Physically-Based Parametric Sound Synthesis and Control

Perry R. Cook

Princeton Computer Science (also Music)

Course Introduction

Parametric Synthesis and Control of Real-World Sounds for

virtual reality

games

production

auditory display

interactive art

interaction design

Schedule

0:00 Welcome, Overview
0:05 Views of Sound
0:15 Spectra, Spectral Models
0:30 Subtractive and Modal Models
1:00 Physical Models: Waveguides and variants
1:20 Particle Models
1:40 Friction and Turbulence
1:45 Control Demos, Animation Examples
1:55 Wrap Up

Views of Sound

- *Sound is a recorded waveform*
  PCM playback is all we need for interactions, movies, games, etc.
  (Not true!!)
- *Time Domain* — *x(t) (from physics)*
- *Frequency Domain* — *X(f) (from math)*
- *Production* — *what caused it*
- *Perception* — *our image of it*

Views of Sound

Time Domain

is most closely related to

Production

Frequency Domain

is most closely related to

Perception

we will see that many hybrids abound

Views of Sound: Time Domain

Sound is produced/modeled by physics, described by quantities of

- Force — force = mass * acceleration
- Position — x(t) — actually < x(t), y(t), z(t) >
- Velocity — Rate of change of position dx/dt
- Acceleration — Rate of change of velocity dv/dt

Examples: ***Mass+Spring+Damper***
Wave Equation

Mass/Spring/Damper

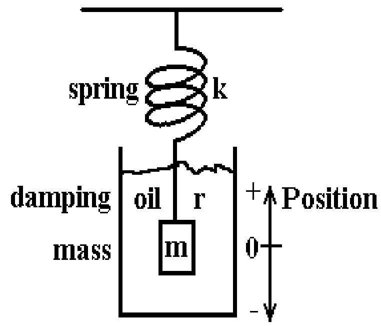
F = ma = - ky - rv - mg

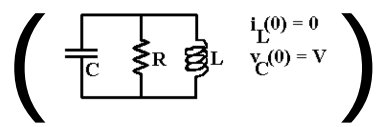
F = ma = - ky - rv

(if gravity negligible)

$$\frac{d^2y}{dt^2} + \frac{r}{m}\frac{dy}{dt} + \frac{k}{m}y = 0$$

$$D^2 + Dr/m + k/m = 0$$





2nd Order Linear Diff Eq. Solution

1) Underdamped:

$y(t) = Y_0 e^{-t/\tau} \cos(\omega t)$

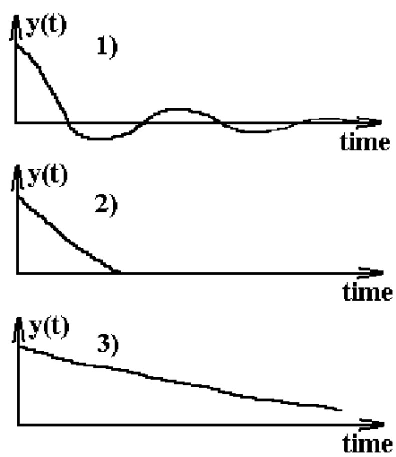
exp. * oscillation

2) Critically damped:

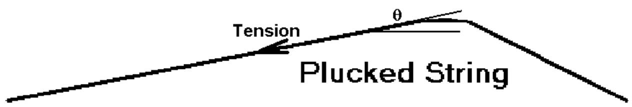
fast exponential decay

3) Overdamped:

slow exponential decay



Wave Equation



- **$df_y = (T \sin\theta)_{x+dx} - (T\sin\theta)_x$**
- ***$f(x+dx) = f(x) + \delta f/\delta x \, dx + \ldots$*** ***(Taylor's series)***
- ***$\sin\theta = \theta$*** ***(for small θ)***
- ***$F = ma = \rho \, dx \, d^2y/dt^2$*** ***(for each dx of string)***

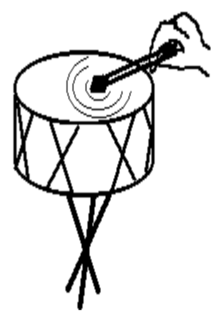
The wave equation: **($c^2 = T / \rho$)**

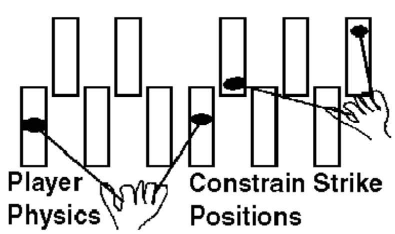
$$\frac{d^2y}{dx^2} = \frac{1}{c^2}\frac{d^2y}{dt^2}$$

Views of Sound: Production

Throughout most of history, some physical mechanism was responsible for sound production.

From our experience, certain <u>gestures</u> produce certain audible results





Examples:
Hit harder --> louder AND brighter
Can't move instantaneously
Can't do exactly the same thing twice

Sound Views: Frequency Domain

Frequency Domain:

- Many physical systems have <u>modes</u> (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

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

(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

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

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

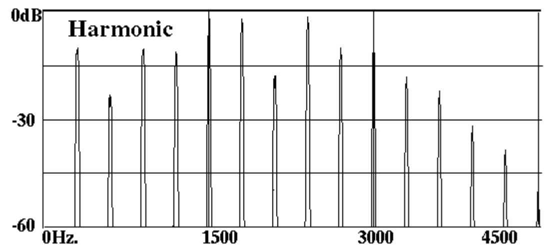
- Especially for transients (attacks, consonants, etc.)

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

Common Types of Spectra

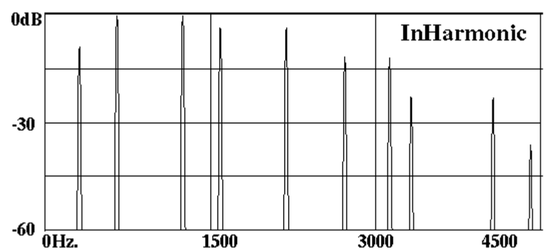
Harmonic

sines at integer multiple freqs.



