

# CSE481i: Digital Sound Capstone

## An Overview



(Material freely adapted from sources far too numerous to mention...)

## Today



- What this course is about
  - Place & time
  - Website
  - Textbook
  - Software
  - Lab
  - Topics
- An overview of Digital Sound Engineering

## Where and when



- We will meet Tuesdays and Thursdays, 10:30-11:50 p.m. in cse 203.
- Lab is Tuesday 2:30-5:20 in CSE 003E
- This week:
  - Check your card access and login in the lab.

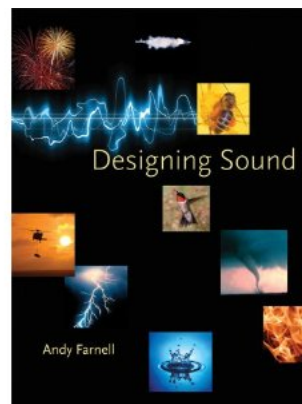
## Recommended Textbook



RECOMMENDED textbook:

**Designing Sound**  
Andy Farnell  
The MIT Press

NOT Required, but recommended  
Some pdf excerpts are available



## Other Useful Books



**The SuperCollider Book**

Scott Wilson, David Cottle, Nick Collins and James McCartney

**The Audio Programming Book**

Edited by Richard Boulanger and Victor Lazzarini

**The Computer Music Tutorial**

Curtis Roads

**Real Sound Synthesis for Interactive Applications**

Perry R. Cook

**Unity Game Development Essentials**

Will Goldstone

More on the website

## Software



Major packages we'll use:

**Pure Data**

graph-based sound language public domain

**SuperCollider**

client-server based sound synthesis language

**The Synthesis ToolKit in C++ (STK)**

physical modeling toolkit

**Audacity**

Sound editor

**Kinect API**

Human Sensing

## The Lab- CSE 003E



- This room is available for your use 24 hours a day
- My office hours are the lab times;  
you can visit me anytime in CSE 464
- You can do your work anytime, but the lab is where you can get help

## Tentative topics



- Introduction and overview
- Fundamentals of Acoustics
- Fundamentals of Digital Audio
- Oscillators
- Modulation
- Sampling
- Non-linear Synthesis
- Granular Synthesis
  
- Sampled sound for animation
- Human sensing—Kinect API, Leap Motion API
- Presentations
- Capstone project planning

# Digital Sound Engineering

## An Overview



## Prehistoric and ancient music



Daniel Maurer / AP

## Music: the old ways:



## Music: the new way:



Or this:



Or this:



Or this:



iPad Apps for Music Making:



Sonoma's StudioTrack



## iPad Apps for Music Making:



Sound Trends' **Looptastic**

## iPad Apps for Music Making:

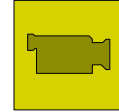


**AC-7 Pro Controller**

## iPad Apps for Music Making:



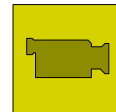
**Air Harp Instrument**



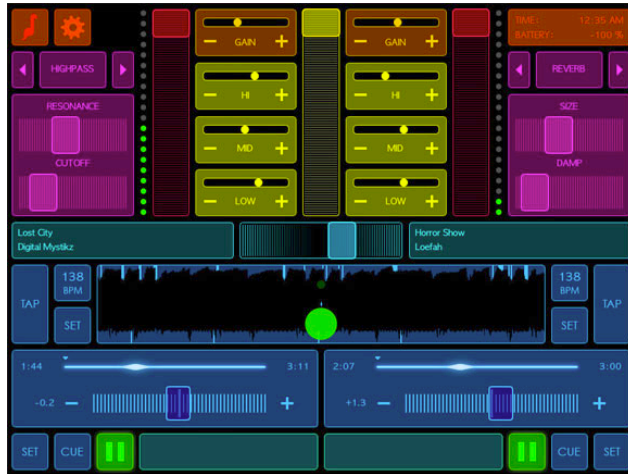
## iPad Apps for Music Making:



**Smule Magic Piano, the spiral/radial piano**

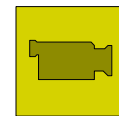


## iPad Apps for Music Making:



Sonorasaurus Rex DJ app

## iPad Apps for Music Making:



iPad app for DJs: VirtualDeck

## So, how does the iPad make music?

Or, How do we produce sound  
with a computer?



