Lecture VI. 10-27-09

**Sound Synthesis** can be defined as the production and manipulation of sounds using mathematical algorithms.

**Unit Generator** – Music III and its descendants Music IV and V introduced Unit Generators. A Unit Generator (Ugen) is a building block of a sound synthesis algorithm. Examples of unit generators are oscillators, filters, multipliers and adders, and amplitude envelope generators. Different complex sonic patterns and sound synthesis algorithms could be implemented by connecting different oscillators.

I. Additive Synthesis: Modulation (Amplitude, Ring Modulation, FM)

**Additive Synthesis** is a synthesis technique derived from the *Fourier* theorem. Mathematically, the Fourier theorem states that a periodic function can be formulated as a sum of sine waves. When applied to computer music, the Fourier theorem can be interpreted as the possibility of creating any complex waveform by summing a set of sinusoidal components: this is the basic idea behind additive synthesis. In computer music, a sine wave is produced by an oscillator whose frequency, amplitude and phase can be varied.

**Amplitude Envelope** – A description of the manner in which the amplitude of a sound changes over time. Attack, Sustain, Decay, Release (ADSR)

**Wave Forms:**

**Sine Wave** - Periodic waveform, simple geometric waveform, which follows the values found in a sine table.

**Triangle Wave** - Like a square wave, the triangle wave contains only odd harmonics. However, the higher harmonics roll off much faster than in a square wave (proportional to the inverse square of the harmonic number as opposed to just the inverse).

**Sawtooth Wave** - A sawtooth wave's sound is harsh and clear and its spectrum contains both even and odd harmonics of the fundamental frequency. Because it contains all the integer harmonics, it is one of the best waveforms to use for synthesizing musical sounds, particularly bowed string instruments like violins and cellos, using subtractive synthesis.
Square Wave or Pulse Wave - A square wave is a kind of non-sinusoidal waveform, most typically encountered in electronics and signal processing. An ideal square wave alternates regularly and instantaneously between two levels.

Noise:

White Noise - is a random signal (or process) with a flat power spectral density. In other words, the signal contains equal power within a fixed bandwidth at any center frequency.

Pink Noise - A random signal of every frequency in which each higher octave drops off 3 dB. The lower octaves have more power, and the higher octaves have less power. Pink noise is used to test loudspeakers and "tune" a room for optimum audio reproduction or masking systems.

Brown Noise - It decreases in power by 6 dB per octave and, when heard, has a "damped" or "soft" quality compared to white and pink noise. The sound is a low roar resembling a waterfall or heavy rainfall.

Filters:

High Pass - a device that attenuates signals of frequency lower than some specified value.

Low Pass - a device that attenuates signals of frequency higher than a specific value.

Band Pass - filter that attenuates signals of frequency within specified limits.

Q - quality of a filter circuit that describes, among other things, the tendency to accentuate signals close to the cut-off frequency.

Play Risset “Mutations” trk.11 CD 2 c.1969 (LCD8975)
**Amplitude Modulation** - If an oscillator is connected to the control input of an amplifier that is processing another oscillator, two interesting possibilities arise. If the control oscillator is running at sub audio frequency, a vibrato of amplitude is produced, similar to that performed on wind instruments. If the control oscillator is running at audio frequency, the result is a mix of the two frequencies, plus extra components at frequencies equal to the sum of the two.

**Ring Modulation** – Produces side bands associated with amplitude modulation but suppresses the original frequency content. The output is a very messy spectrum, especially if one of the input signals is harmonically complex to start with. A further refinement of the same device suppresses sidebands above or below the input signal, producing **frequency shifting**. (The upper sidebands are produced by adding some constant value to the frequencies of all components of one input, pushing the whole spectrum off of the harmonic series)
FM:
John Chowning while at Stanford’s Artificial Intelligence Lab, which later became the Center for Computer Research in Music and Acoustics (CCRMA) started to explore the musical potentials of computer-generated sounds. While playing with combined oscillators, he discovered what is still nowadays the most successfully commercial sound synthesis technique known as FM synthesis.