Inharmonic

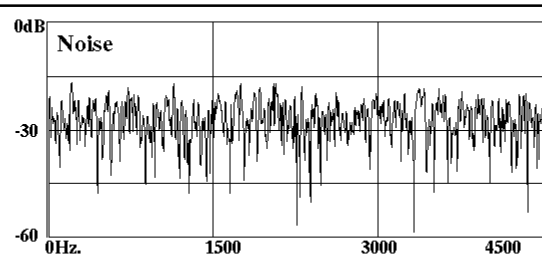
sines (modes), but not integer multiples



Common Types of Spectra

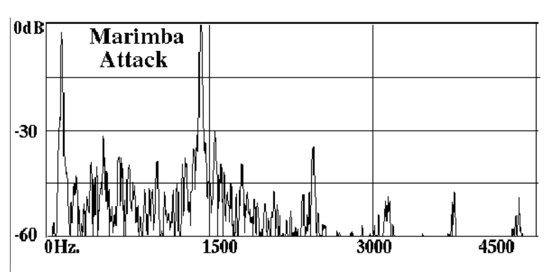
Noise

random amplitudes and phases



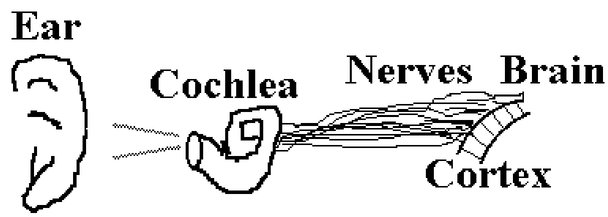
Mixtures

(most real-world sounds)



Views of Sound: Perception

Human sound perception:



Brain:
Higher level cognition, object formation, interpretation

Ear:
receive 1-D waves

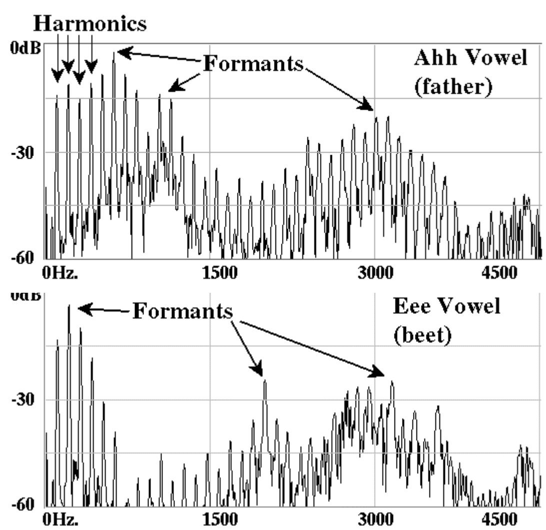
Cochlea:
convert to frequency dependent nerve firings

Auditory cortex:
further refine time & frequency information

Perception: Spectral Shape

Formants (resonances) are peaks in spectrum.

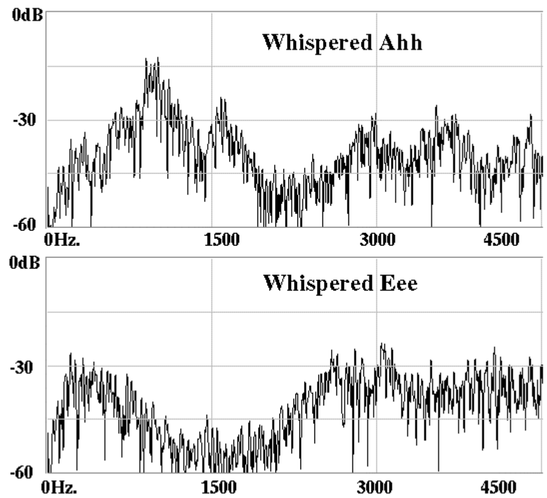
Human ear is sensitive to these peaks.



Spectral Shape and Timbre

Quality of a sound is determined by many factors

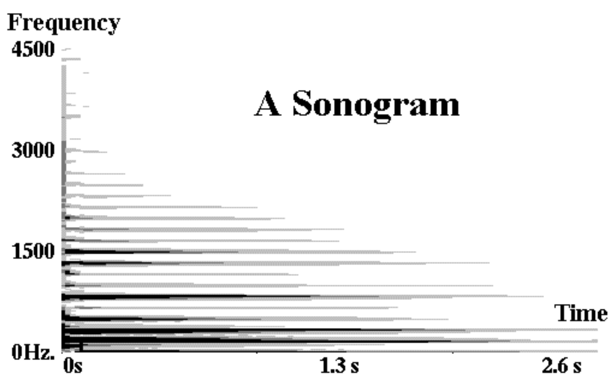
Spectral shape is one important attribute



Spectra Vary in Time

***Spectrogram** (sonogram)*

amplitude as darkness (color) vs. frequency and time

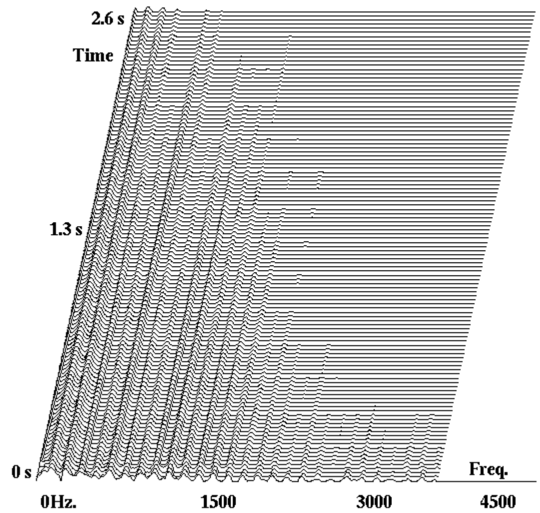


Spectra in Time (cont.)

Waterfall Plot

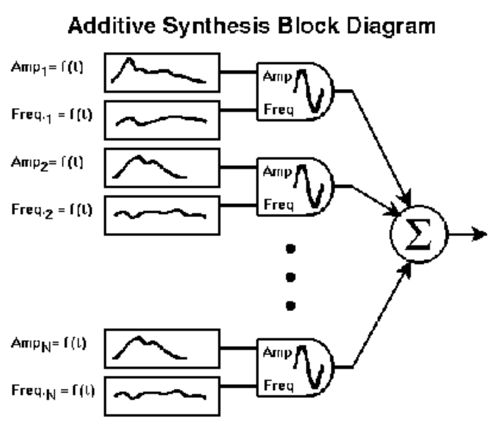
pseudo 3-d
amplitude as height vs. freq. and time

Each horizontal slice is an amplitude vs. time magnitude spectrum



Additive Synthesis

Control the amplitude and frequency of a set of oscillators



The sinusoidal model:

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

R : number of sinewave components,
$A_r(t)$: instantaneous amplitude,
$\theta_r(t)$: instantaneous phase

Sinusoidal Modeling

Vocoders — ***Dudley '39, Many more since***

Sinusoidal Models — ***Macaulay and Quatieri '86***

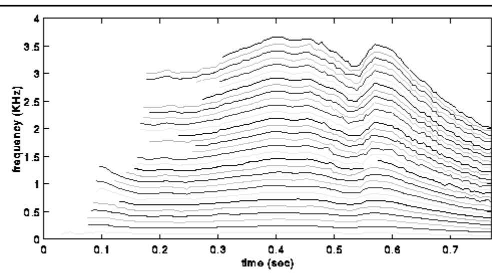
SANSY/SMS Sines + Stochastic — ***Serra and Smith '87***

