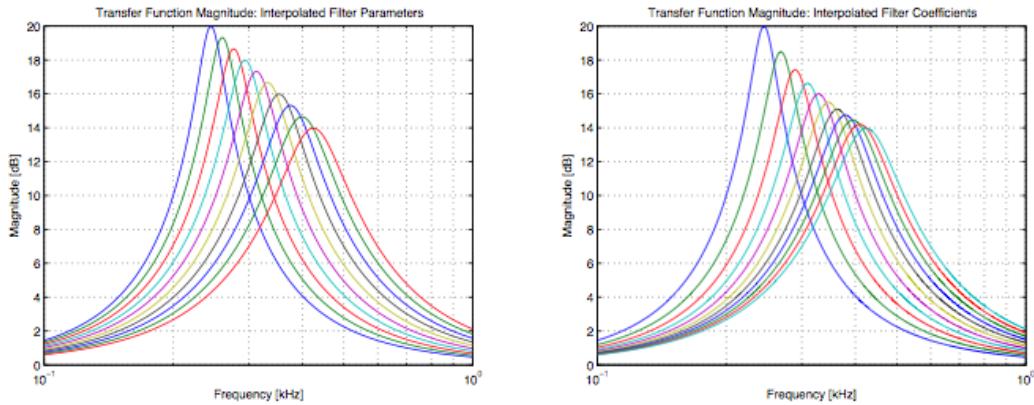
- The polynomial

$$a(z) = 1 + a_1 z^{-1} + a_2 z^{-2}$$

will have stable roots if a_1, a_2 lie in the shaded region above.

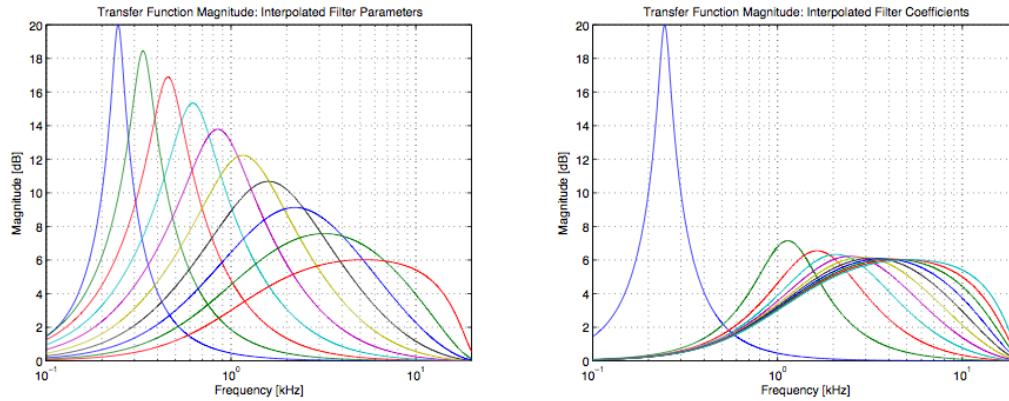
- Linearly interpolating between stable coefficient sets produce stable intermediate filters.

Parameter, Coefficient Interpolation



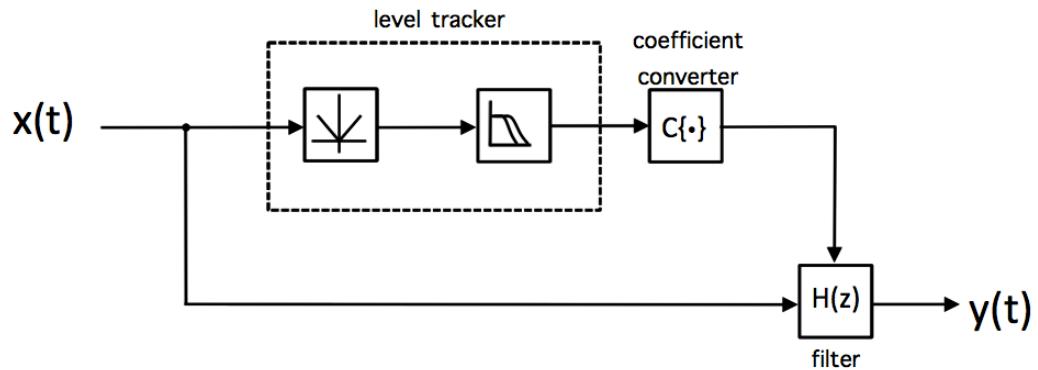
- When starting and ending filters are similar, the intermediate filters produced by parameter interpolation and coefficient interpolation will also be similar.