## Computer Audio

- An interdisciplinary field including
  - Music
  - Computer Science
  - Electrical Engineering (signal processing)
  - Physics (musical acoustics)
  - Psychology (psychoacoustics, music perception)
  - Physiology (hearing, speech and singing)



## Computer Audio Taxonomy



- Signal Processing
  - Sound Analysis and Resynthesis
  - Physical Modeling of Musical Instruments and Speech
  - Musical Effects
  - 3D Spatialization
  - Audio Coding and Compression
  - Audio Signal Separation
  - Music Signal Pitch Detection
- AI
  - Machine Recognition of Audio and Music
  - Musical Instrument Recognition
  - Music Perception and Cognition
  - Psychoacoustics
  - AI and Music

## Computer Audio Taxonomy



- Software
  - Music Visualization
  - Music Composition Systems and Tools
  - Music Programming Languages
  - Algorithmic Composition
  - Music Notation and Printing
  - Music on the Internet
  - Music in Computer Games
  - Sound Effects in Computer Games
  - Computer Music and Digital Art
- Database
  - Music Information Retrieval
  - Musical Informatics
  - Music Databases

## Computer Audio Taxonomy



- Computer Engineering
  - Audio Hardware
  - Music Performance Interfaces (new musical instruments)
  - Interactive Performance Systems
  - Real Time Performance Systems
  - Music Workstations
  - Soundcards
  - Music Synthesizers
  - Music and Audio on Mobile Phones
  - Wireless Audio Systems
  - Music Networks
  - MIDI

## Computer Audio Taxonomy

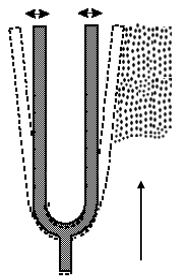


- Theory/Science
  - Music Data Structures and Representation
  - Musical Tuning and Intonation
  - Music Grammars
  - Musical Acoustics
  - Acoustics of Musical Instruments and the Voice

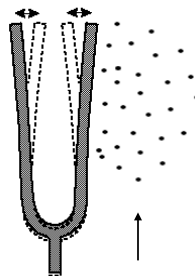
## What is Sound?



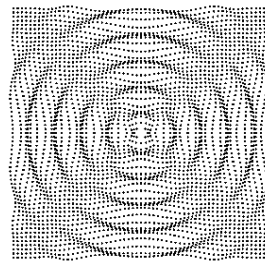
- Variation in air pressure caused by compression and decompression of molecules
- Caused by friction producing force (stick striking symbol, finger plucking guitar string)
- 'Waves' produced by cohesion of molecules, which fall on eardrum or microphone
- Directly and through reflection off surfaces in room
- Ear can detect frequencies in the range 20Hz to 20kHz
- Ear has very high dynamic response compared with eye (ie ability to detect changes in pressure)
- Requires much higher sampling rates to digitize audio compared with images



(a) compression



(b) rarefaction



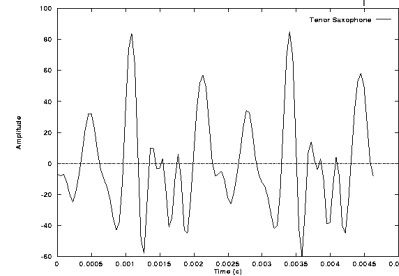
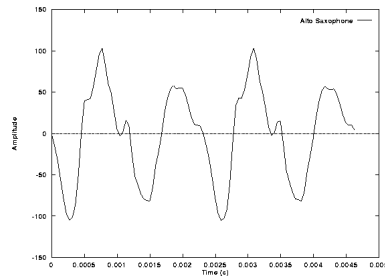
(c) wave propagation of a tuning fork  
as seen from above

## Properties of sound...



- Waveform – distinctive pattern of variations in air pressure
  - Musical instruments produce orderly repeating waveforms
  - Noise produces random chaotic waveforms
- Fourier demonstrated how any wave form can be decomposed into a series of component sine waves of different frequencies
- Different frequency components, or pure tones, which are added together to produce a complex waveform are called the **frequency spectrum** of that waveform

