Spectral Modeling

Perry R. Cook

Princeton Computer Science
(also Music)

Sound Views: Frequency Domain

Frequency Domain:

- Many physical systems have modes (damped oscillations)
- Wave equation (2nd order) or Bar equation (4th order) need 2 or 4 "boundary conditions" for solution
- Once boundary conditions are set solutions are sums of exponentially damped sines the sinusoids are Modes

The (discrete) Fourier Series

A time waveform is a sum of sinusoids

$$x(n) = \sum_{n=0}^{N-1} A_m \exp(\frac{j2\pi nm}{N}) \quad \textit{(A is complex)}$$

$$= \sum_{n=0}^{N-1} B_m \sin(\frac{2\pi nm}{N}) + C_m \cos(\frac{2\pi nm}{N})$$

$$= \sum_{n=0}^{N-1} D_m \cos(\frac{2\pi nm}{N} + \theta_m)$$

The (discrete) Fourier Transform

A Spectrum is a sinusoidal decomposition of a signal

$$A(m) = X(SRATE * m / N) = \sum_{n=0}^{N-1} x(n) \exp(\frac{-jnm2\pi}{N})$$

This transform is unique and invertible
(non-parametric representation like sampling)

Spectra: Magnitude and Phase

Often only magnitude is plotted

- Human perception is most sensitive to magnitude
  - *Environment corrupts and changes phase*
- 2 (pseudo-3) dimensional plots easy to view

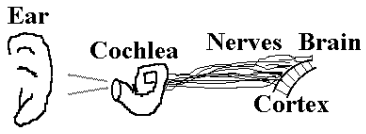
Phase is important, however

- Especially for transients (attacks, consonants, etc.)

If we know instantaneous amplitude and frequency, we can derive phase

Views of Sound: Perception

Human sound perception:



Brain: Higher level cognition, object formation, interpretation

Ear: receive 1-D waves

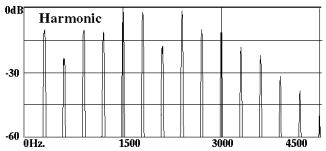
Cochlea: convert to frequency dependent nerve firings

Auditory cortex: further refine time & frequency information

Common Types of Spectra

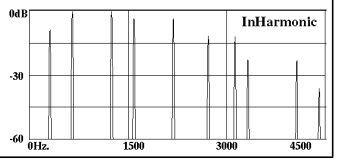
Harmonic

sines at integer multiple freqs.



Inharmonic

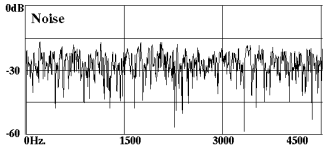
sines (modes), but not integer multiples



Common Types of Spectra

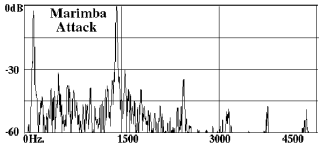
Noise

random amplitudes and phases



Mixtures

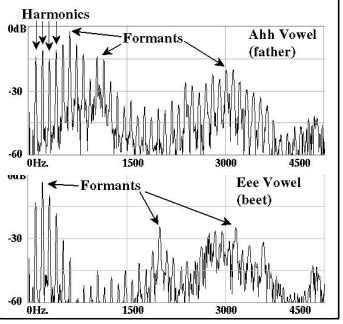
(most real-world sounds)



Perception: Spectral Shape

Formants (resonances) are peaks in spectrum.

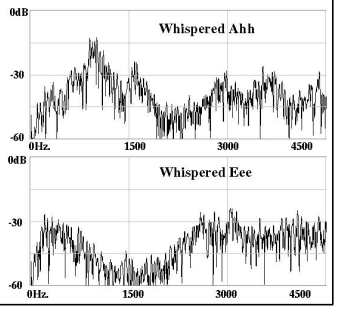
Human ear is sensitive to these peaks.