Lemur — ***Fitts and Hakken '92***

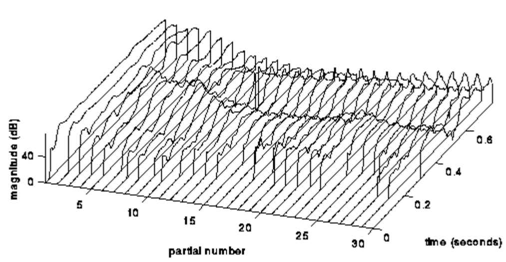
FFT^{-1} — ***Freed, Rodet and Depalle '96***

Transients — ***Verma, Meng '98***

Sinusoidal Analysis "Tracks"

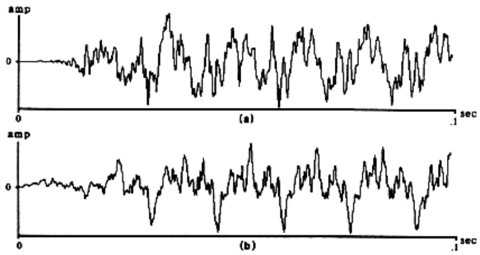


frequency of partials



magnitude of partials

Magnitude-only synthesis



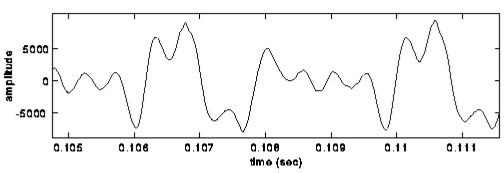
original sound

magnitude-only synthesis

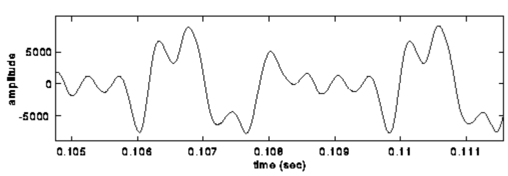
$$\hat{A}(m) = \hat{A}^{l-1} + \frac{(\hat{A}^{l} - \hat{A}^{l-1})}{S} m \qquad \hat{\omega}(m) = \hat{\omega}^{l-1} + \frac{(\hat{\omega}^{l} - \hat{\omega}^{l-1})}{S} m \qquad \hat{\theta}_r(m) = \hat{\theta}_r(l-1) + \hat{\omega}_r(m)$$

$$s^{l}(m) = \sum_{r=1}^{R^{l}} \hat{A}_r^{l}(m) \cos\left[\hat{\theta}_r^{l}(t)\right]$$

Magnitude and Phase Synthesis



original sound
x(n)



synthesized sound with phase matching
s(n)

$$\theta_r(t) = \int_0^t \omega_r(\tau) d\tau + \theta_r(0) + \varphi_r$$

$\omega_r(t)$: instantaneous frequency
$\theta_r(0)$: initial phase value
φ_r : fixed phase offset

Deterministic plus Stochastic Synthesis (SMS)

model: $$s(t) = \sum_{r=1}^{R} A_r(t)\cos[\theta_r(t)] + e(t)$$

$A_r(t)$, *$\theta_r(t)$*: instantaneous amplitude and phase of r[th] sinusoid,
e(t) : residual component.

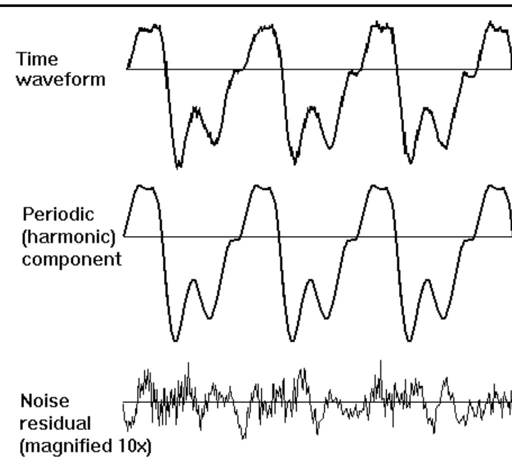
when sinusoids are very stable, the instantaneous phase can be calculated by: $$\theta_r(t) = \int_0^t \omega_r(\tau)d\tau$$

otherwise:
$$\theta_r(t) = \int_0^t \omega_r(\tau)d\tau + \theta_r(0) + \varphi_r$$

$\omega_r(t)$: instantaneous radian frequency
$\theta_r(0)$: initial phase value,
φ_r : fixed phase offset

Residual (stochastic component)

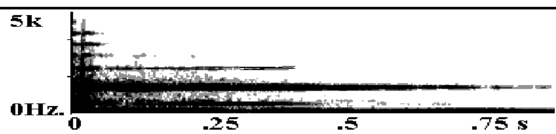
Resynthesis (with phase) of sine components allows extraction and modeling of residual component

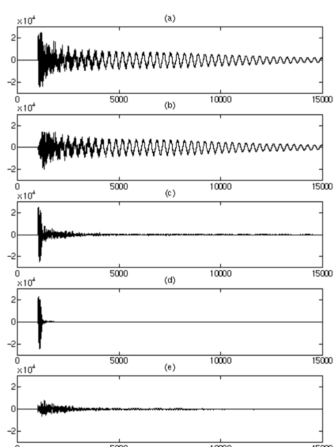


SMS High level attributes

- **Basic SMS parameters**
  - Instantaneous frequency and amplitude of partials
  - Instantaneous spectrum of residual
- **Instantaneous attributes**
  - Fundamental frequency
  - Amplitude and spectral shape of sinusoidal components
  - Amplitude and spectral shape of residual
  - Degree of harmonicity
  - Noisiness
  - Spectral tilt
  - Spectral centroid
- **Region Attributes**

Transients





- ***Transients are vertical stripes in spectrogram***
- ***Use DCT to transform back to time domain, then do “sinusoidal” track analysis on that***
- ***Detection is the hard part***

Sines + Noise + Transients Strengths and Weaknesses

Strengths:

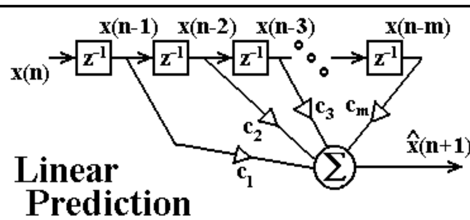
- General signal model
  (doesn't care what made it)
- Closed form identity analysis/resynthesis
- Perceptual motivations (somewhat, not all)