## Same note.. different waveforms



- Both figures show an 'A' note, left played on an alto sax and the right on a tenor sax.
- Both have additional frequencies as well as the main 440Hz

## Physical and subjective attributes..



- Important to distinguish between the properties of a stimulus and those of a subjective response to that stimulus
- A linear increase in the stimulus value does **not** necessarily produce a similar increase in the subjective response

<u>Stimulus value</u>	<u>Subjective response</u>
(luminance)	(brightness)
Amplitude of wave	Loudness of sound
Frequency of wave	Pitch of sound
Several attributes (hard to define)	Timbre of sound



## Amplitude and Frequency



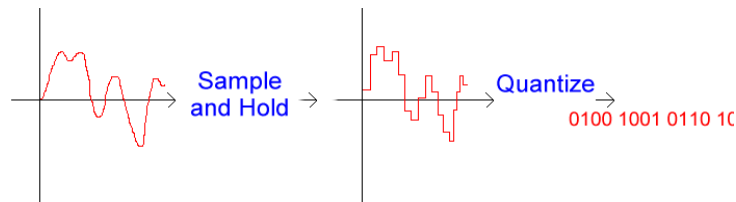
- **Amplitude** measured in decibels
  - The louder a sound is, the more it will mask or dominate other other sounds adjacent to it in time
- **Frequency** measured in cycles per second (Hertz – Hz)
  - More digital information required to encode higher frequency sounds, lower pitched sounds degraded less by low sample rates
- **Timbre**
  - loosely defined by ‘tone’, ‘color’, ‘texture’ of sound that enables brain to differentiate one tone from another
  - Affected by acoustic properties of instruments and room

## Digitizing sound



- Analog signal is sampled and converted to a series of digital values (A to D converter)
- Digital values later converted back to analog for playback through speakers (D to A conversion)
- Parameters are **frequency** at which samples are taken and the **resolution** of each sample (i.e number of bits used to encode analog signal value)
- Nyquist’s theorem prescribes minimum sample rate in order to be able to re-construct analog signal
- If maximum frequency in the waveform is  $n$  Hz, then minimum sample rate should be  $2n$  Hz

## Sampling and Quantizing



- Sampling – process of acquiring an analog signal
- Quantizing – conversion of held signal into sequence of digital values

## Sample rates

- If upper range of ear is 20Khz, then there is no need to faithfully reproduce frequency components in signals higher than this.
- **CD quality:** at least  $2 \times 20\text{KHz} = 44.1\text{KHz}$
- The human voice has few frequency components lower than 100Hz, or higher than 3000Hz - a **bandwidth** of 2900Hz
- **Speech:** at least  $2 \times 2.9\text{KHz} = 8\text{KHz}$

## Sample data rates



- For CD quality,
  - Rate = 44.1Khz (44100 samples per second)
  - Resolution = 16 bits
  - Stereo = 2 channels
- Data rate =  $44100 * 16 * 2$  bits/second = 1411200 bits/sec
- (10Mb storage for 1 minute of recorded sound)

## Examples of data rates and quality



Sample Rate	Resolution	Stereo/Mono	Bytes (1 min)
44.1 KHz	16 bit	Stereo	10.1 Mb
44.1 KHz	8 bit	Mono	2.6 Mb
22.05 KHz	16 bit	Stereo	5.25 Mb
22.05 KHz	8 bit	Mono	1.3 Mb
11 KHz	8 bit	Mono	650 Kb
5.5 KHz	8 bit	Mono	325 Kb

CD quality audio

As good as a TVs audio

As good as a bad phone line

## Digitized vs. Synthesized



Multimedia sound comes from two sources:

- Digitized – from an external (sampled) real life sound
- Synthesized – created from waveforms in a sound card for example

Traditional analog sound synthesis is achieved by

- Creating a waveform using an oscillator, which sets the basic frequency
- Adding an "envelope", by specifying parameters such as attack, decay, sustain, release
- Then sending through filter(s) to modify timbre

