Additive Music Synthesis

The foundations of acoustic theory were described in a preceding chapter, "Audio Machines." This chapter introduces basic acoustic concepts such as *frequency*, *amplitude*, and *resonance*. These concepts provide the foundations for music synthesis.

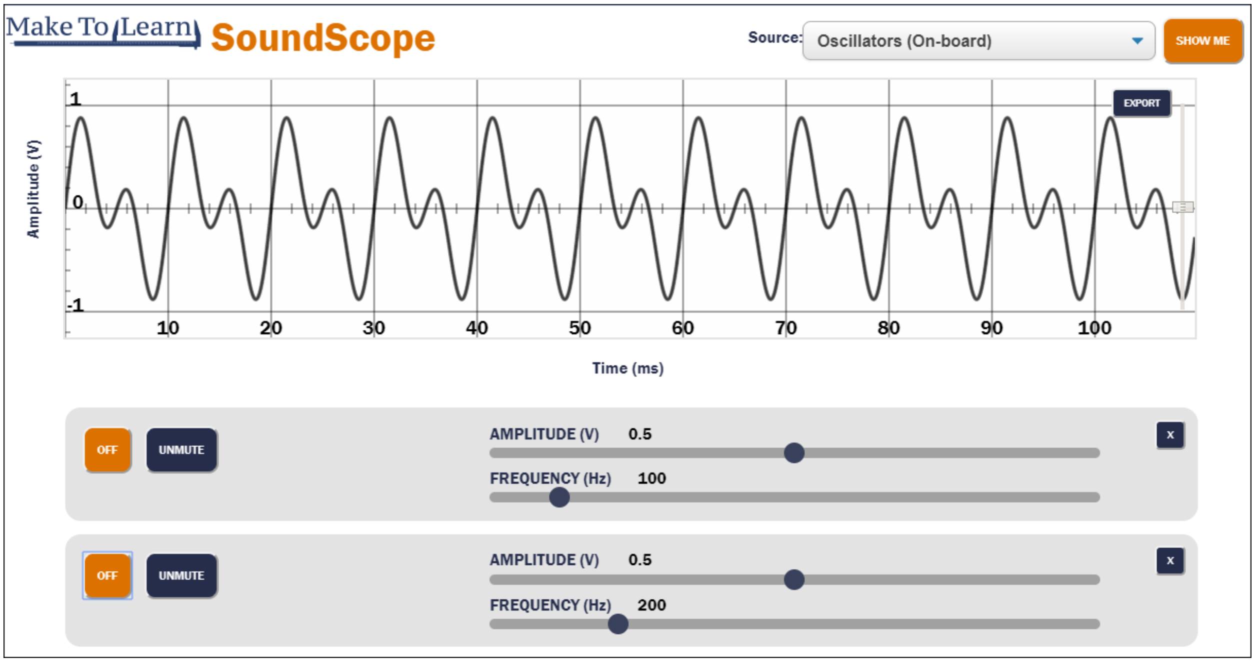
Subsequent chapters introduced methods for sequencing notes to create songs using the programming language, *Scratch*, and the open source music synthesis program, *Caustic*. In addition to sequencing notes to create songs, Caustic can also synthesize sounds.

A music synthesizer employs oscillators to reconstruct complex musical tones. Two or more tones can be combined to create a new tone. The process of mixing tones together to create new sounds is known as music synthesis.