Elsea, Page 128 - 130

**Frequency Modulation** is a very powerful algorithm for creating sounds. The heart of the technique is the way extra tones (sidebands) are created when one oscillator is used to modulate the frequency of another. These sidebands are symmetrically spaced about the frequency of the carrier, and the size of the spaces is equal to the frequency of the modulator. Increasing modulation increases the number of sidebands, but the amplitude of the sidebands varies in a rather complex way as the modulation changes.

In FM, the instantaneous frequency of a carrier wave is varied according to a modulation wave, such that the rate at which the carrier varies is the frequency of the modulating wave, or modulating frequency. The amount the carrier varies around its average, or the peak frequency deviation (Index), is proportional to the amplitude of the modulating wave.

- **Carrier** - the oscillator we listen to
- **Modulator** – an oscillator that changes the frequency of the carrier at an audio rate.
- **Index** also called **modulation depth** - indicates by how much the modulated variable varies around its ‘original’ level.
II. Institute de Recherche et Coordination Acoustique Musique (IRCAM)

The Sogitec 4X digital synthesizer developed at IRCAM during the 1980's by Peppino Di Guigno, was the direct result of Luciano Berio's demand that “live electronic sound ought to consist of at least a 1000 sine wave oscillators, in order to be interesting for the composer and the audience alike.” It was the last huge hardware, that IRCAM instigated before it turned to the ISPW and subsequently to software solutions in the shape of jMax. It’s eight internal custom-built processors were capable of altogether 200 MIPS (2), which being the equivalent of 1000 sine waves, 500 filters or 450 second order filters (3). Each processor contained a data-memory, an address-memory, a micro-program memory and a function--memory. For calculations it used 24 bit fixed-point units consisting of a multiplier, an arithmetic and logic unit. It also had 256 internal (programmable) clocks and a large dual buffer for recording and playing. The algorithms were cross-developed on DEC en Sun mainframes in Fortran and C.

Show Picture of Sogitec 4X

III. Granular Synthesis

In musical terms, a sound grain can be defined as a short sonic snippet of about ten to a hundred milliseconds, an elementary particle as opposed to a complex soundscape. By combining different grains over time, and by overlapping several grains at the same instance of time, interesting sonic effects can be produced. The synthesis technique in which different sound grains are combined is known as granular synthesis. One of the pioneers of the use of granular synthesis in computer music is Curtis Roads. Working together with his teacher Iannis Xenakis, he investigated the idea of composing with sound particles, an idea mainly inspired by the theory of the Nobel Prizewinner Dennis Gabor, who claimed that all sounds can be considered as being made of elementary sound particles limited in time, frequency and amplitude.

Granular Synthesis uses statistical methods to form sounds from masses of brief acoustical events (usually between 2 and 50 ms) called grains.
In asynchronous granular synthesis grains are scattered probabilistically over a specific duration within regions called clouds.

**Granular Synthesis parameters:**

1. **Start time and duration of the cloud;**

2. **Amplitude envelope of the cloud;**

3. **Grain duration** this may be constant (the same for every grain in the cloud), randomly variable within boundaries (10 to 100 ms), or frequency dependent (the duration of each grain is tied to its fundamental frequency period, as in synthesis with wavelets) - shorter grains will generally produce wider spectrum clouds, midsize grains produce fluttering/gurgling effects, and longer grains produce AM/tremolo-like results - in general, grain duration affects the bandwidth of the cloud.

4. **Grain envelope** usually a bell-shaped (Gaussian) envelope, to minimize sharp transitions between grains.

5. **Density of grains per second** which can vary over the duration of the cloud - its effect will depend also on the grain duration, since longer grains will overlap more with each other.

6. **Frequency band of the cloud** usually given as high and low frequency bounds within which grains are scattered (cumulus), but also restricted to discrete sets of pitches (stratus) - narrow bands and high densities create pitched streams with formant spectra, medium bands (like intervals of several semitones) and high densities generate colored noises, wide bands (an octave or more) and high densities form massive blocks of sound.