Weaknesses:

- No gestural parameterization
- No physics without lots of extra work
- No guaranteed compression
  or understanding

Subtractive Synthesis: LPC

$$x(n) = \sum_{k=1}^{m} c_k x(n-k)$$ LPC



$$e(n) = x(n) - \hat{x}(n)$$

Error

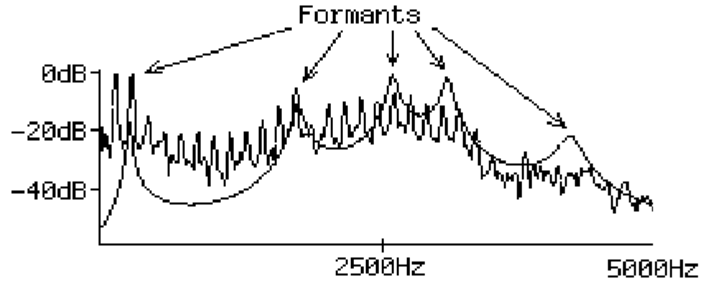
$$\hat{x}(n) = \sum_{k=1}^{m} c_k x(n-k)$$

Prediction

Mean Squared Error over block length P $$E = \frac{1}{P}\sum_{n=0}^{P} e(n)^2$$

LPC continued

LPC is well suited to speech



Also well suited to sounds with resonances

LPC filter envelope (smooth line) fit to human vowel sound / i / (eee)

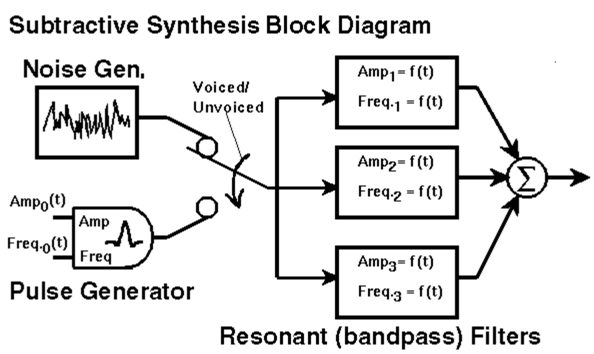
Subtractive Synthesis: Formants

Factor LPC into "resonators"

Eigenmode motivations

Perceptual motivations

Vocal production motivations



Excite with pulse(s), noise, or residual

Modal Synthesis

Systems with resonances
(eigenmodes of vibration)

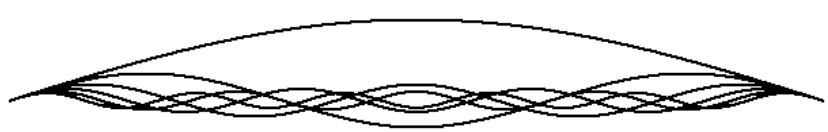
Bars, plates, tubes, rooms, etc.

Practical and efficient, if few modes

Essentially a subtractive model in that
there is some excitation
and some filters to shape it.

Modal Synthesis: Strings

Strings are pinned at both ends

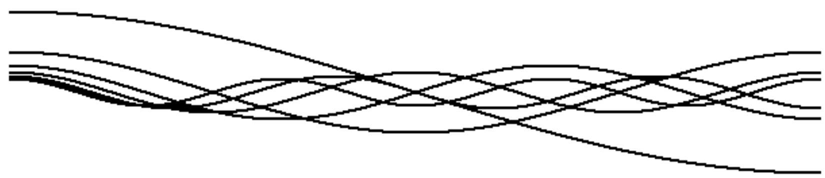


Generally harmonic relationship

Stiffness can cause minor stretching of harmonic frequencies

Modal Synthesis: Bars

Modes of Bars: Free at each end

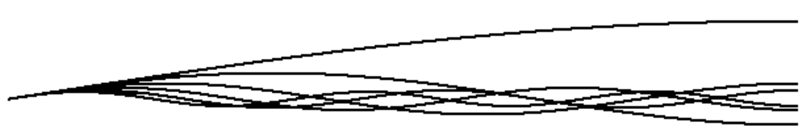


These would be harmonic, but stiffness of rigid bars stretches frequencies.

Modes: 1.0, 2.765, 5.404, 8.933

Modal Synthesis: Tubes

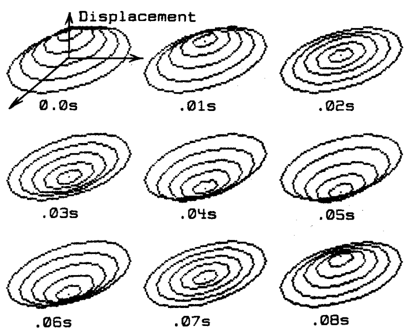
Open or closed at each end, same as strings and bars, but harmonic because speed of sound is constant with frequency



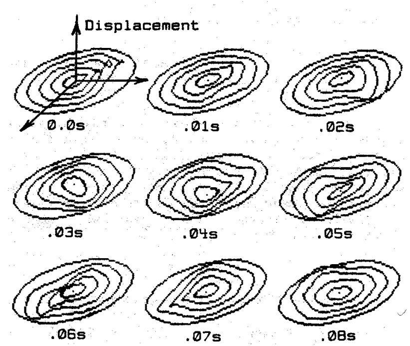
Open + Closed:odd multiples of fundamental   (quarter wavelength)

Modal Synthesis: Plates, Drums

Modes of Plates: inharmonic (round = Bessel)



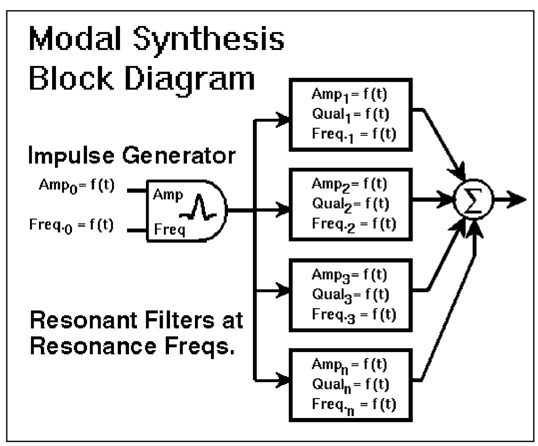
Center strike



Edge strike

Modal Synthesis: Block Diagram

- Impulse generator excites filters
- Residue can be used for excitation
- Filters shape spectrum, model eigenmodes
- Filter parameters can be time-varying