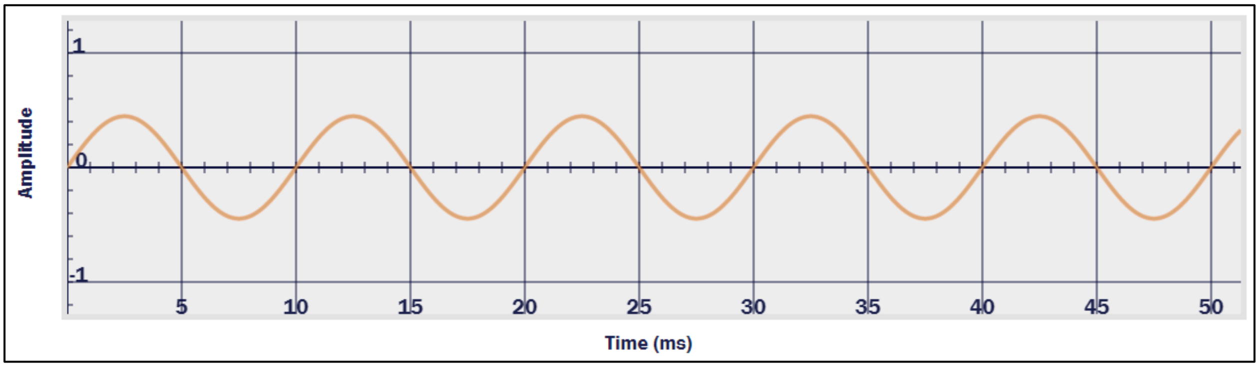
The *Make to Learn Tone Generator* can be used to generate tones and mix them together in this way. In this chapter, the *Make to Learn Tone Generator* will be used to demonstrate basic music synthesis concepts that will later be applied in Caustic. It is available at:

https://maketolearn.org/tools/soundscope/

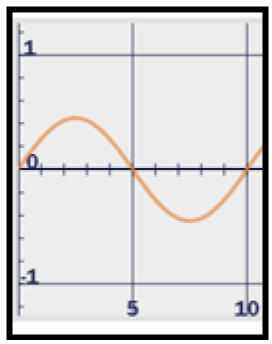
It can be used to mix tones together. The result of combining two more or tones can be heard. The resulting waveforms can also be observed on a visual display.



Begin by creating a 100 Hz tone. The graph of a pure tone that consists of only a single frequency produces a sine wave. In this instance, a cycle of the waveform is completed in 10 milliseconds. Since there are 1000 milliseconds in a second, 100 cycles of the tone will be completed in one second.



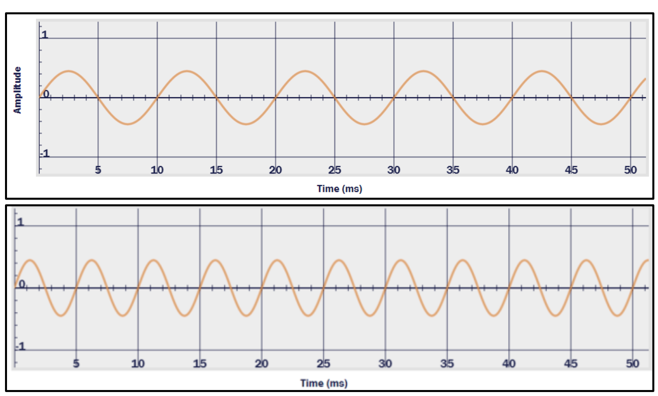
A single cycle of the sine wave represents one back-and-forth movement. When the tone is sent to a speaker, the positive peak will move the speaker cone in one direction and the negative peak will move the cone in the opposite direction.



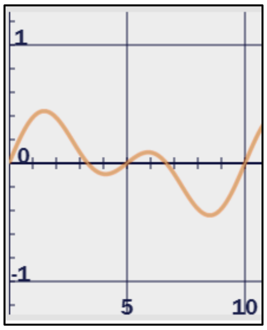
The number of cycles can be calculated by counting the peaks. In this instance, there are ten positive peaks in one-tenth of a second. This means that there were ten complete cycles of the waveform in that time period. If there are ten cycles in one-tenth of a second, multiplying by 10 tells us that there would be 100 complete cycles in one full second. Since this is a 100 Hz tone, the result obtained in this way matches our expectation.

Next we will compare the 100 Hz tone with a 200 Hz tone. Set the tone generator to 200 Hz.

