Sound and Voice Synthesis and Analysis

Human Computer Interface Technology

November 10, 1999

Perry R. Cook
Princeton University

Sound and Voice, New and Expected Frontiers

- **Sound and Voice on Computers is Strange (Historically)**
  - **We've Always Known Our Computers Will**
    - **Speak and**
    - **Understand Speech**

    Of course, they do neither very well
  - **Sound is expected. We notice if**
    - **sound isn't there, or**
    - **doesn't fit well, or**
    - **is too loud, or**
    - **isn't synchronized**
  - **but sound is quite subversive**
    - **Can influence emotion**
    - **Can affect other senses**
    - **Can distract from, or enhance, bad displays**

PCM Waveform (Wavetable) Synthesis

uses prerecorded (or synthesized) waveforms, stored in memory or on disk. Defined as such, it's difficult to call it synthesis, but advanced techniques to modify and combine recorded sounds bring PCM squarely into camp of flexible synthesis algorithms.

PCM is:

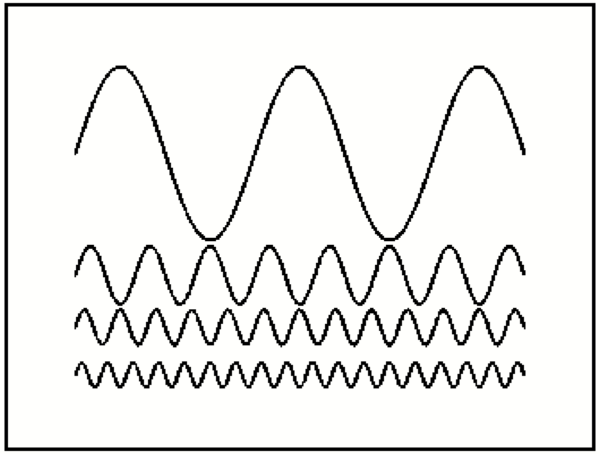
- **Easy, because you can just record things and play them back**
- **Hard, because**
  - **if you can't record them, it's difficult to synthesize them**
  - **in the limit, it requires infinite memory and time**

Requirements:

- **Storage**
- **Filters**
- **Interpolator**
- **DACs**

Additive (Fourier) Synthesis

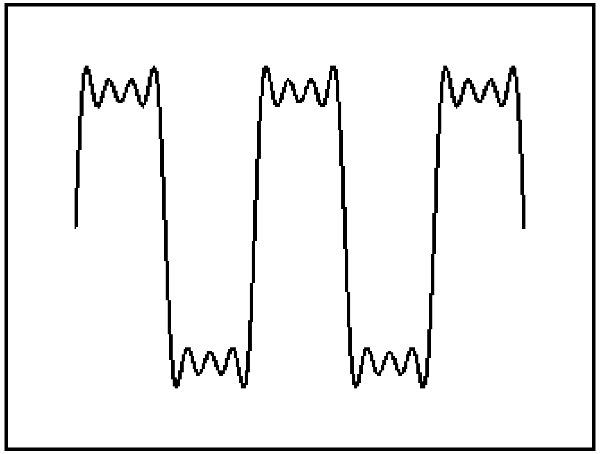
Fourier Analysis: Represent Time Waveforms as Linear Combination of Sine Waves



**Example:
Construct a square wave by adding sine waves of odd integer frequency relationships and amplitude of 1/frequency.**

sin(Ct) + sin(3Ct)/3 + sin(5Ct)/5 + sin(7Ct)/7

Additive (Fourier) Synthesis



sin(Ct) +
sin(3Ct)/3 +
sin(5Ct)/5 +
sin(7Ct)/7

Addition of odd sine waves approximates a square-wave.

More components yields better fit.

Additive (Fourier) Synthesis

Applying Fourier Analysis Yields a "Frequency Spectrum"

Handy Fourier Facts:

- **Any waveform can be represented by a combination of sinusoids.**
- **But! It might take an infinite number of sinusoids.**
- **For a digitally sampled signal of length N**
  It will take at most N/2 sinusoids to represent it.

Three Types of Spectra

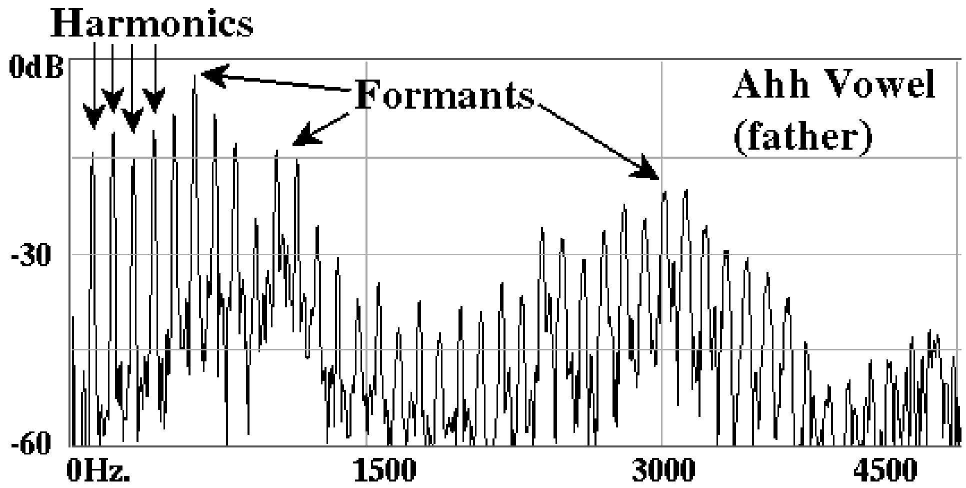
Harmonic
Pitched Periodic Sounds: Vowels, Trumpets, etc.
These types of sounds give us a strong sense of pitch.

InHarmonic
Bells, Gongs, Some Drums. Weak or ambiguous sense of pitch.

Noise
Consonants, Some Percussion Instruments,
Attacks of Many Harmonic and Inharmonic Sounds
Sense of high or low, but no clear sense of pitch.