Parameter, Coefficient Interpolation



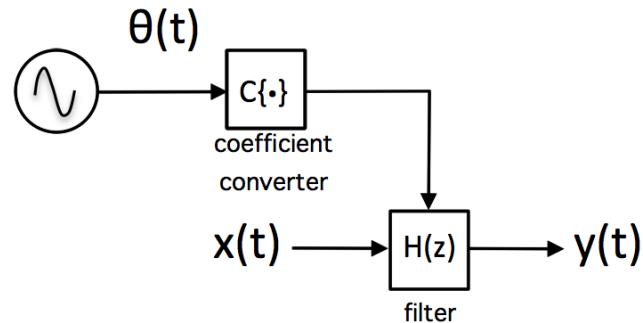
- When the parameters are very different, the intermediate filters produced by parameter and coefficient interpolation are likely to be dissimilar.

Envelope Filter



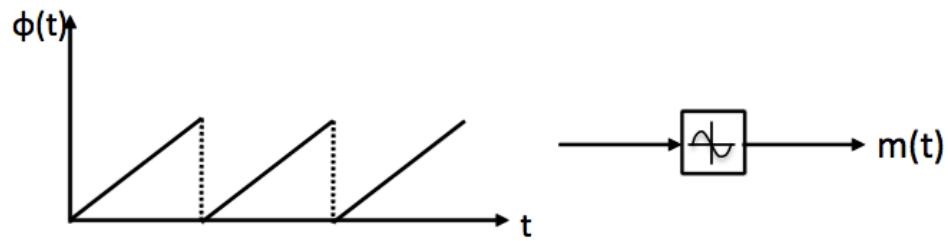
- Filter adjusted based on signal level.
- Example - auto wah.

LFO - Controlled Filter



- The filter parameters may be controlled by an LFO, “low-frequency oscillator”.
- Example - phasor.

LFO Implementation



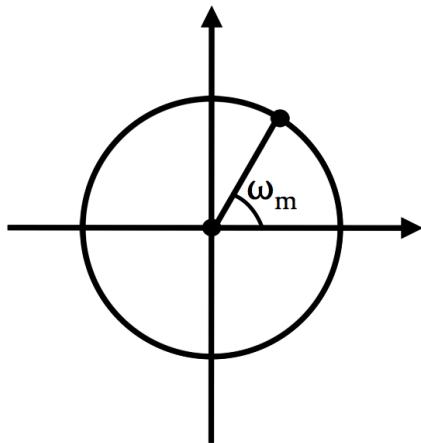
- An LFO may be implemented by generating a ramp according to a frequency ω_m ,

$$\varphi(n) = \text{rem}(\varphi(n) + \omega_m, 2\pi)$$

and applying a transformation, e.g.,

$$m(n) = |\varphi(n)|$$

Magic Circle



- Sinusoids may be efficiently generated by iteratively multiplying by $e^{j\omega_m}$:

$$x(n) = \cos \omega_m \cdot x(n - 1) + \sin \omega_m \cdot y(n - 1)$$

$$y(n) = -\sin \omega_m \cdot x(n - 1) + \cos \omega_m \cdot y(n - 1)$$

- Magic circle

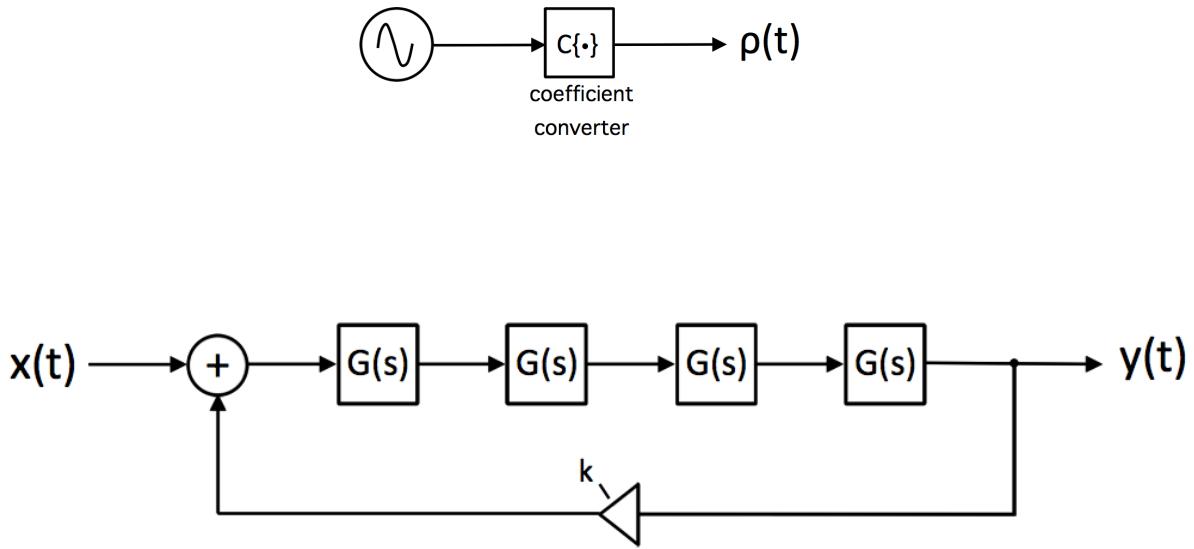
$$x(n) = x(n - 1) + \epsilon \cdot y(n - 1)$$

$$y(n) = -\epsilon \cdot x(n) + y(n - 1)$$

$$\epsilon = 2 \sin(\omega/2)$$

- Numerically robust—forms closed, nearly circular loops with 12-bit integer arithmetic.

Example - Phasor



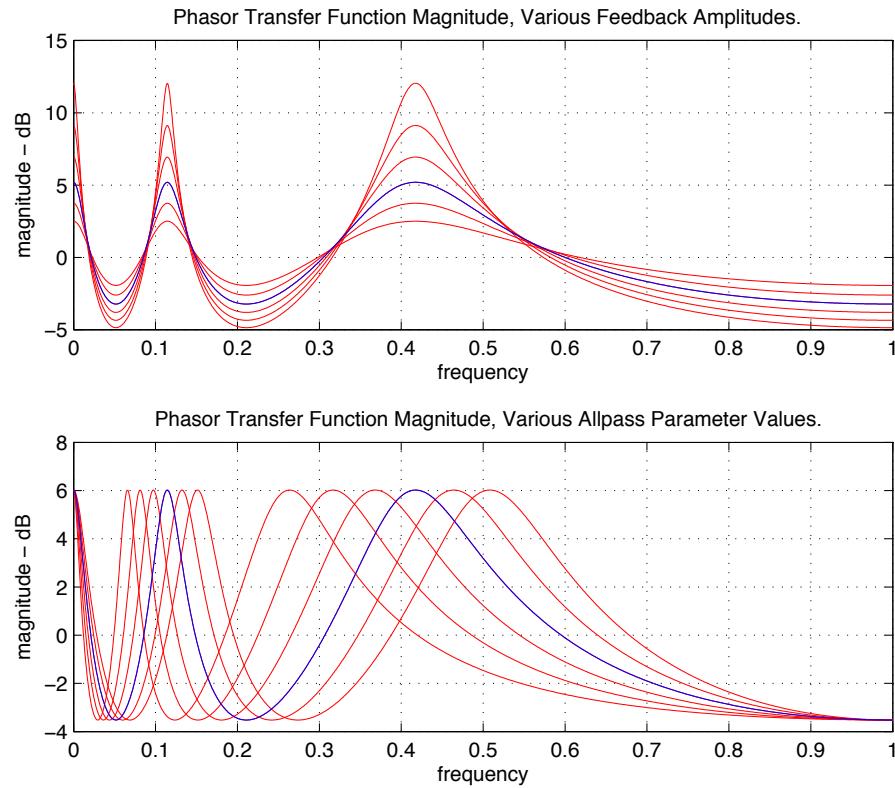
- $G(s)$ is an allpass filter parametrized by ρ ,

$$G(s) = \frac{s - \rho}{s + \rho}$$

- The transfer function is given by

$$H(s) = \frac{G^4(s)}{1 - kG^4(s)}$$

Example - Phasor



- Phasor transfer function magnitudes, varying ρ and k .