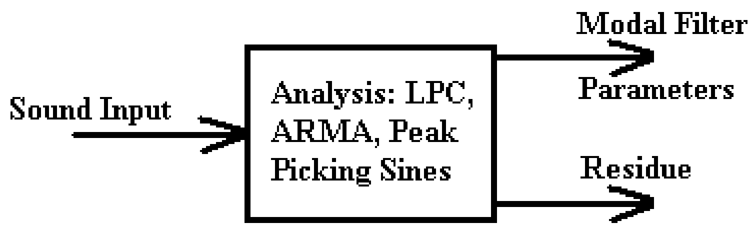
Modal: Residual Excitation

Linear source/filter decomposition

•"Parametric sampling"

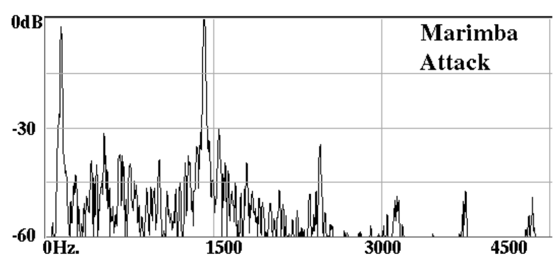
–Drive filter with source and get back identity

–Can modify the parameters in interesting ways

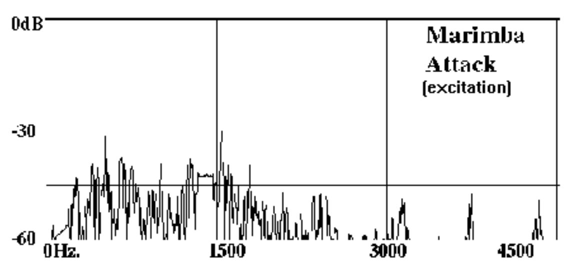


Residual Extraction, Example

Original struck bar



After modal subtraction



Modal Synthesis

Strengths:

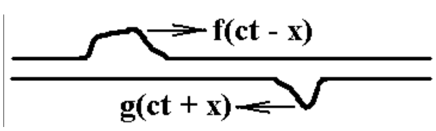
- Generic, flexible, cheap if only a few modes
- Great for modeling struck objects of metal, glass, wood

Weaknesses:

- No spatial sampling
- No (meaningful) phase delay
- Hard to interact directly and continuously
  (rubbing, damping, etc).

Physical Modeling: Strings

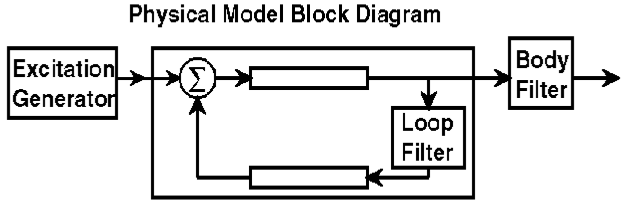
Plucked string model



Use D’Alembert Solution
of 2nd order wave equation
(left and right going components)

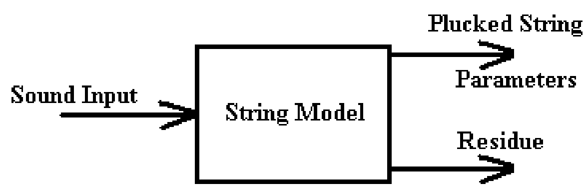
- Digital Waveguide Filter Solution

Simple Plucked String
Physical Model Block Diagram



Physical Models: Strings

Can extract model parameters and residual excitation for plucked strings

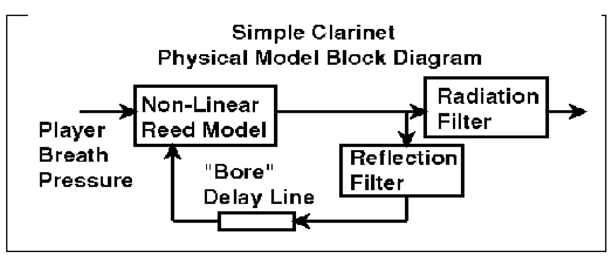


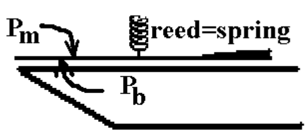
Bowed Strings: Model stick-slip friction curve of bow-string interaction

Physical Models: Tubes

Clarinet, Flute, Brass

Same D'Alembert solution as strings, but add non-linear oscillator





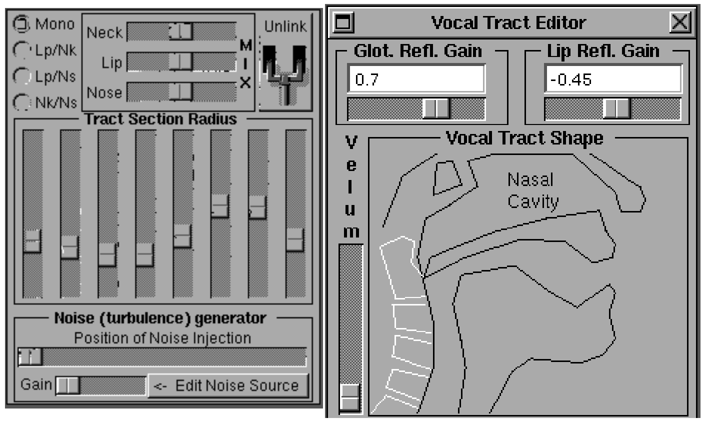
Non-linear Reed Model

Reed, Lip, Flute Jet acts as spring and switch for air

Physical Models: Voice

Acoustic tube model of vocal tract

Oscillator or mass/ spring model of source(s)



Physical Models: Waveguides for Strings, Tubes, Voices

Strengths:

- Cheap in both computation and memory
- Parametrically meaningful, extensible for more realism

Weaknesses:

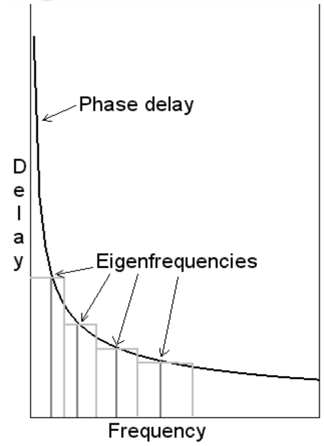
- Little in the real world looks or sounds exactly like a plucked string, flute, clarinet, or human voice
- Each family needs a different model

Physical Models: Bars

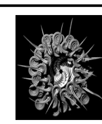
Stiffness in bars makes wave propagation frequency dependent

Idea:

- Model each mode with filter and delay
- Merge modal with waveguide
- Preserve spatial sampling
- Can strike, damp, bow



Banded Waveguides (Georg Essl)