Spectral Shape and Timbre

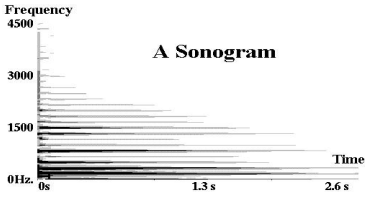
Quality of a sound is determined by many factors

Spectral shape is one important attribute



Spectra Vary in Time

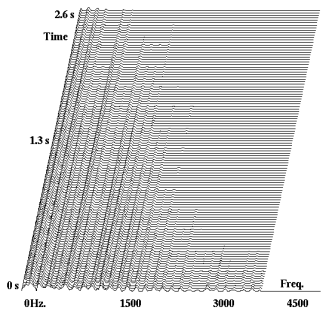
***Spectrogram** (sonogram)*
amplitude as darkness (color)
vs. frequency and time



Spectra in Time (cont.)

Waterfall Plot
pseudo 3-d amplitude as height vs. freq. and time

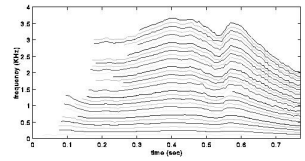
Each horizontal slice is an amplitude vs. time magnitude spectrum



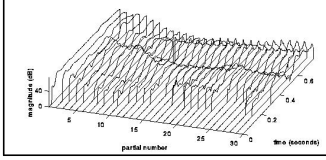
Sinusoidal Modeling

Vocoders	***Dudley '39, Many more since***
Sinusoidal Models	***Macaulay and Quatieri '86***
SANSY/SMS Sines + Stochastic	***Serra and Smith '87***
Lemur	***Fitts and Hakken '92***
FFT^{-1}	***Freed, Rodet and Depalle '96***
Transients	***Verma, Meng '98***

Sinusoidal Analysis "Tracks"



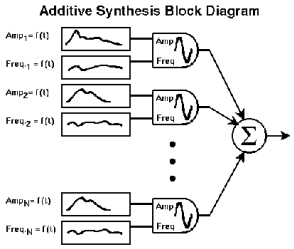
frequency of partials



magnitude of partials

Additive Synthesis

Control the amplitude and frequency of a set of oscillators



The sinusoidal model:

$$s(t) = \sum_{r=1}^{R} A_r(t)\cos\left[\theta_r(t)\right]$$

R : number of sinewave components,
$A_r(t)$: instantaneous amplitude,
$\boldsymbol{\theta}_r(t)$: instantaneous phase

References: Spectral Models

Dudley, H. 1939, "The Vocoder," Bell Laboratories Record, December.

Moorer, A. 1978. "The Use of the Phase Vocoder in Computer Music Applications." Journal of the Audio Engineering Society, 26 (1/2), pp. 42-45.

Dolson, M. 1986, "The Phase Vocoder: A Tutorial," Computer Music Journal, 10 (4), pp. 14 - 27.

Robert J. McAulay and Thomas Quatieri 1986, "Speech Analysis/Synthesis Based on a Sinusoidal Representation," IEEE Trans. Acous, Speech, Signal Processing, vol ASSP-34, pp. 744-754.

Xavier Serra, 1989, "A System for Sound Analysis/Transformation/Synthesis Based on a Deterministic Plus Stochastic Decomposition," Ph.D. dissertation, Dept. of Music, Stanford University, Stanford CA.

Kelly Fitz, Lippold Haken, and Bryan Holloway,1995, "Lemur - A Tool for Timbre Manipulation ," to appear in Proc. Intl. Computer Music Conf..

Adrian Freed, Xavier Rodet, and Phillipe Depalle 1993, "Synthesis and Control of Hundreds of Sinusoidal Partials on a Desktop Computer without Custom Hardware," Proc. ICSPAT.

T. Verma, T. Meng, 1998 "An Analysis/Synthesis Tool for Transient Signals that Allows a Flexible Sines+Transients+Noise Model for Audio," 1998 IEEE ICASSP-98. Seattle, WA, USA.

K. Van den Doel and D. Pai, "Synthesis of Shape Dependent Sounds with Physical Modeling," *Proc. Intl. Conference on Auditory Display*, Santa Clara, CA, 1997.