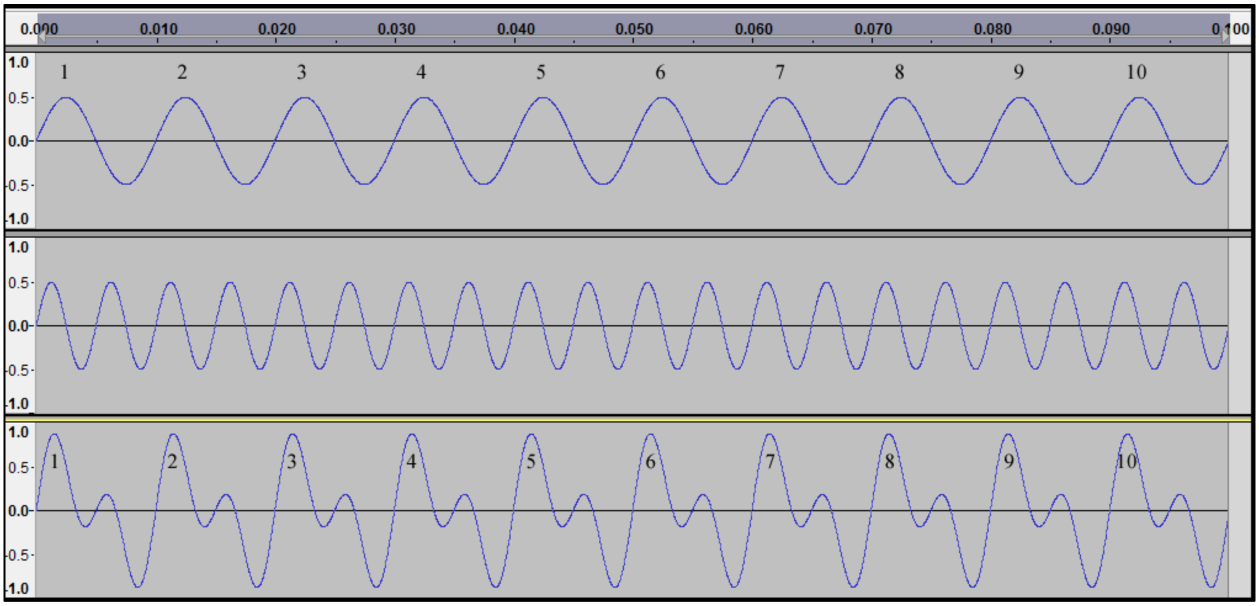
As we might expect, the 200 Hz tone has twice as many cycles in one-tenth of a second as the 100 Hz tone (20 cycles versus 10 cycles). The 200 Hz tone has a higher pitch.