Example - Phasor

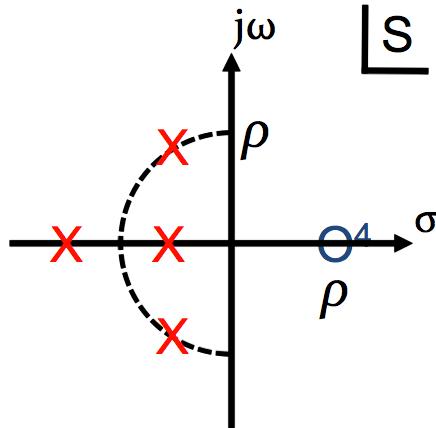
- Phasor transfer function,

$$H(s) = \frac{G^4(s)}{1 - kG^4(s)}$$

$$= \frac{(s - \rho)^4}{(s + \rho)^4 - k(s - \rho)^4}$$

- Phasor poles

$$s_n = -\rho \cdot \frac{1 + \sigma_n}{1 - \sigma_n}, \quad \sigma_n = k^{1/4} \cdot e^{j\frac{\pi}{2}n}, \quad n = 0, 1, 2, 3.$$

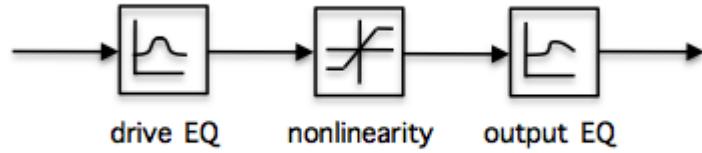


- Biquad factorization

$$H(s) = \frac{(s - \rho)^2}{\left(s + \rho \frac{1+k^{1/4}}{1-k^{1/4}}\right) \left(s + \rho \frac{1-k^{1/4}}{1+k^{1/4}}\right)} \cdot \frac{(s - \rho)^2}{s^2 + \rho \cdot s \cdot \frac{1-k^{1/2}}{1+k^{1/2}} + \rho^2}$$

28. Distortion Processing

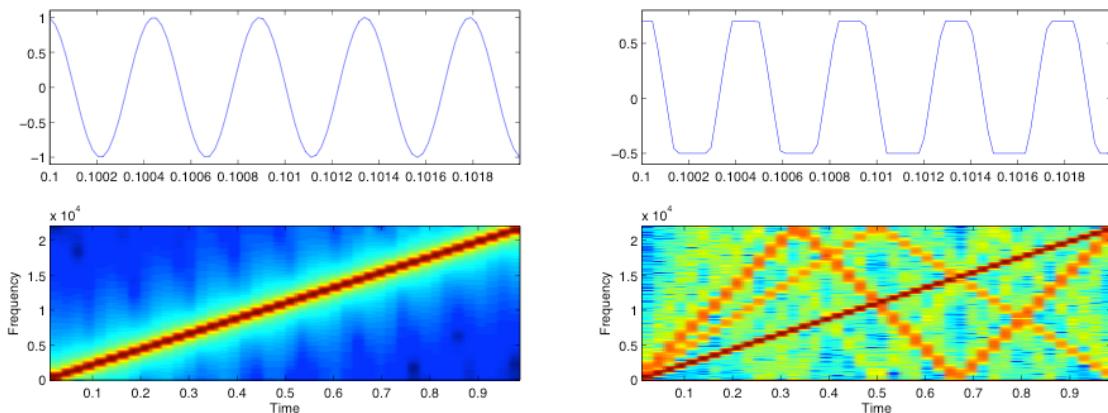
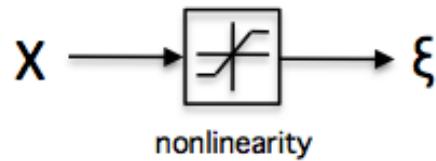
Signal Model Distortion Processor



$$y(t) = q(t) * \nu(x(t) * c(t))$$

- Applying a memoryless nonlinearity and filtering in cascade provides guitar amp-like sounds.
- Setting $c(t)$ to be a differentiator and $q(t)$ to be an integrator, and $\nu(\cdot)$, a saturating nonlinearity, a slew rate limiter results.