## MIDI – Musical Instruments



- Digital Interface – supported by many instruments/ computers/ manufacturers (1980)
- Defines set of messages indicating note/ instrument/ pitch/ attack etc
- Sound card/ Synthesizer takes this symbolic message and 'creates' matching sound
- Sampled sounds can be stored by users on better equipment
- Compare waveforms to bitmapped images, midi to vector graphics

# Digital Sound Synthesis Methods



Depending on your age, you might think the first synthesizer looked something like this:



However, it looked more like this:



## Synthesis Definition

The Oxford Classical Dictionary defines **synthesis** as:

- Combination, composition, putting together
- Building up of separate elements into connected whole



## Synthesis Definition



- Generally, most people associate synthesis purely with subtractive synthesis
- Very limiting way to look at sound synthesis by electronic means

## The Bigger Picture

- Theoretically sound divisions, but practically limiting
- Techniques of different types applicable to others

### Subtractive

Sampling

Analog

Granular

Waveshaping

FM

Physical Modeling

### Additive



## In the beginning...

- Additive synthesis
- Principle first utilized in cathedral organs:



## Additive Synthesis

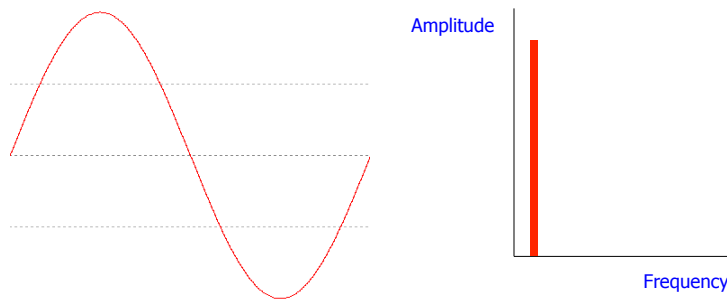
- Mathematical basis:
  - 1822: Jean Baptiste Joseph, Baron de Fourier published theory:
    - *Any arbitrarily complicated periodic waveform can be deconstructed into combinations of sine waves of different amplitudes, frequencies and phases*
- This is accomplished by the *Fast Fourier Transform*: **FFT**



## Additive Synthesis



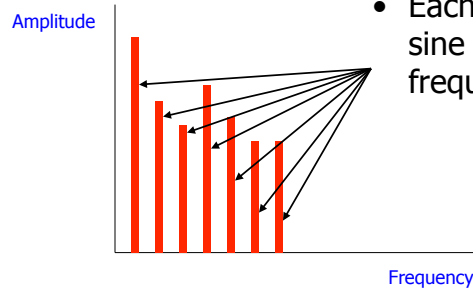
- Sine wave = simplest possible waveform
- Contains only the fundamental



## Additive Synthesis



- A more complex waveform will be composed of any number of sines of varying frequencies and amplitudes:



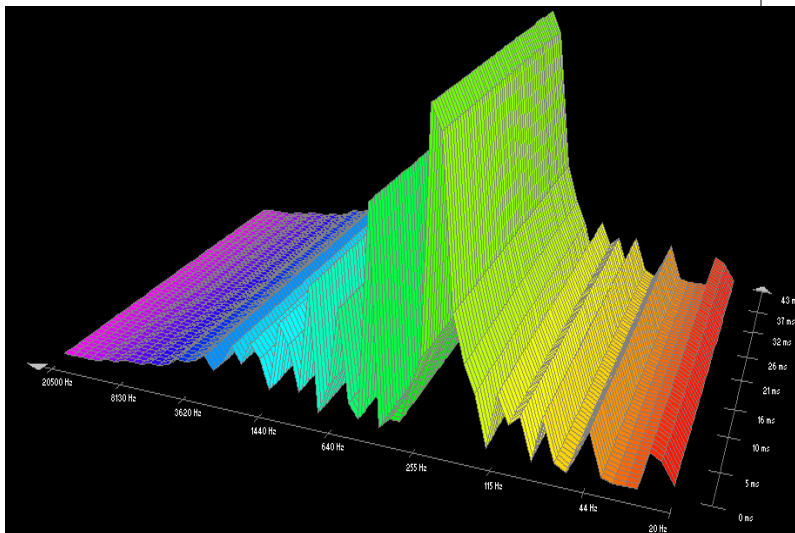
- Each line represents a sine at a specific frequency and amplitude