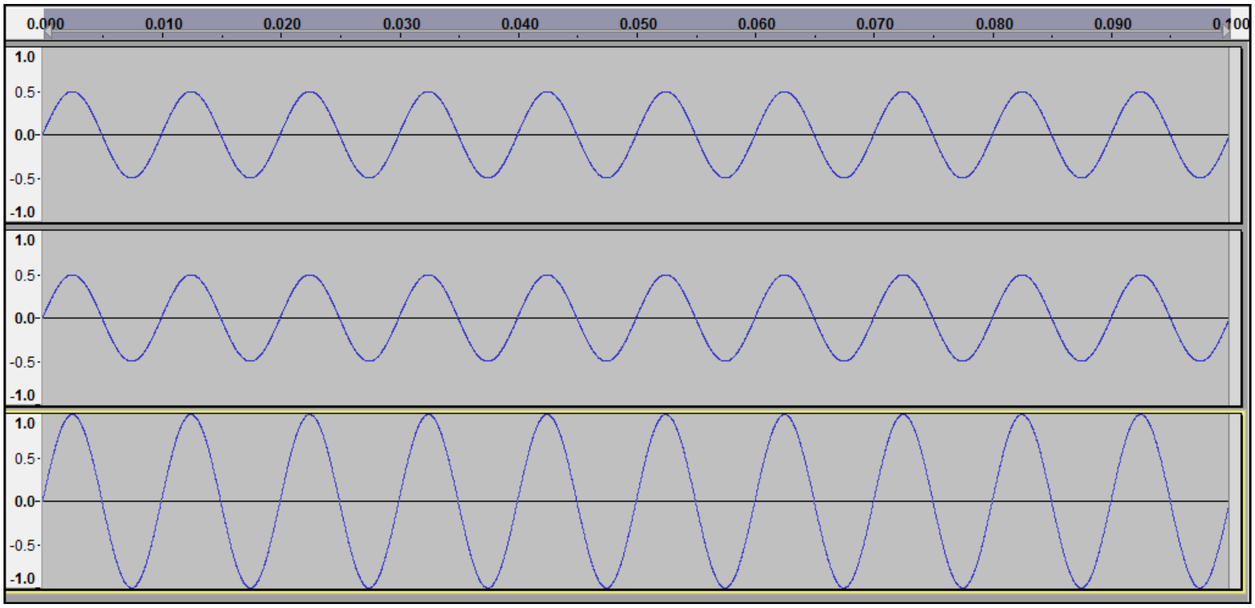
What do the two tones sound like when they are combined? Add a second oscillator. Set one oscillator to 100 Hz and the second oscillator to 200 Hz. A single cycle of the combined waveform looks like this. A complete cycle still takes the same amount of time (one-tenth of a second). However, the waveform is no longer a pure tone. Combining the 100 Hz tone with a 200 Hz tone results in a *complex wave form*. A complex wave form consists of two or more frequencies mixed together. Addition of the 200 Hz tone results in a double peak in a single cycle – one large peak and a second smaller one.



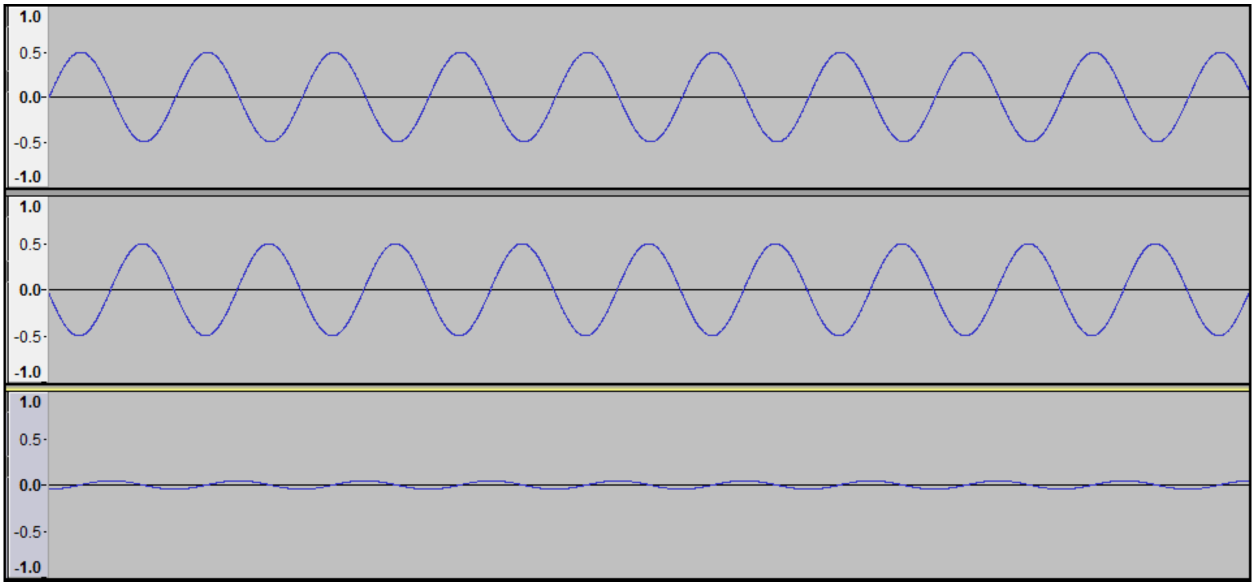
When counting complete cycles, only the highest peak is counted. Therefore, we can calculate that there are ten complete cycles of the combined waveform in one-tenth of a second. The complex waveform consists of a tone with a fundamental frequency (F0) of 100 Hz with a 200 Hz overtone.