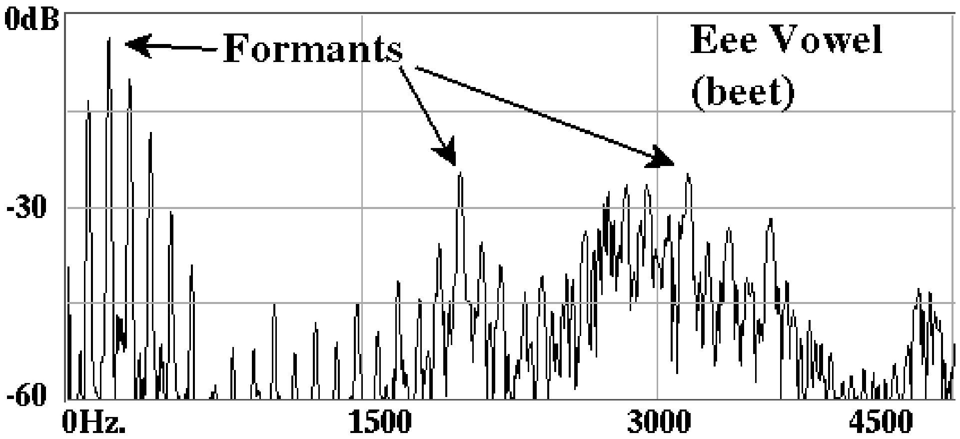
Real sounds are a mixture of sources and different spectral types.

Harmonic Spectra



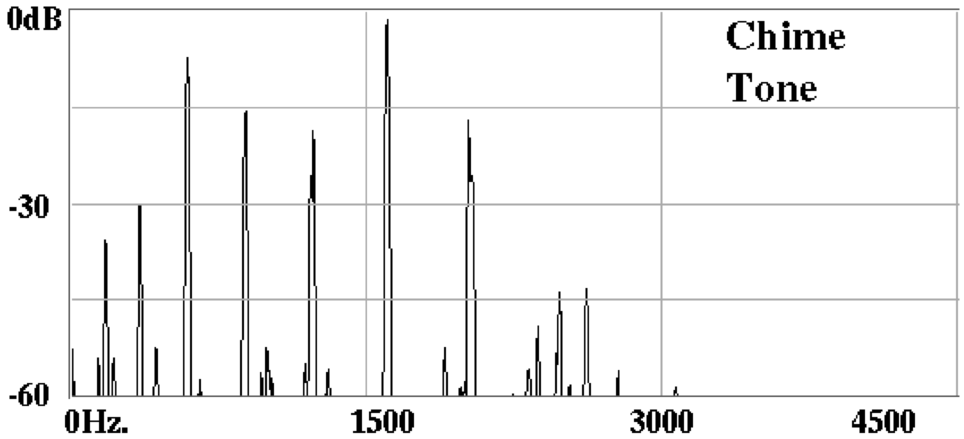
Fourier spectrum of voiced vowel ahh (as in father), showing "harmonic" sinusoidal peaks at integer multiples of a fundamental frequency.

Harmonic Spectra



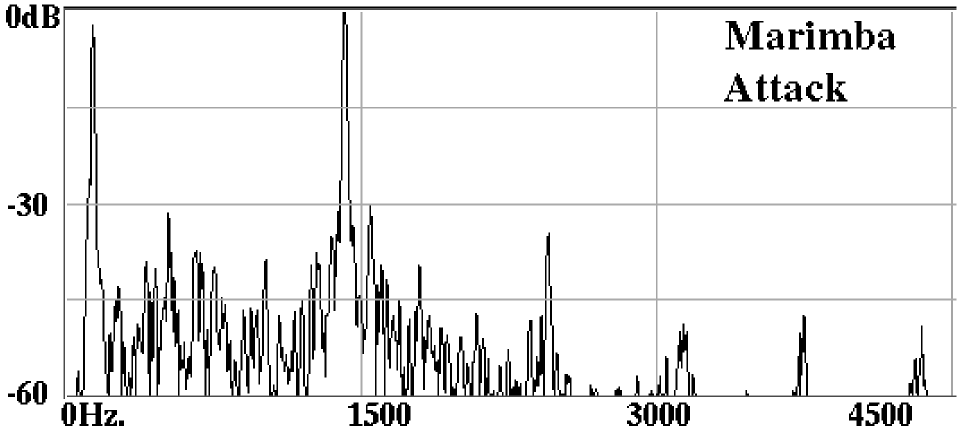
Fourier spectrum of voiced vowel eee (as in beet), showing harmonics.
Note that the overall spectral shape differs from the ahh vowel.
This distinguishes those two sounds perceptually.

Inharmonic Spectra



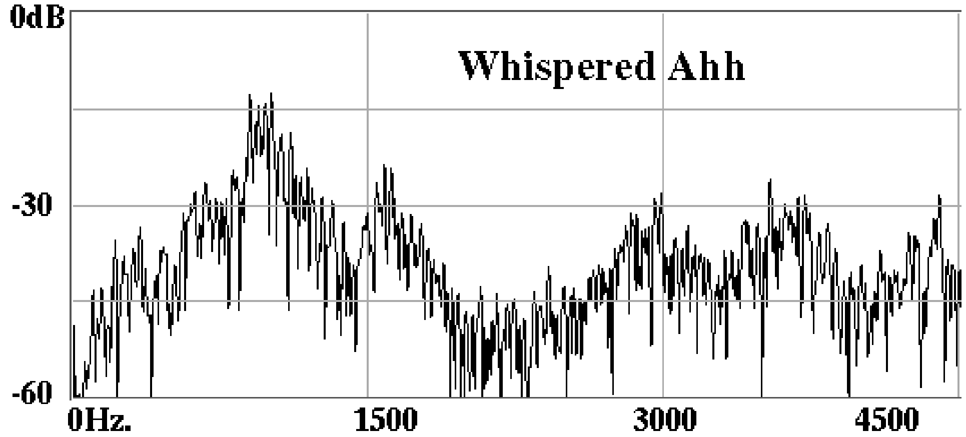
Fourier spectrum of chime note, showing "partials", which are sinusoids, but not harmonically related.
Such sounds have perceptually ambiguous pitch.