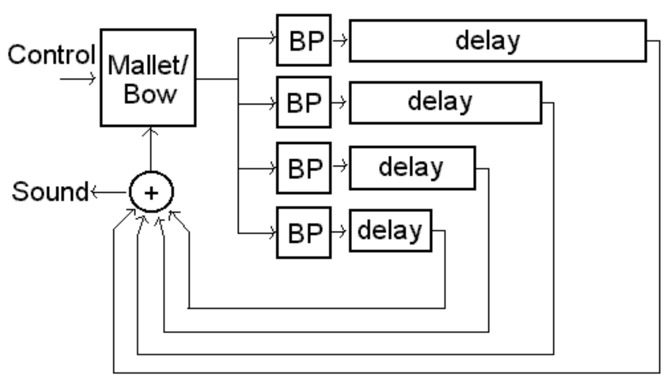
Acoustics:

Approximate model of wave train closure and neighboring frequencies

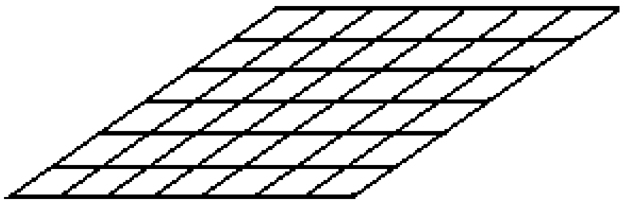
Filter:

Comb filter with only one resonance

(high vs. low Q)



Physical Models: 2D surfaces



2 (N) Dimensional Waveguide Meshes or

Finite Elements and Finite Differences

- *Discretize objects into cells (elements)*
- *Express interactions between them*
- *Express differential equation for system*
- *Solve by discrete steps in space and time*

2D Meshes, Finite Elements and Differences

Strengths

- (somewhat) arbitrary geometries
- Less assumptions than parametric forms
- Can strike, damp, rub, introduce non-linearities at arbitrary points

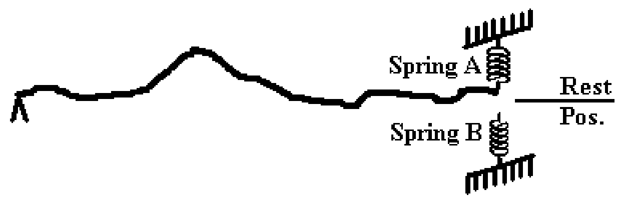
Weaknesses:

- Expensive
- Don’t know all the physics/solutions
- Sampling in space/time
- Dispersion is strange (diagonals vs. not)

Physical Models: Non-linearity

Add spring with position dependent constant

(one spring for positive displacement, another negative)



Acts to spread spectral components

Physical Models: Bars , Plates, Non-Linearities

Strengths:

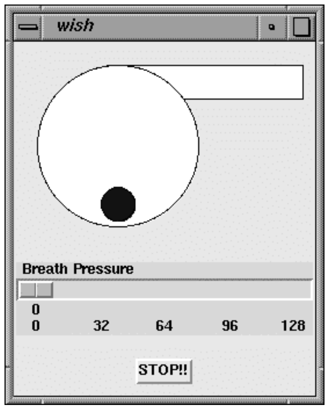
- Variety of sounds
- Computationally cheap
  (somewhat, under conditions)
- Physical (or pseudo) variables
- Naturalness of interaction

Weaknesses:

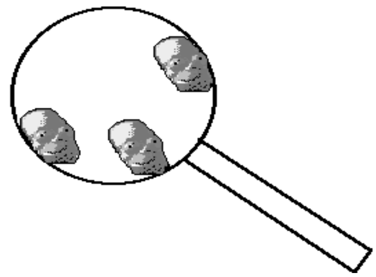
- No general system-ID techniques (2D + non-linear)
- Repeatability (not necessary a weakness)

Physical Models: Particles

Whistle: Single particle influences oscillator



Homeraca: Many particles launch PCM or parametric sounds



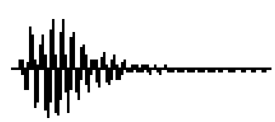
Stochastic Event Synthesis

Run model w/ lots of beans

Collect statistics -> Poisson

System energy decays exponentially.

Particle collision causes decaying burst of filtered noise



PhISEM Algorithm

- **Exponentially decaying system energy**
- **Particle sound energy is exponentially (fast) decaying white noise, but sum of exponentially decaying noises is an exponentially decaying noise.**
- **Each time step, compute likelihood (based on number of particles) of new sound-producing event**
  - ***If so, add to net particle sound envelope energy***
- **Filter result with net system resonances, with reallocation if needed**

Particles and PhISM

Strengths:

- Cheap
- Meaningful parameters
- Good for lots of real-world sounds

Weaknesses:

- No complete system ID (analysis) process (yet)

Related Techniques

Granular Synthesis ***(Many authors)***

Cut sound and randomly remix

Wavelets ***(Miner '99)***

Time/Freq transform (next slide)

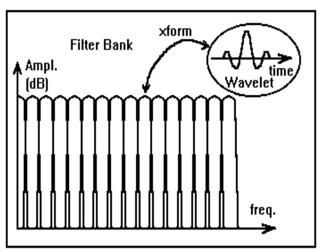
Independent Components Analysis ***(Casey 98)***

Interactive Sinusoidal Modeling ***(Pai et al)***

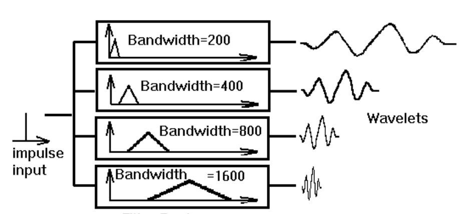
Wavelets

Filterbank view:

FFT is a filterbank but with evenly spaced frequencies and poor freq/time tradeoffs



Wavelets help to "cheat" time / frequency tradeoff and uncertainty



Physically Oriented Library of Interactive Sound Effects

Immersive Virtual Reality (games)
Augmented Reality (games)
Telepresence (games)
Interactive Art
Movie Production (or games)
Auditory display
Other Structured Audio Applications
Because we can, and want to
Much interesting work to do:
friction, psychoacoustics, ...