What happens when two 100 Hz tones are combined? The values of each waveform are added together to obtain the combined waveform. Since the value of the first positive peak is 0.5, the value of the combined peaks is 1.0. Similarly, since the value of the first negative peak is – 0.5, the value of the combined peaks is – 1.0. The amplitude of the combined waveforms is double the amplitude of each individual waveform.



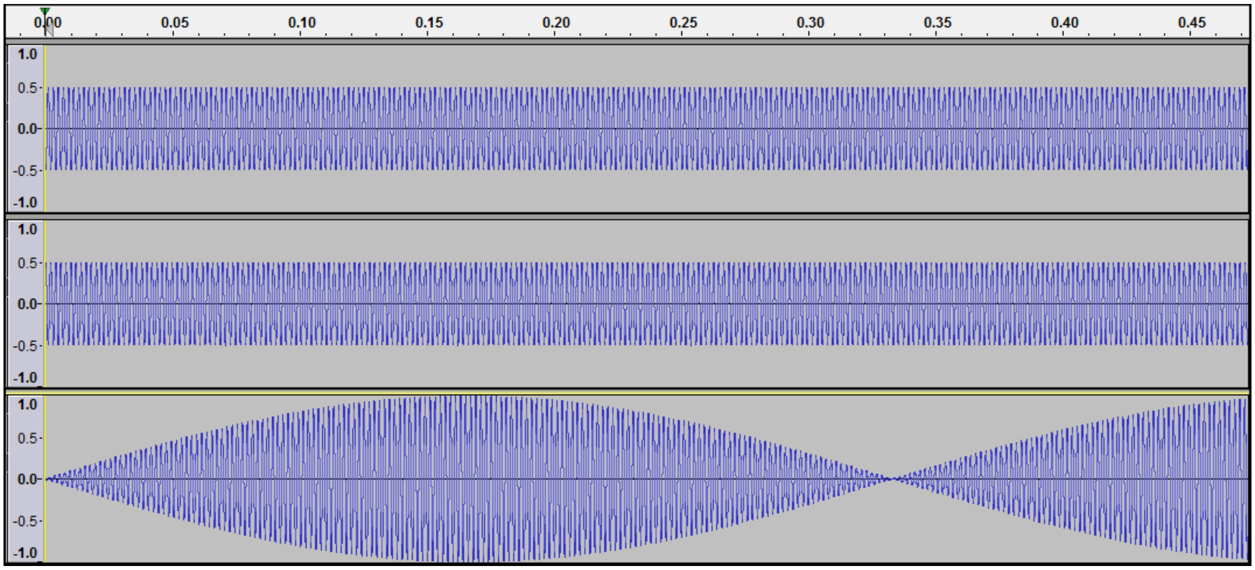
In the previous example, the positive peaks are aligned with one another. If a positive peak and a negative peak oppose one another instead, the positive peak and the negative peak will cancel one another out. The intensity of the combined tones is reduced rather than amplified. The point in the cycle at which a tone begins is known as its *phase*. In this case, the two tones are *out of phase*. This principle is used by sound-canceling headphones to achieve their effect.