Inharmonic Spectra



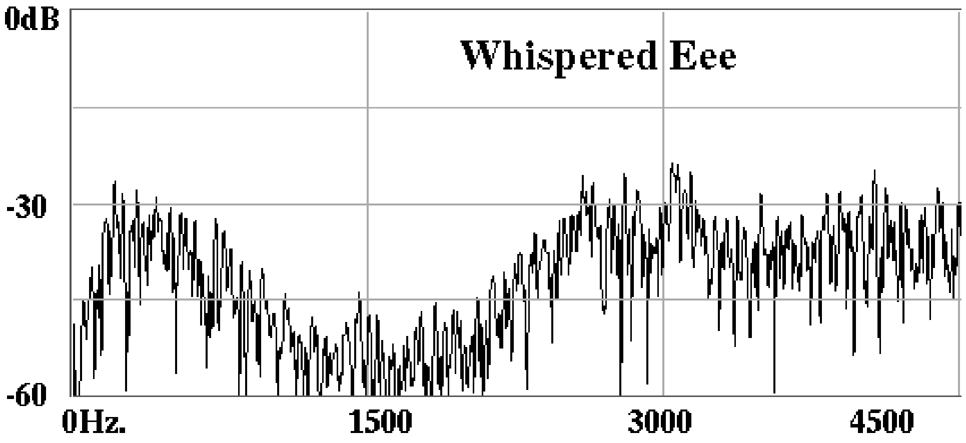
Fourier spectrum of marimba note, showing some inharmonic partials and noise from the stick strike.

Noise Spectra



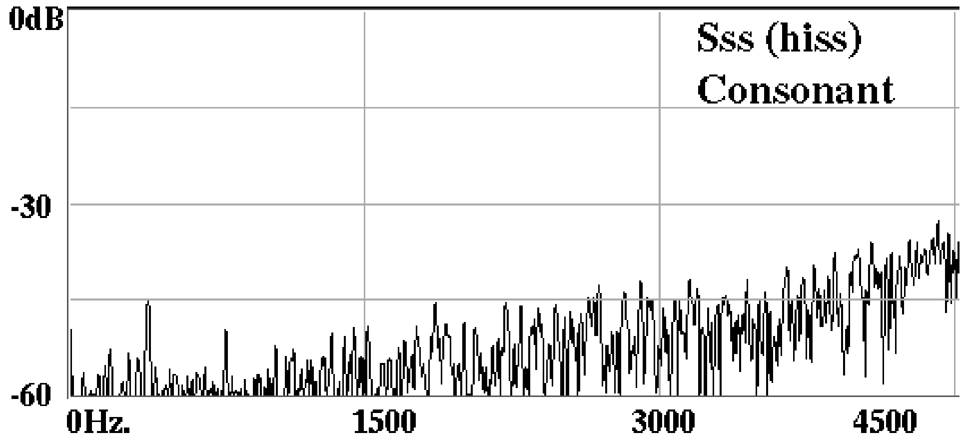
Fourier Spectrum of whispered vowel ahh (as in father). This spectrum shows no clear partials, but a noise spectrum in the same ahh shape as the harmonic vowel spectrum shown previously.

Noise Spectra



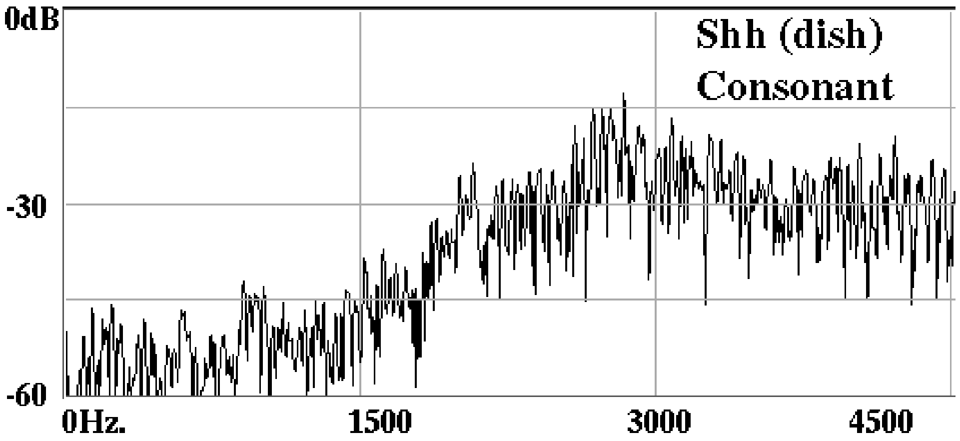
Fourier spectrum of whispered vowel eee (as in beet).

Noise Spectra



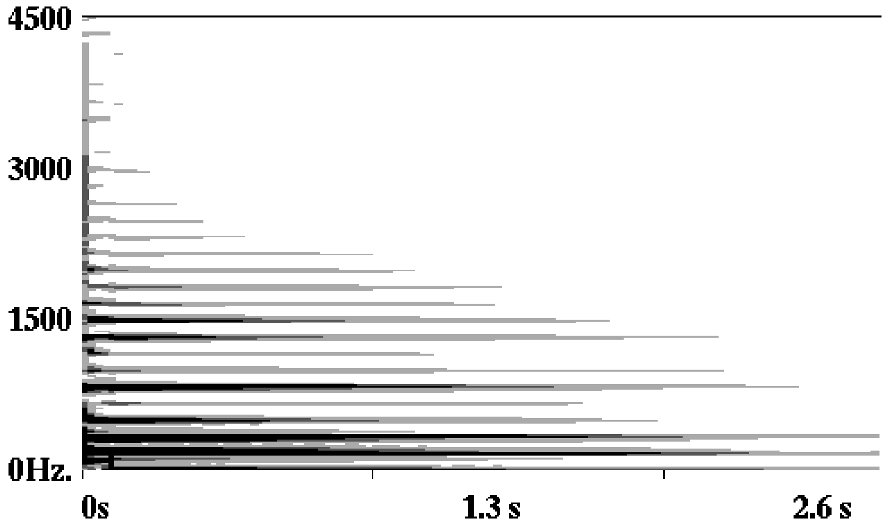
Fourier spectrum of consonant sss (as in hiss).

Noise Spectra



Fourier spectrum of consonant shh (as in dish).