7. **Waveforms** any waveform may be inserted into grains, beginning with sine waves, but including sampled waveforms - grain waveforms may change within the duration of a cloud, creating statistically controllable mutation effects.

8. **Spatial dispersion** of the cloud among output channels.

**Time granulation of sampled sounds:**

1. **Time/rate changing effects** the playback rate of a granulated signal is changed with respect to the rate at which it was sampled. To double duration but not pitch, clone each grain of the original once, and to halve
duration but not pitch, delete every other grain. To shift pitch without changing duration, change the playback sample rate while cloning or deleting a proportional number of grains.

2. Real-time granulation of an input sound a programmable delay line is tapped to furnish the various grains, enriching the input sound with the spectral products of the granulation window and the mixture of the original and delayed sound.

3. Asynchronous granulation of different samples, to create mutations or cross-sampling effects.

Give example using SuperCollider

Play “nscor” by Charles Curtis c.1980 remix 1986 (LCD200)

IV. Sound Modeling

Among the different synthesis techniques, sound modeling techniques have seen the greatest interest from acoustician, engineers, computer scientists and composers. While acousticians are interested in understanding how different musical instruments produce sound, engineers and computer scientists are interested in developing efficient yet accurate algorithms to simulate such sounds, and musicians and composers are interested in using modeling techniques to extend the sonic possibilities offered by traditional instruments.

Sound modeling techniques are commonly divided into spectral models and physical models. While spectral models simulate how a sound is perceived by the listener, physical models reproduce the source sound production mechanism.

Spectral Modeling:

Fourier analysis - any periodic waveform (time domain) can be expressed as the sum of one or more harmonically related sine waves called its spectrum (frequency domain), each with a particular frequency, amplitude, and phase. The mathematical procedure for converting a waveform into its spectral components is called the Fourier Transform. The Inverse Fourier Transform converts a spectrum back to a waveform without any loss of information. The harmonic frequencies are called "bins". The amplitudes (or magnitudes) of each bin are known as the "real" part of the spectral signal; the phases are called the "imaginary" part, because imaginary numbers (that use the square root of -1) are required to compute phase.
Linear Predictive coding (LPC) can also be considered as a spectral modeling technique especially useful for voice analysis and synthesis. Originally developed for speech in the late 1960s and early 1970s, it is an example of technology transfer that tends to occur between computer music and the larger research groups in telecommunications and speech. In LPC the main resonances of a voiced sound are represented in terms of a digital filter. By removing such resonances, the so-called residual part remains. By modifying the parameters of the filters, it is possible to obtain interesting sonic variations.

Play “Idle Chatter” by Paul Lansky c.1985 (LCD200)

Play “Any Resemblance is Purely Coincidental” by Charles Dodge c.1980 (LCD1899)
Dodge used a technique of source separation on a 1907 recording of Leoncavallo’s aria ‘Ridi Pagliachi’ to separate Enrico Caruso’s voice from the instrumental accompaniment. Dodge manipulated the voice using LPC, creating new contours and chorus effects.

Physical Modeling:

Sound synthesis by physical modeling is a class of synthesis techniques in which the source sound production mechanism is mathematically simulated. As opposed to spectral models, physical models do not consider the way the sound is perceived by the ear, but how it is produced by a vibrating object.

An interesting aspect of physical modeling is the decomposition of a vibrating object into exciter and resonator. The exciter is intended as the source of energy imposed to the system, while the resonator is the object which produces sound. When the exciter and resonators are connected in a feedback loop – this is the case for self-sustained oscillators such as the violin, in which there is a continuous interaction between the bow and the string. In contrast, in percussion instruments the interaction between the player and the instrument is transient for each particular stroke, which means that the player interacts with the instrument for a finite amount of time, and then the instrument is left to resonate. The exciter-resonator approach is particularly interesting from a musical perspective, since unnatural exciters and resonators can be combined together, to create augmented virtual instruments.

Play “Silicon Valley Breakdown” by David Jaffe c.1982 (LCD125)