## Additive Synthesis

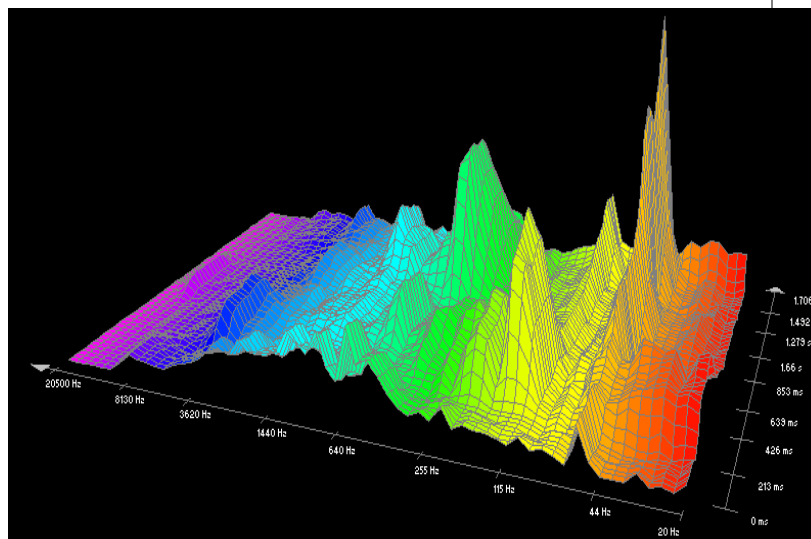
- But this simple approach hides many difficulties
- Theory shown so far deals with a single moment in a sound's duration
- Most sounds are complex and evolving



## Sawtooth Wave



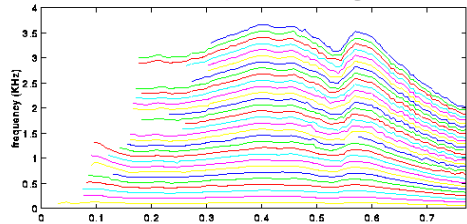
## Complex Wave



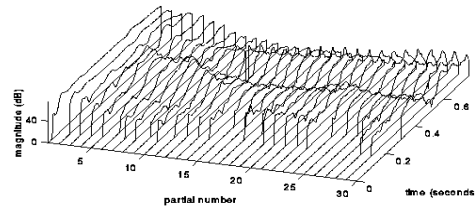
## Additive Synthesis

- Thus, will have multiple slices depending on:
  - Length of waveform
  - Rate of change of waveform
- Control data therefore massive
- Very hard to create sounds using additive synthesis
- Holy Grail: **Analysis-Based Resynthesis**

# Sinusoidal Analysis "Tracks" (McAuley and Quatieri)



frequency of partials



magnitude of partials

# Sinusoidal Additive Synthesis



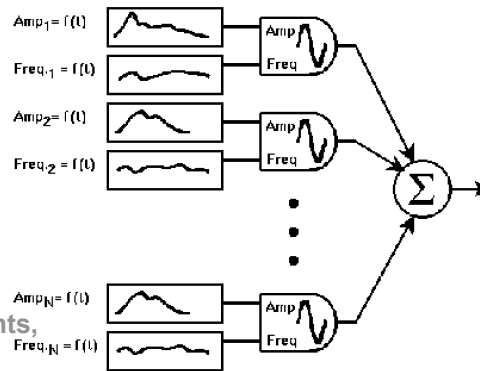
Control the amplitude and frequency of a set of oscillators

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

Additive Synthesis Block Diagram



## The Bigger Picture

- Theoretically sound divisions, but practically limiting
- Techniques of different types applicable to others