The Sonogram

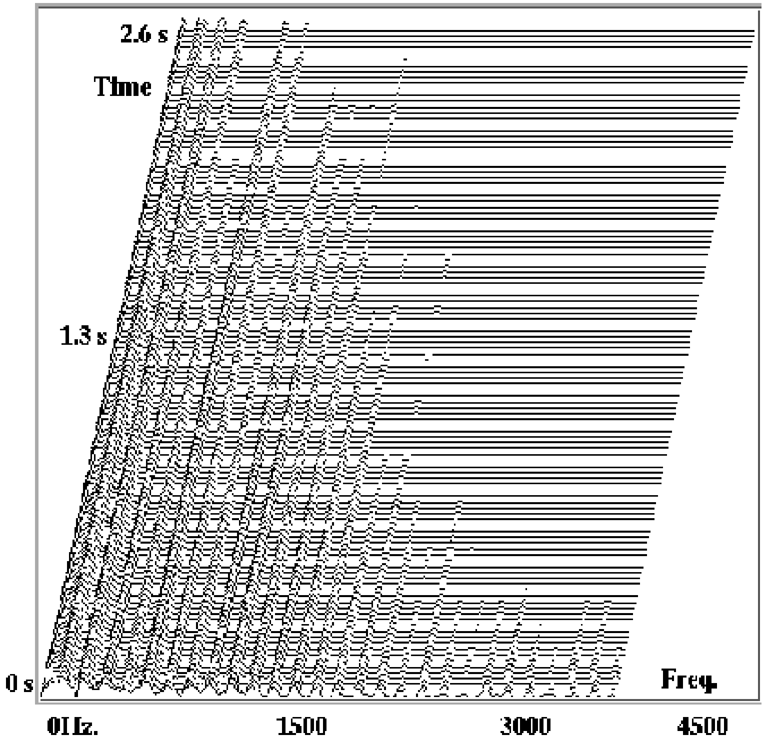


Sonogram plot of single guitar tone, showing frequency on the vertical axis and time on the horizontal axis.

The Waterfall Plot

Waterfall spectrum plot of single guitar note. Frequency is shown on the horizontal axis, and time is shown angling upward to the right.

What is shown here is that a plucked guitar string begins with a noisy attack, which very soon gives way to harmonic oscillation. Also, the higher harmonics decay more quickly than the lower frequencies.

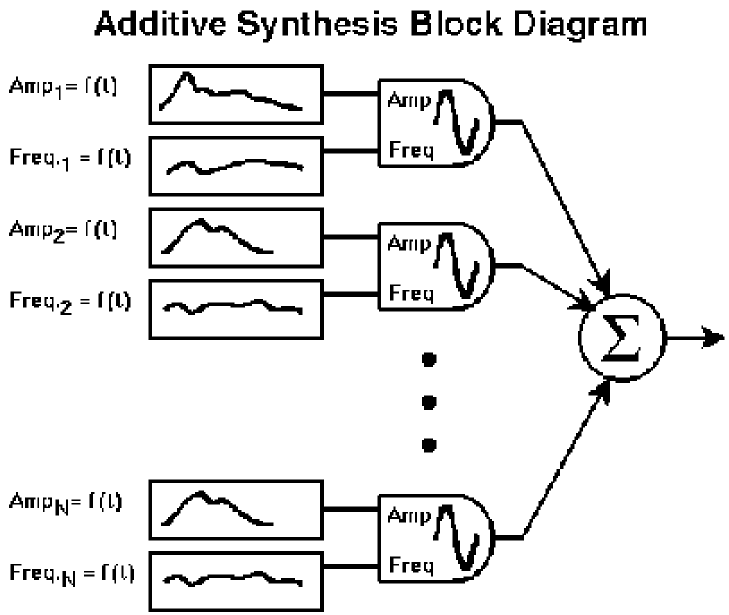


Additive Synthesis Diagram

Additive Synthesis Can Be Used to Model Spectra Containing a Few Harmonic and/or Inharmonic Partials.

Sine Waves as They Evolve in Time are Modeled by Individual Oscillators With Time-Varying Pitch and Amplitude.

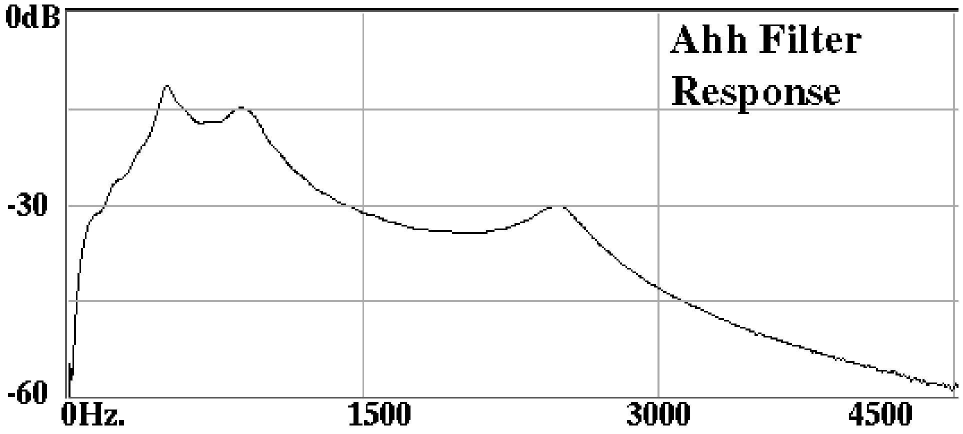
Sound Example:
Electronic Organ



Subtractive Synthesis

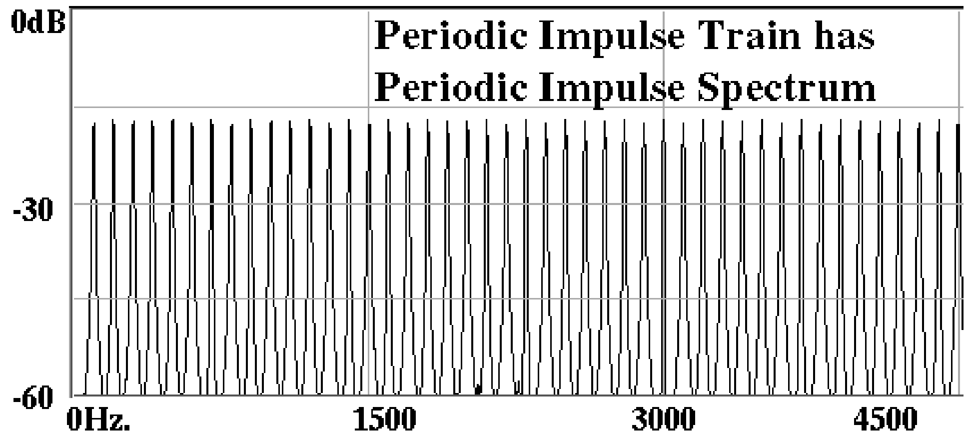
Generate Complex Source and Filter it to Desired Spectral Shape

Sources: Noise, Pulse Wave, Other Wideband Sounds.