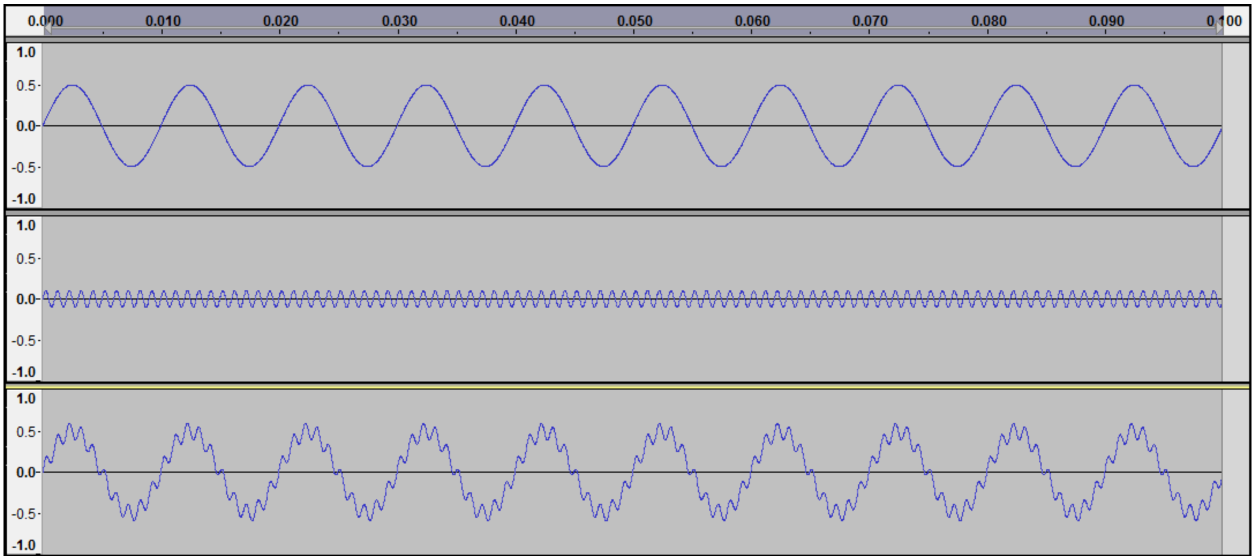
The *frequency*, *amplitude*, and *phase* of two tones that are combined are the factors that determine the quality of the sound that is produced. This type of music synthesis is known as *additive synthesis* because it is the result of adding the two tones together.

If two tones differ in frequency by a small amount, they will drift in and out of phase. For example, generate a 500 Hz tone that is 30 seconds long. Generate a second tone of the same duration that is 503

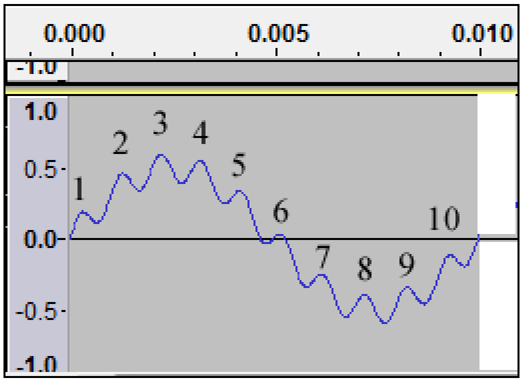
Hz. Then play the two tones together. The intensity of the combined tones will get louder and softer as they drift in and out of phase. This phenomenon is known as *beats*.



Beats are a special instance of one wave *modulating* a second wave. In another example, a 100 Hz tone with an amplitude of 0.5 is combined with a 1000 Hz tone with an amplitude of 0.1.



There are ten complete cycles of the 1000 Hz tone for every cycle of the 100 Hz tone. In this instance, the ten peaks of the 1000 Hz tone riding on top of a single cycle of the 100 Hz tone can be clearly seen.



A tone that affects a second tone in this way is said to modulate the second tone. In this case, the 1000 Hz tone is modulating the 100 Hz tone, changing its quality. The 100 Hz tone is said to be a *carrier wave* that carries the 1000 Hz tone riding on top of it.