Synthesis ToolKit in C++

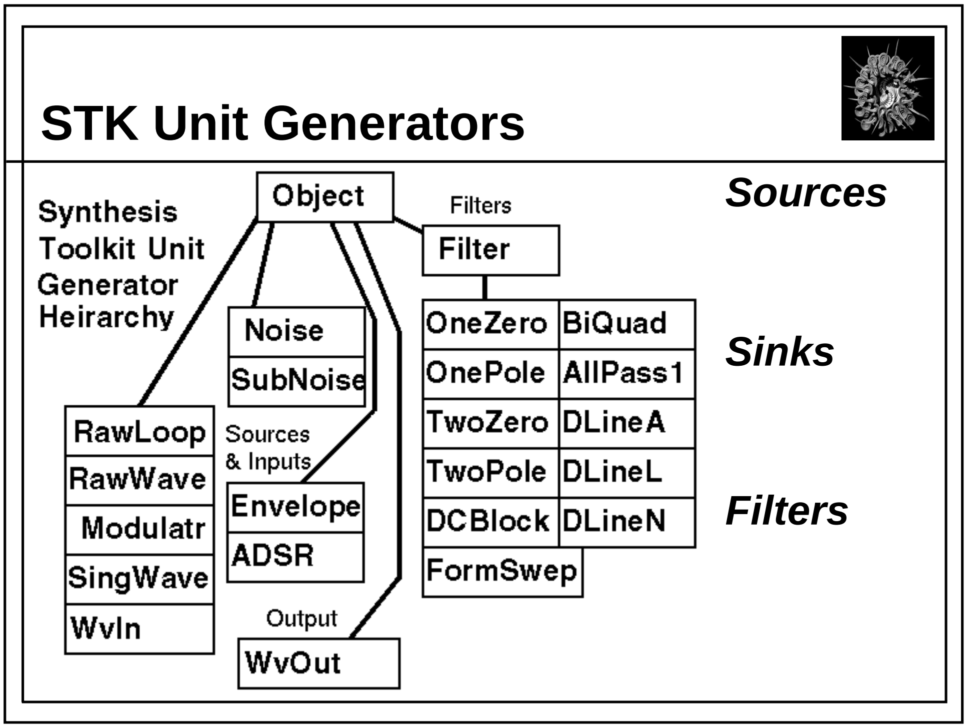
STK is a set of classes in C++ which allow rapid experimentation with sound synthesis and processing. Available for free:

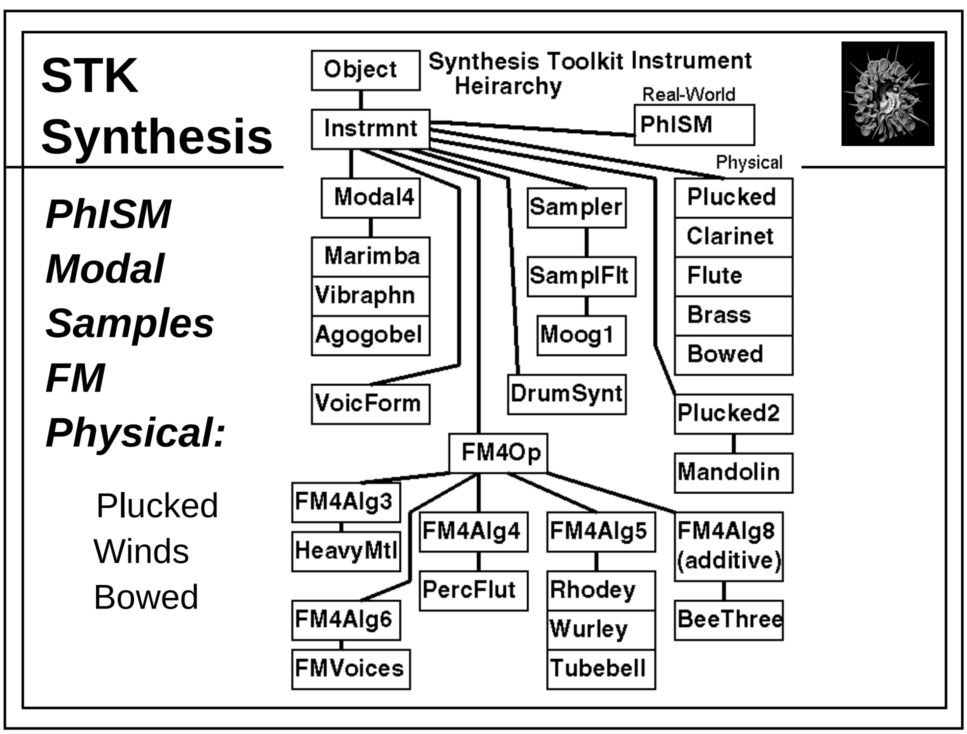
http://www.cs.princeton.edu/~prc

http://www-ccrma.stanford.edu/~gary

"Unit Generators" are the classical computer music/sound building blocks:

Oscillators, Filters, Delay Lines, etc.


STK Unit Generators
Synthesis Toolkit Unit Generator Heirarchy
Object
Filters
Filter
Noise
SubNoise
OneZero
BiQuad
OnePole
AllPass1
TwoZero
DLineA
TwoPole
DLineL
DCBlock
DLineN
FormSwep
RawLoop
RawWave
Modulatr
SingWave
WvIn
Sources & Inputs
Envelope
ADSR
Output
WvOut
Sources
Sinks
Filters


STK Synthesis
PhISM
Modal
Samples
FM
Physical:
Plucked
Winds
Bowed
Object
Synthesis Toolkit Instrument Heirarchy
Instrmnt
Real-World
PhISM
Physical
Modal4
Marimba
Vibraphn
Agogobel
VoicForm
Sampler
SamplFlt
Moog1
DrumSynt
Plucked
Clarinet
Flute
Brass
Bowed
Plucked2
Mandolin
FM4Op
FM4Alg3
HeavyMtl
FM4Alg4
PercFlut
FM4Alg5
Rhodey
Wurley
Tubebell
FM4Alg6
FMVoices
FM4Alg8 (additive)
BeeThree

SKINI:
Synthesis toolKit Network Interface

Double Precision floats for:

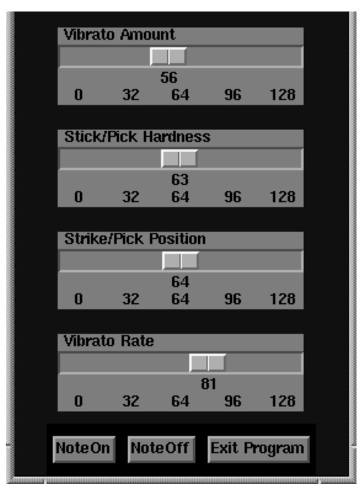
- Note Numbers (micro tuning or fine pitch control)
- Control Values (more precision)
- Delta times

Text Based (easy creation, editing, debugging)

Sockets (Pipes)

- Connection on local machine is same as on remote
- SKINI sources:
  - GUIs, MD2SKINI, Scorefiles, Any formatted text generator
- SKINI11.cpp parses SKINI messages

STK GUIs in TCL / TK



Common simple controls for all algorithms

PICOs (sound effects controllers)

K-Frog

J-Mug

P-Pedal

PhilGlas

P-Grinder

Pico Glove

T-bourine

Data Driven Sound: "Music for Unprepared Piano"