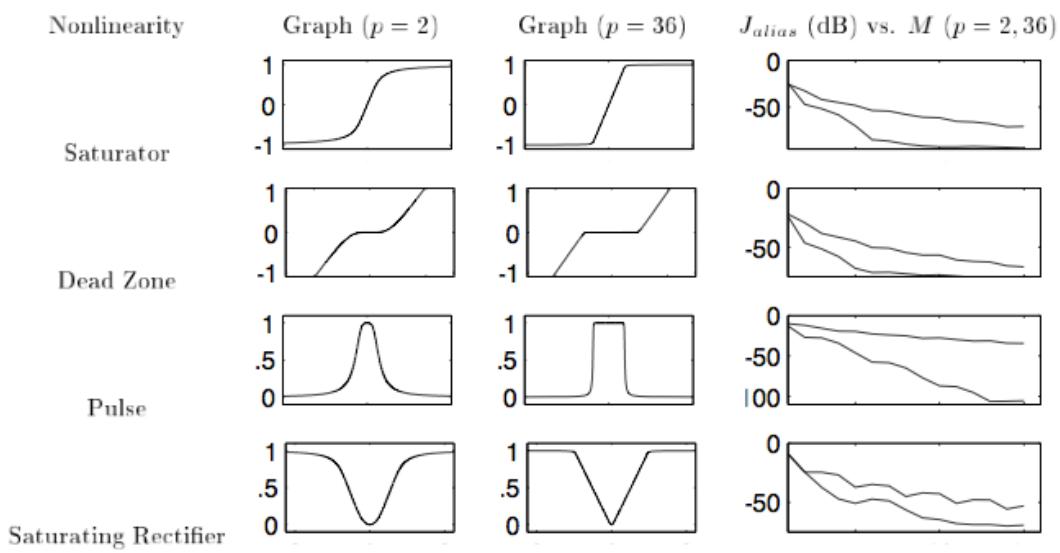
Memoryless Nonlinearity



- Memoryless nonlinearities can greatly increase the bandwidth of the input, providing aliasing.

Memoryless Nonlinearity Examples

Saturator	$y = S_p(x)$	$x = y + y^{2p} \gamma_p(y)$
Dead Zone	$y = D_p(x)$	$y = x - S_p(x)$
Pulse	$y = P_p(x)$	$y = \frac{d}{dx} S_p(x)$
Saturating Rectifier	$y = R_p(x)$	$y = 1 + \frac{(S_p(x-1/2) - S_p(x+1/2))}{2}$

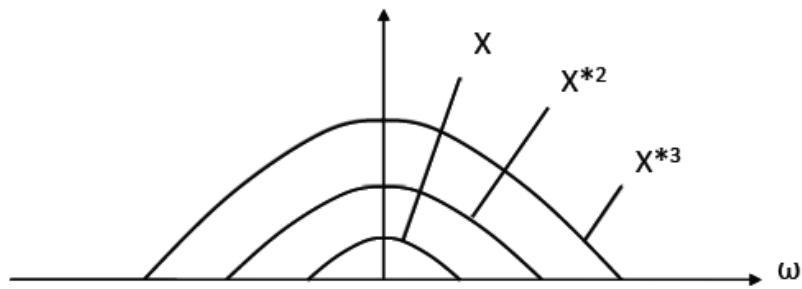


Memoryless Nonlinearity Model

$$\nu(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \nu^{(k)}(0) x^k$$

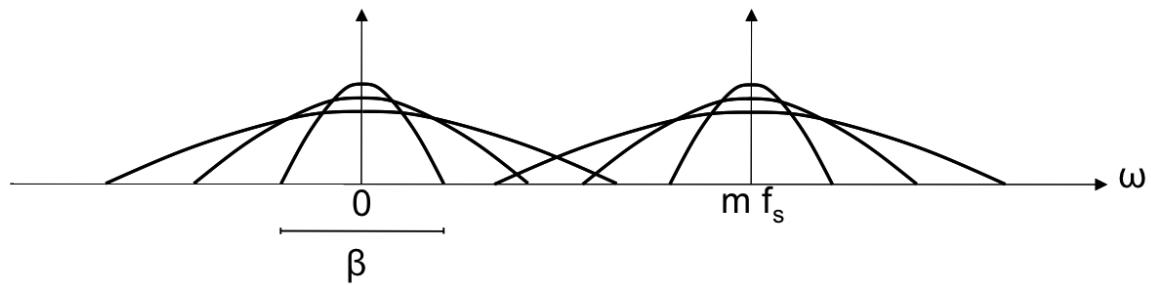
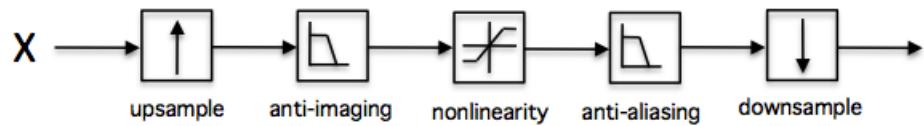
$$y(t) = \nu(x(t))$$