### Subtractive

Sampling

Analog

Granular

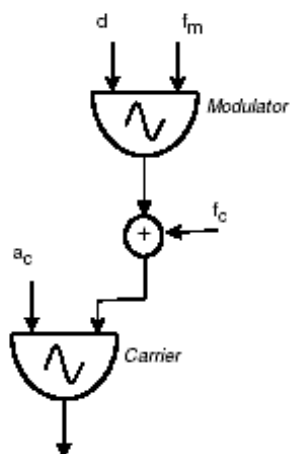
Waveshaping

FM

Physical Modeling

### Additive

## FM Synthesis

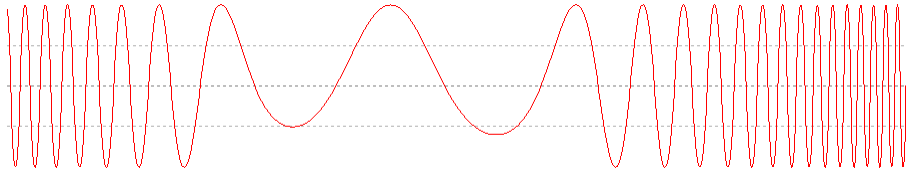


- Simple FM: carrier oscillator has its frequency modulated by the output of a modulating oscillator.
- Sidebands produced around carrier at multiples of modulating frequency.
  - Number generated depends on the amplitude of the modulator.

## FM Synthesis



- Nothing more than an extreme form of vibrato:



- When the modulation is fast enough, we no longer hear the rise and fall of the vibrato
- Instead, we perceive the changes in pitch as changes in the timbre of the sound

## Modulator : Carrier Ratio



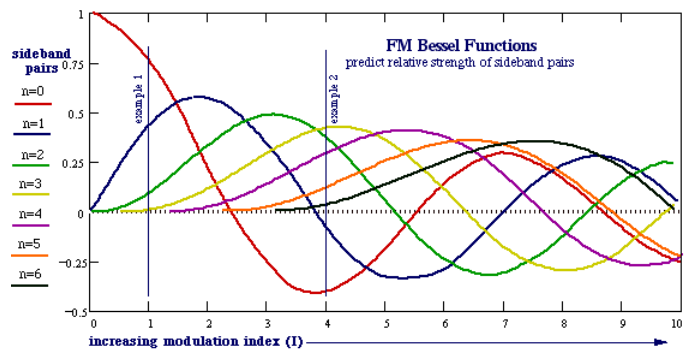
- Sidebands at  $C +$  and  $- (n * \text{Modulator})$
- Ratio of M:C determines whether spectrum is harmonic or not.
  - Simple integer ratio = harmonic
  - Non-integer ratio = inharmonic

## Modulation Index and Bandwidth

- The *bandwidth* of the FM spectrum is the number of sidebands present.
- The bandwidth is determined by the *Modulation Index*
  - $I = \text{depth of modulation} / \text{modulator}$
  - $D$  depth of modulation, which depends on the amount of amplitude applied to modulating oscillator. ( $D = A \times M$ )
- If the index is above zero, then sidebands occur.

## FM Synthesis

$$J_n(\beta) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{\beta}{2}\right)^{n+2k}}{k! \cdot (n+k)!}$$



## FM Synthesis



- Unfortunately, the relationship between these is not predictable without experience:
  - as the Index changes, the amplitude of each sideband pair evolves in a different pattern
  - some sidebands gain amplitude, others lose amplitude
  - there may also be cancellation effects caused by phase-inverted sidebands.
- This remains the most significant barrier to learning FM synthesis
- Nevertheless a powerful technique for creating complex sounds

## Granular Synthesis

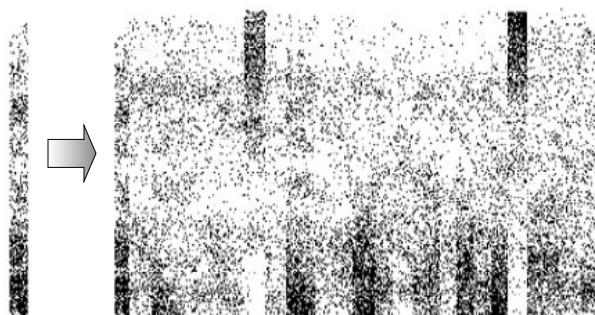


- Attempt to deal with the shortcomings of additive synthesis to deal with changes in the sound over time
- 1947: Dennis Gabor, physicist formulated theory:
  - sound is perceived as a series of short, discrete bursts of energy, each slightly changed in character from the last
- Rooted in quantum physics – coexistence of the wave and photon in light
- Sonic equivalent of the photon is the **grain**



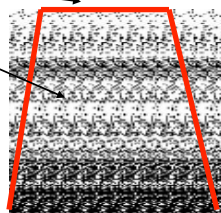
## Granular Synthesis

- *Definition:* generation of thousands of short sonic grains which are combined linearly to form large scale audio events
- Grain = tiny piece of sonic data, duration: 10 to 50 ms.