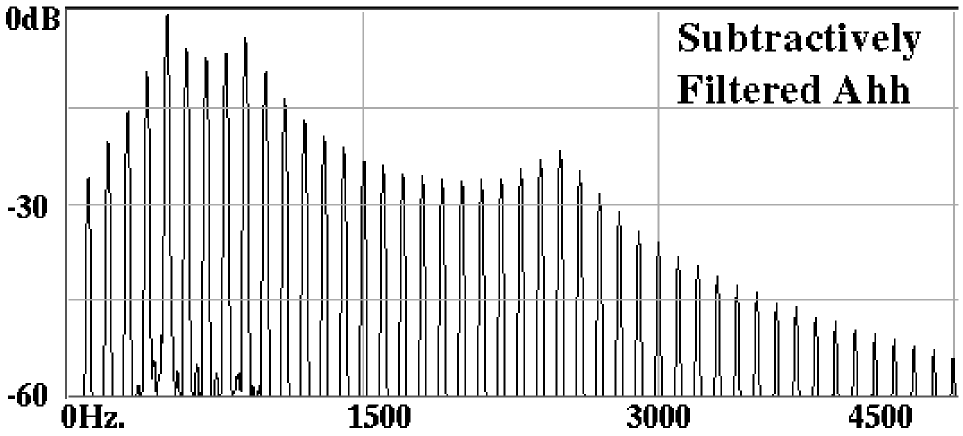
Filter in spectral shape of "Ahh"

Subtractive Synthesis



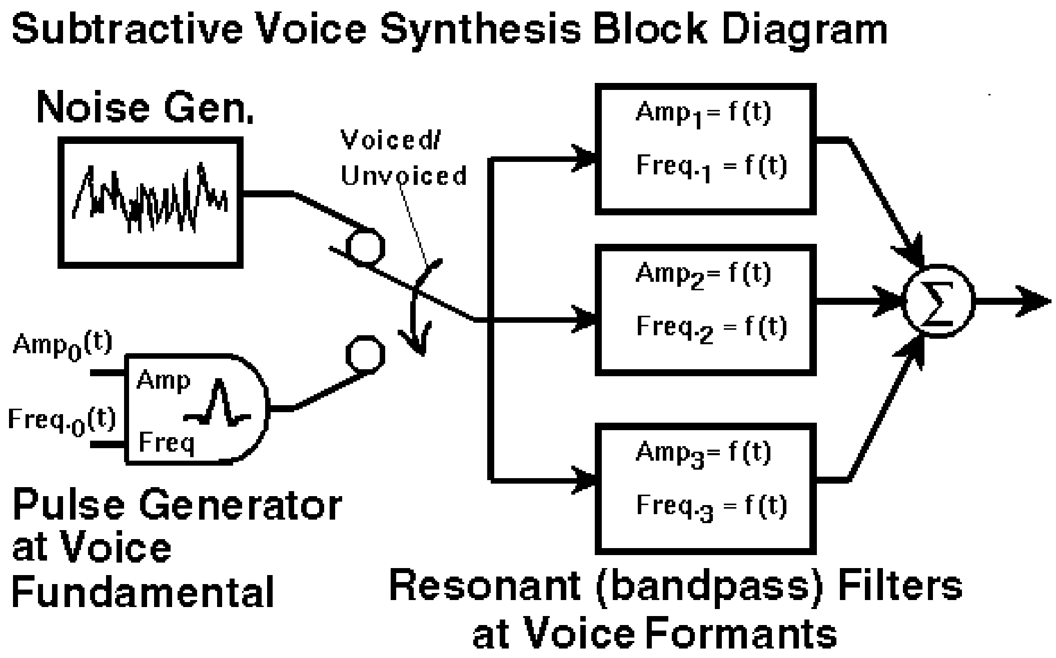
Spectrum of periodic impulse train
(all harmonics equal amplitude).

Subtractive Synthesis



Applying a filter with the spectral shape of an ahh vowel to a periodic impulse train results in a subtractively synthesized ahh sound.

Subtractive Voice Synthesis



Additive and Subtractive Analysis

- **The Fourier Transform Can be Used to Analyze Sounds for Additive Resynthesis**
- **Techniques such as Linear Predictive Coding (LPC) can decompose sound into Filter (Spectral Shape) and Source (What's Left Over)**
- **Sound Examples:**