$$Y(\omega) = \sum_{k=0}^{\infty} \frac{1}{k!} \nu^{(k)}(0) X(\omega)^{*k}$$



- Expanding the nonlinearity using a Taylor series shows the increased output bandwidth.

Memoryless Nonlinearity Discretization



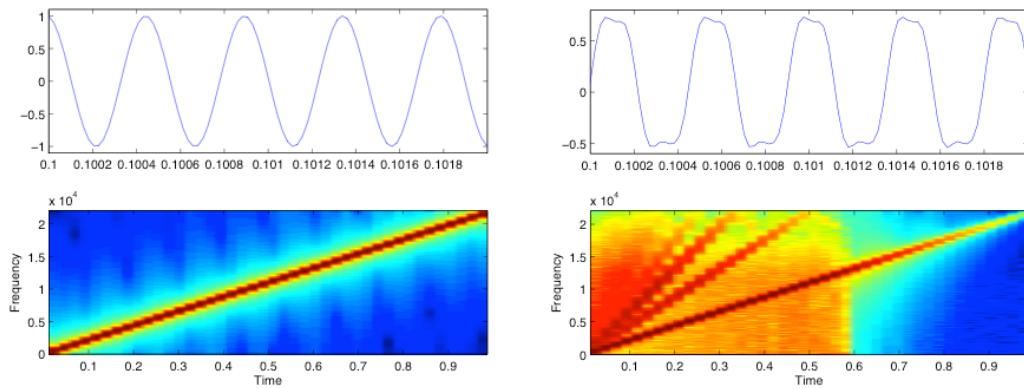
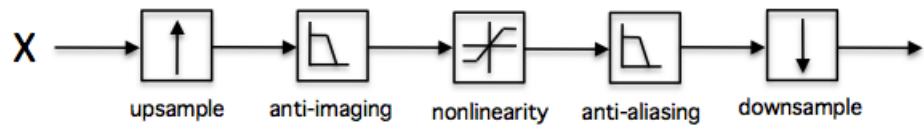
- With a signal bandwidth β , an upsampling factor of m , and a p^{th} order nonlinearity, no aliasing will occur, provided that

$$m \cdot f_s - p \cdot \beta/2 \geq f_s/2$$

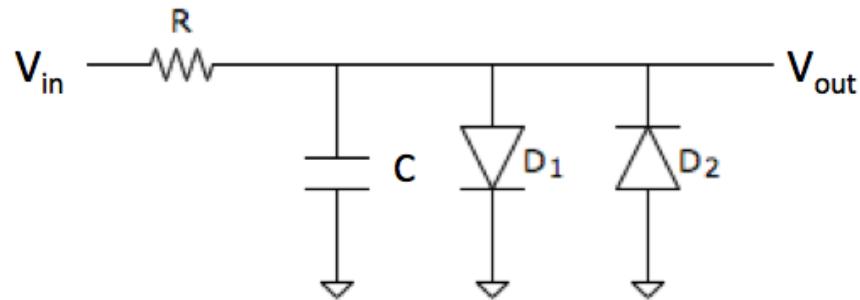
- If

$$\beta = f_s, \quad m \geq \frac{p+1}{2}$$

Anti-aliased Memoryless Nonlinearity



Circuit Model Distortion



$$\frac{V_{in} - V_{out}}{R} = C \frac{dV_{out}}{dt} + I_s \cdot \sinh\left(\frac{V_{out}}{V_T}\right)$$

- Solve the differential equation; sample the output.
- Numerically integrate the differential equation.