## Granular Synthesis

- Two components:
  - Envelope
  - Contents



- **NB:** Grain Density – number of grains per second
  - Low density leads to rhythmic effects



## Granular Synthesis



- Subject to same fundamental problem as additive synthesis, though:
  - Tension between precision and control
  - Massive number of grain events
- Basic unit -> **grain cloud** rather than grain itself
  - ~ Set of rules for generating and controlling grains
- It has some of the drawbacks of FM synthesis as well:
  - Unpredictable results
- But capable of creating sound textures that no other form of synthesis can

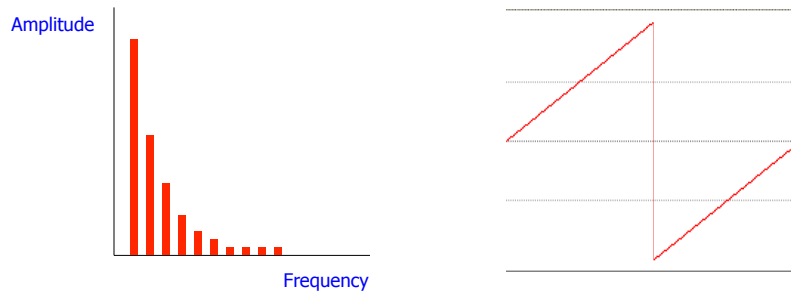
## Subtractive Synthesis



- Well understood and widely employed
- Begin with a harmonically rich sound source and remove frequencies by means of **filtering**
- While any sound source can be employed, traditionally associated with certain waveshapes

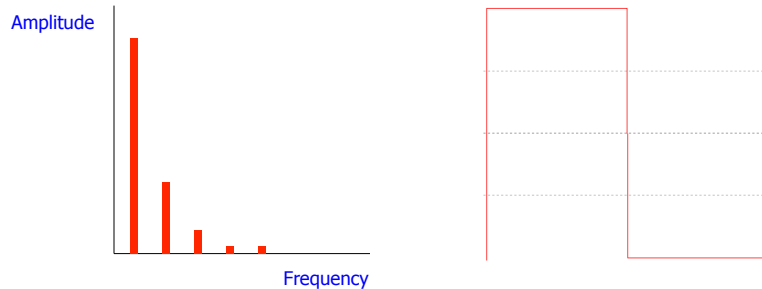
## Subtractive Synthesis

- Sawtooth: contains all harmonics, with amplitude  $1/n$ :