  Voice
  Marimba
  Metal Sounds
  Wood Sounds

FM and Waveshaping

Non-Linear Waveform Synthesis

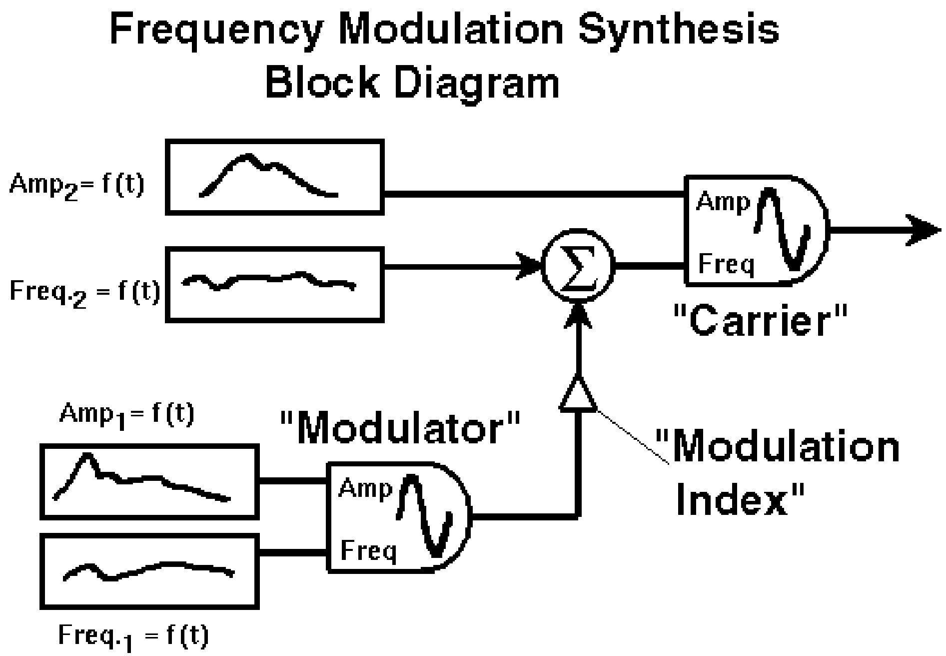
Generate Complex Spectra by Non-Linear Warping Function

$$y(t) = F(x(t))$$

In Simple FM, both functions are sinusoids

Two sinusoidal oscillators are used, with one functioning as the modulator, and the other as the carrier. The carrier generates a sinewave which is warped by the modulation of its frequency.

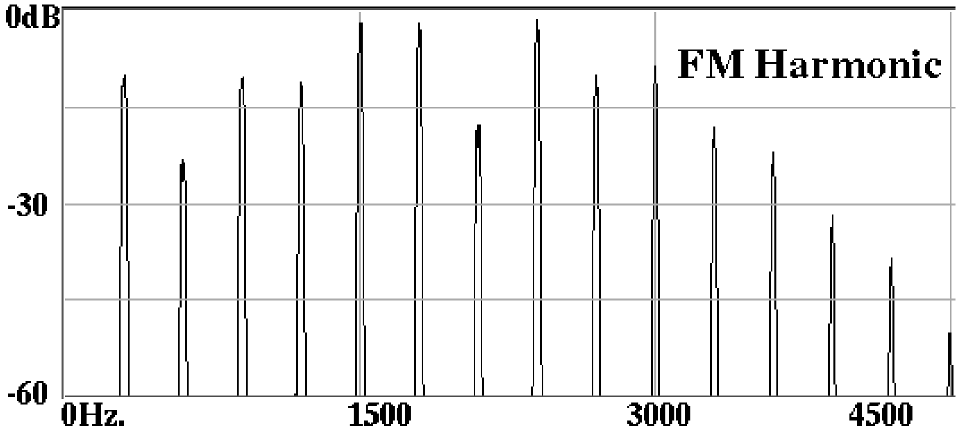
Simple FM



FM Basics

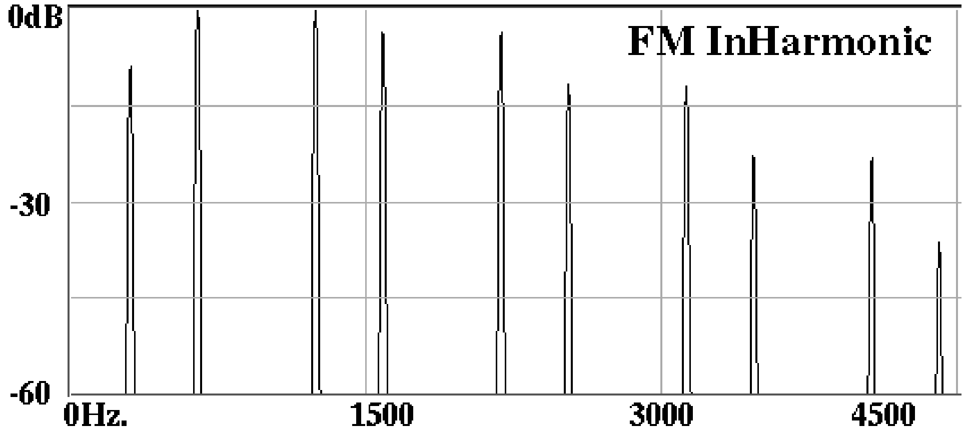
- When using simple sinusoidal FM, sideband sinusoidal partials appear around the carrier frequency, spaced at integer multiples of the modulation frequency.
- The number of significant sidebands grows (nearly) linearly with the amount of modulation.
- The ratio of the frequencies of the carrier and modulator oscillators determines the components of the spectrum.

Harmonic FM Spectrum



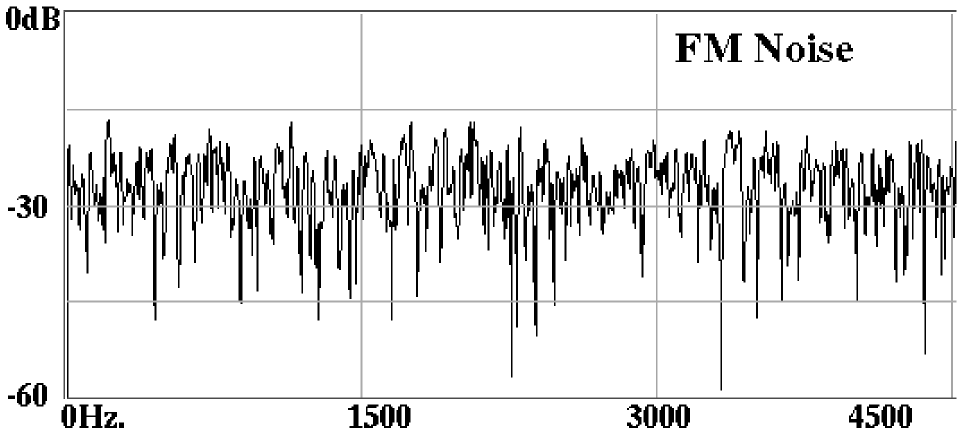
By setting the carrier frequency to an integer multiple of the modulator frequency,harmonically related sidebands result.