SIGGRAPH 98

with

Bargar

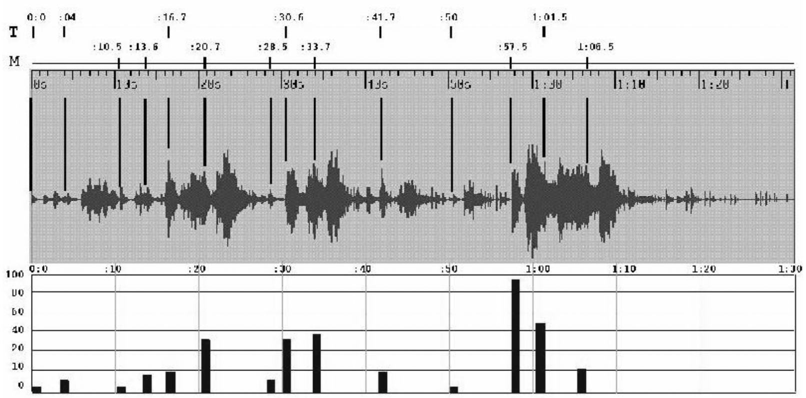
Choi

Betts

(NCSA)

Music for Unprepared Piano

The "Score"



References: Sinusoidal 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**.

References: LPC and Subtractive

Atal, B. 1970. "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave." Journal of the Acoustical Society of America 47.65(A).

Markel, J. and A. Gray, 1976, *Linear Prediction of Speech*, New York, Springer.

Moorer, A. 1979, "The Use of Linear Prediction of Speech in Computer Music Applications," Journal of the Audio Engineering Society 27(3): pp. 134-140.

Rabiner, L. 1968. "Digital Formant Synthesizer" Journal of the Acoustical Society of America 43(4), pp. 822-828.

Klatt, D. 1980. "Software for a Cascade/Parallel Formant Synthesizer," Journal of the Acoustical Society of America 67(3), pp. 971-995.

Carlson, G., Ternström, S., Sundberg, J. and T. Ungvary 1991. "A New Digital System for Singing Synthesis Allowing Expressive Control." Proc. of the International Computer Music Conference, Montreal, pp. 315-318.

Kelly, J., and C. Lochbaum. 1962. "Speech Synthesis." Proc . Fourth Intern. Congr. Acoust. Paper G42: pp. 1-4.

Cook, P. 1992. "SPASM: a Real-Time Vocal Tract Physical Model Editor/Controller and Singer: the Companion Software Synthesis System," Computer Music Journal, 17: 1, pp. 30-44.

References: Modal (Filter) Synthesis

Rossing, T. 1976. "Acoustics of Percussion Instruments - Part 1," The Physics Teacher, 14, pp. 546-556.

Serra, X. 1986. "A Computer Model for Bar Percussion Instruments," Proc. ICMC, The Hague, pp. 257-262.

Wawrzynek, J. 1989. "VLSI Models for Sound Synthesis," in Current Directions in Computer Music Research, M. Mathews and J. Pierce Eds., Cambridge, MIT Press.

Adrien, J.M. 1991, "The Missing Link: Modal Synthesis", in: G. De Poli, A. Picalli, and C. Roads, eds. *Representations of Musical Signals.* MIT Press, Cambridge, Massachusetts.

Doutaut V. & A. Chaigne 1993. "Time Domain Simulations of Xylophone Bars," Stockholm Music Acoustics Conference, pp.574-579.

Larouche, J. & J. Meillier 1994. "Multichannel Excitation/Filter Modeling of Percussive Sounds with Application to the Piano," IEEE Trans. Speech and Audio, pp. 329-344.

P. Cook 1997, “Physically Inspired Sonic Modeling: (PhISM): Synthesis of Percussive Sounds,” *Computer Music Journal*, 21:3 (expanded from ICMC 1996).

References: Waveguide Modeling

For many references to specific physical models, please consult the reference sections of the papers attached as appendices to the course notes. Other physical modeling papers:

Computer Music Journal, 1992-3, Two Special Issues on Physical Modeling, MIT Press, Vol. 16 No. 4 and Vol. 17 No. 1, Winter 1992 and Spring 1993.

Van Duyne, S. and J. Smith 1993. "Physical Modeling with the 2-D Digital Waveguide Mesh." In Proceedings of the ICMC, Tokyo, pp. 40-47.

J.O. Smith, “Acoustic Modeling Using Digital Waveguides,” in Roads et. al. eds., Musical Signal Processing, Netherlands, Swets and Zeitlinger, 1997.

Pierce, J. R. and Duyne, S. A. V. 1997, A passive non-linear digital filter design which facilitates physics-based sound synthesis of highly nonlinear musical instruments. *Journal of the Acoustical Society of America*, 101(2):1120-1126.

References: Friction

Siira J. and Pai D.K. 1996, “Haptic Textures, A Stochastic Approach,” IEEE International Conference on Robotics and Automation.

Fritz, J.P and Barner K. E. 1996, “Stochastic Models for Haptic Texture,” Proceedings SPIE Intl. Symposium on Intelligent Systems and Advanced Manufacturing.

Hayward, V., Armstrong, B. 1999. A new computational model of friction applied to haptic rendering. Preprints of ISER'99 (6th Int. Symp. on Experimental Robotics).

References: Confined Turbulence

Verge, M. 1995. *Aeroacoustics of Confined Jets, with Applications to the Physics of Recorder-Like Instruments*. Thesis, Technical University of Eindhoven, also available from IRCAM.

References: PhISEM, Wavelets, Granular, ...

Rodet, X. 1984. "Time-Domain Formant-Wave-Function Synthesis," Computer Music Journal 8 (3), pp 9-14.

Roads, C. 1991. "Asynchronous Granular Synthesis" In G. De Poli, A. Piccialli, and C. Roads, eds. 1991. Representations of Musical Signals. Cambridge, Mass: The MIT Press, pp. 143-185.

C. Cadoz, A. Luciani and J. Florens, 1993, “CORDIS-ANIMA: A Modeling and Simulation System for Sound Image Synthesis-The General Formalization” Computer Music Journal, Vol. 17, No. 1, pp. 21 - 29.

P. Cook, 1997, “Physically Inspired Sonic Modeling: (PhISM): Synthesis of Percussive Sounds,” Computer Music Journal, 21:3.

N. Miner, 1998, Creating Wavelet-based Models for Real-time Synthesis of Perceptually Convincing Environmental Sounds, Ph.D. Diss., University of New Mexico.

M. Casey, 1998, Auditory Group Theory with Applications to Statistical Basis Methods for Structured Audio, Ph.D. Dissertation, MIT Media Lab.