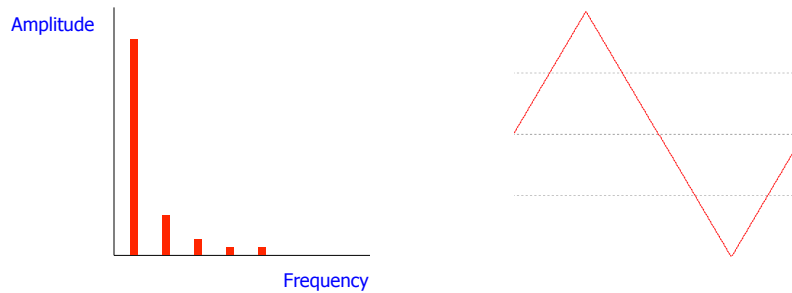
## Subtractive Synthesis

- Square: only odd harmonics present, also with amplitude  $1/n$



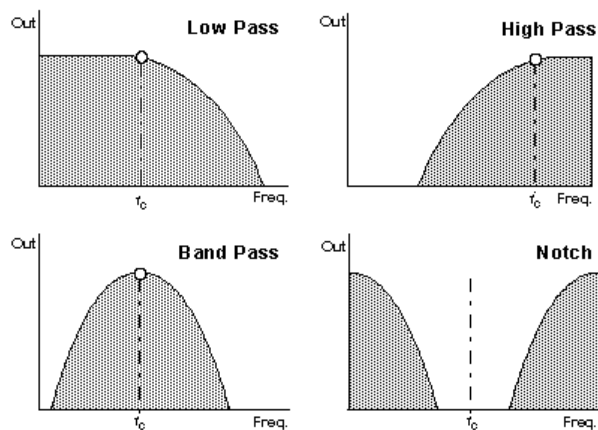
## Subtractive Synthesis

- Triangle: only odd harmonics present, but with amplitude  $1/n^2$



## Subtractive Synthesis

- Basic Filters



## Waveshaping Synthesis



- Sound of a waveform determined primarily by its harmonic content
- Can create new harmonics by passing waveform through non-linear element: *waveshaper*
- Often a Chebyshev polynomial

## Physical Modeling




- Modeling sound generation
  - more expressive and realistic sounds
  - ideal for software implementation
  - no need for dedicated hardware
- **Brute force approach**
  - solve equations of motion with respect to boundary conditions
- **Better Approach**
  - Partway solve equations for changing parameters
  - lookup tables
  - lumped processes
  - novel algorithms



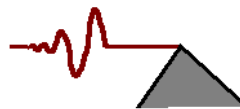
Yamaha VL1

## Digital Wave-guide Modeling

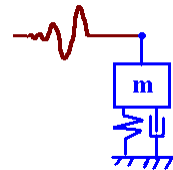
D'Alembert's Solution .. Waves travel in equal and opposite directions

$$f(ct + x) + g(ct - x)$$


Reflected and attenuated at boundary



Boundary behaviour frequency dependant



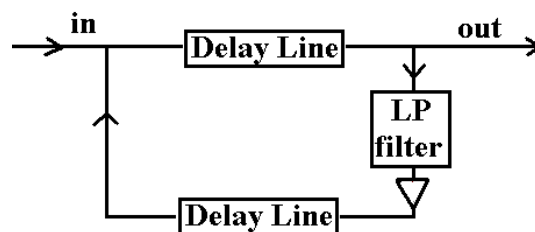
## Digital Wave Guide Modeling

Energy in (pluck, hammer, bow, etc)

**Delay line** simulates time wave travels

Feed back loop simulates reflection

Filter simulates frequency dependant attenuation



## Modeling an electric guitar



Array represents wave-guides

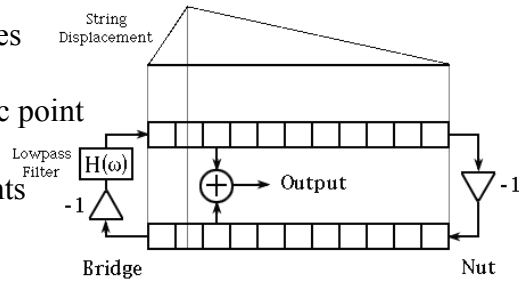
Add displacement at specific point

Data passes between elements

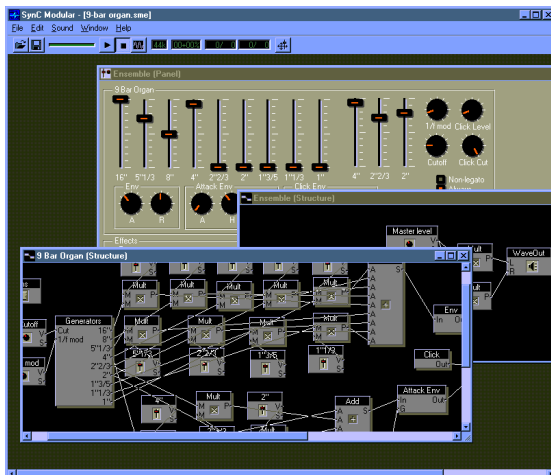
Low Pass Filter for bridge

Tap data at pickup point

Can hence delay, modify and add back as feedback



## Modular and Virtual Modular Synthesis



SynC modular

A modular paradigm allows for additive, subtractive, fm and sampling synthesis techniques to be used together