Inharmonic FM Spectrum



By setting the carrier frequency to an irrational multiple of the modulator frequency, inharmonically related sidebands result.

Noise FM Spectrum

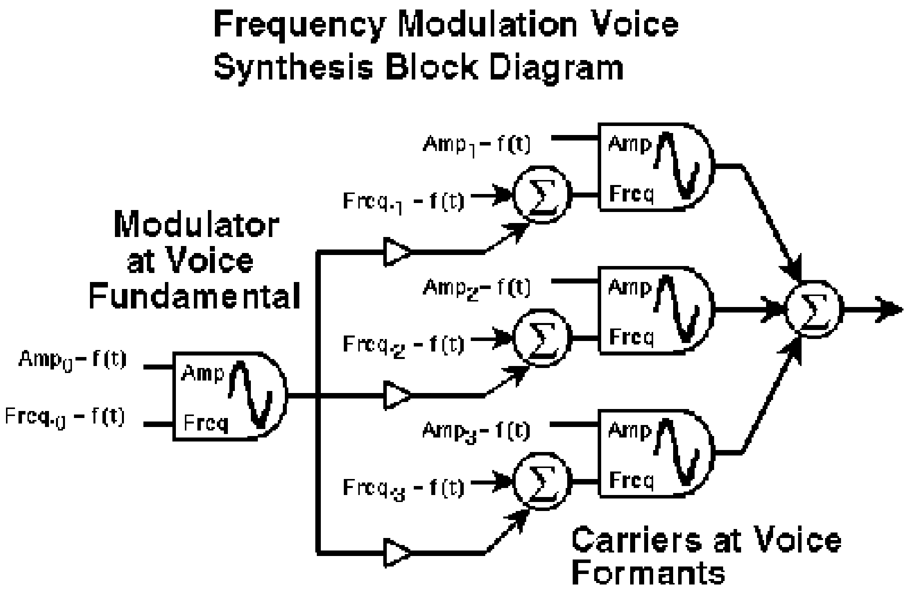


If the amount of modulation is set to a significantly high level, aliasing in the sidebands creates a dense, noise-like spectrum.

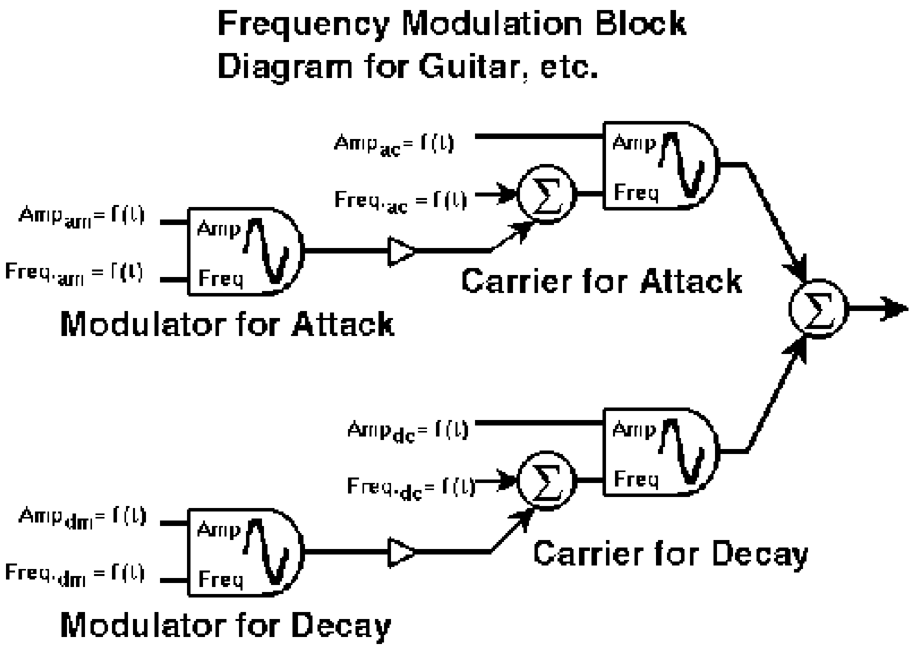
FM: More Carriers and Modulators Yield Better Spectral Fits

- **FM vocal model places a carrier near each formant resonance, and modulates all carriers with a common oscillator at the voice fundamental pitch.**
- **Instruments with a noisy attack and exponential decay of sinusoidal partials are modeled well with one simple FM pair for the attack, and another pair for the decay. The nature of FM lends itself to the higher harmonics or partials dying away more rapidly than the lower spectral components.**

FM Voice Synthesis



A Popular FM Algorithm



FM Sound Examples

Inharmonic Sound Example: Simple FM Bell

Harmonic + Noise Sound Example: Electric Piano

Physical Modeling

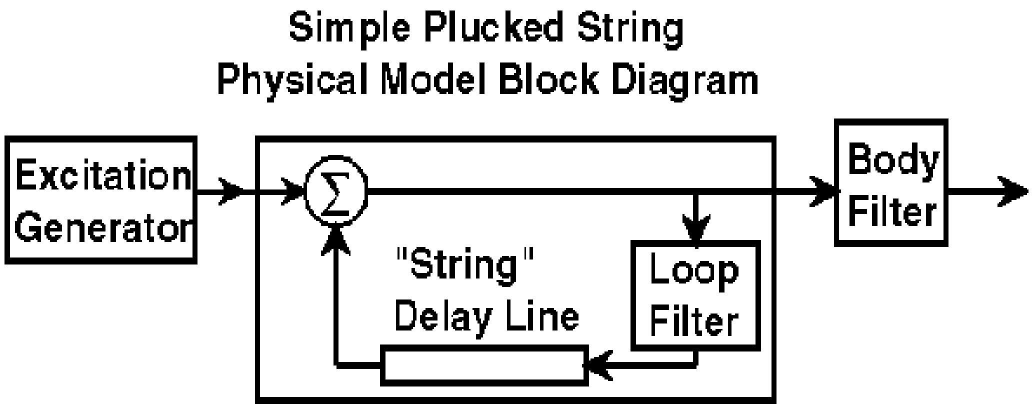
Model not the Waveform, Not the Spectrum, but the Time Domain Physics of the Instrument

Voice: Late 1950s
Strings: Late 60s
Winds: Late 70s

Can take advantage of one-dimensional paths in many systems. Strings,narrow pipes, and other such paths can often be replaced with delay lines (waveguides).

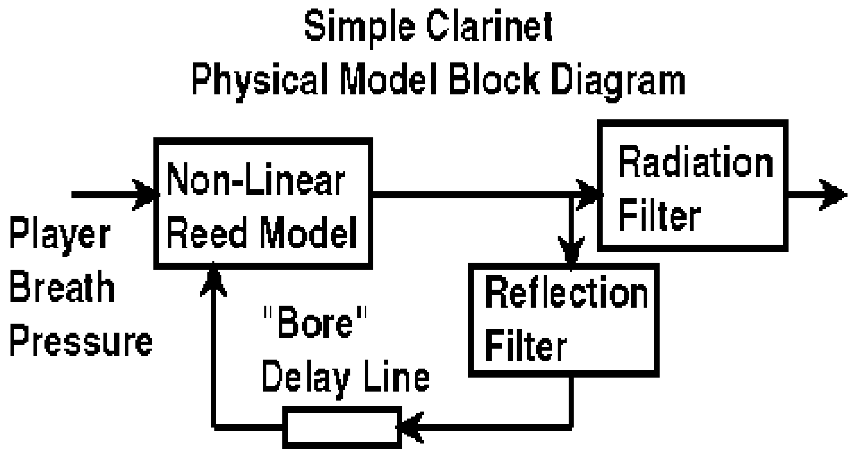
Any losses and some non-linearities along these paths can be lumped into calculations at at connection points.

Plucked String Model



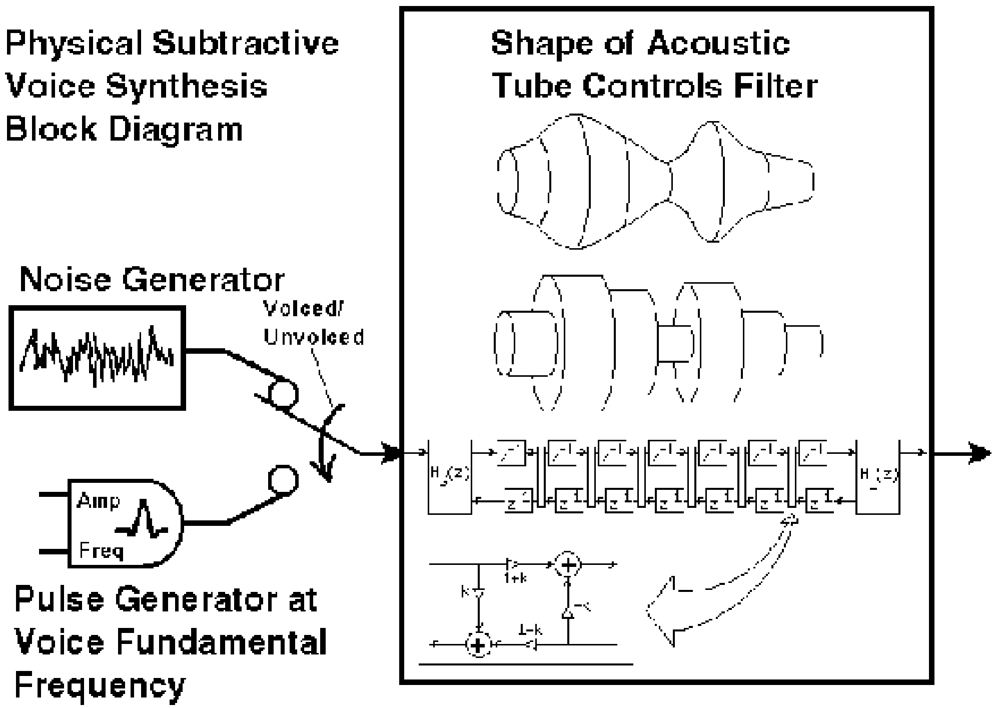
**Simple plucked string model. Delay models round-trip time around string,
filters model effects of instrument body. Excitation can be as simple as a burst of noise, or more elaborate for more realistic sound synthesis.**

Clarinet Model

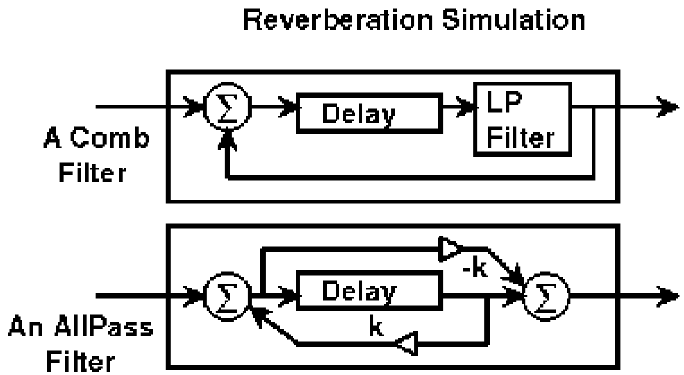


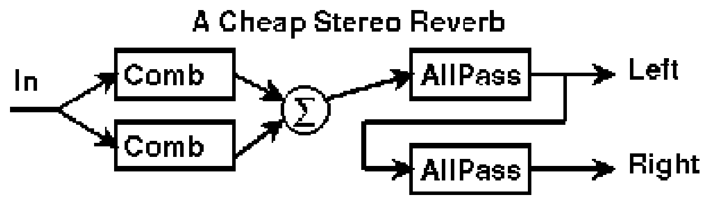
Simple clarinet wind instrument model. Delay-line models round-trip time around tube. Filters model effects of toneholes and bell. Non-linear "reed" function is the heart of most wind instrument models.

Voice Model



Delay-Based Effects





Comb filters model resonances between parallel surfaces, like walls in a room. AllPass filters model dispersion.

Physical Modeling Sound Examples

Waveguide String Sound Example: Mandolin

Waveguide Wind Sound Example: Clarinet

Particle Model Sound Example: Shakers

Particle Model Sound Example: Crunchy Sounds

Particle/Deterministic Sound Example: Ratchets

Back to Sound

That's All For Now

Roll Back To Auditory Display