Oscillator modules in a music synthesizer are configured in just this manner, with one oscillator generating a waveform that modulates a waveform generated by a second oscillator. The oscillators in the SubSynth module of Caustic are set up in this manner. The frequency of the second oscillator can be adjusted by amounts that vary by *octaves*, *semitones*, and *cents* (fine adjustment). An octave is the interval between one musical pitch and another note with double its frequency. A semitone is the interval between one musical note and the next in this scale. For example, C is adjacent to C♯; the interval between them is a semitone. One hundred cents is equal to one semitone. The cents control provides the capability for fine adjustments.

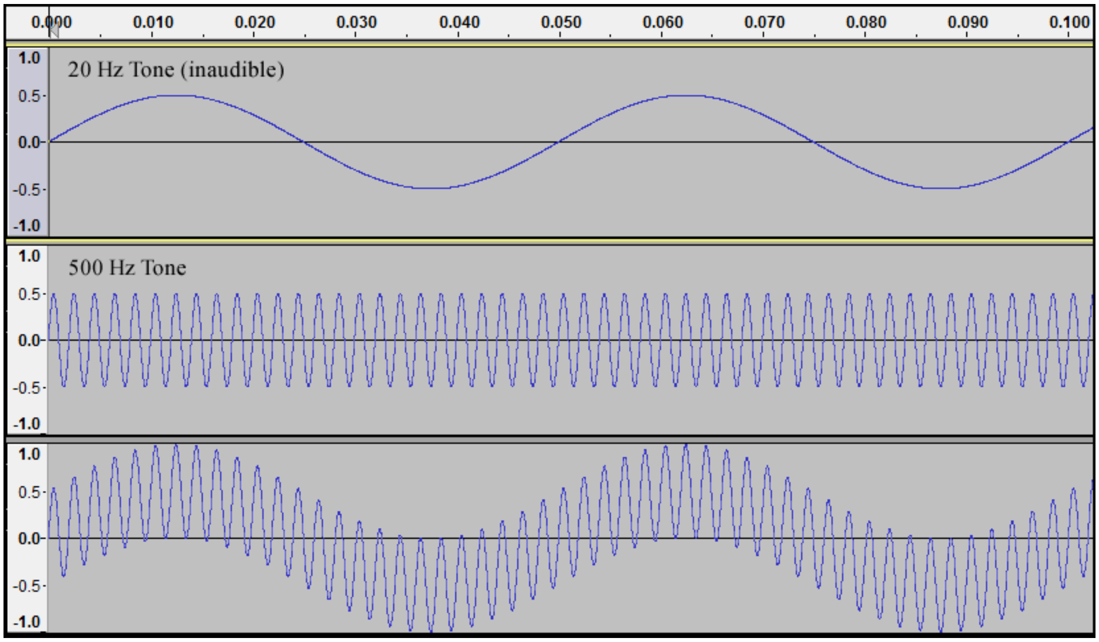
The *phase* control is the remaining adjustment available on the second oscillator. This adjustment allows the phase of the second oscillator to be adjusted with respect to the first oscillator.

The *Mix* control on the first oscillator controls the relative intensity of the first oscillator with respect to the intensity of the second oscillator. The *Bend* control bends the frequency upward or downward. Guitar players learned that they could increase the frequency at the beginning of a note by bending the guitar string. This technique was used to good effect in rock music and in other genres such as the blues. The Bend control on the oscillator reproduces this effect.

All of the examples discussed thus far involve *amplitude modulation*. That is, the amplitude of the second waveform is used to vary the amplitude of the first wave form. Other types of modulation such as frequency modulation are also possible. In this case, the frequency of the second waveform is used to modulate the frequency of the first waveform. A control in the first oscillator (currently set to Amplitude Modulation) can be used to adjust the type of modulation that is employed.

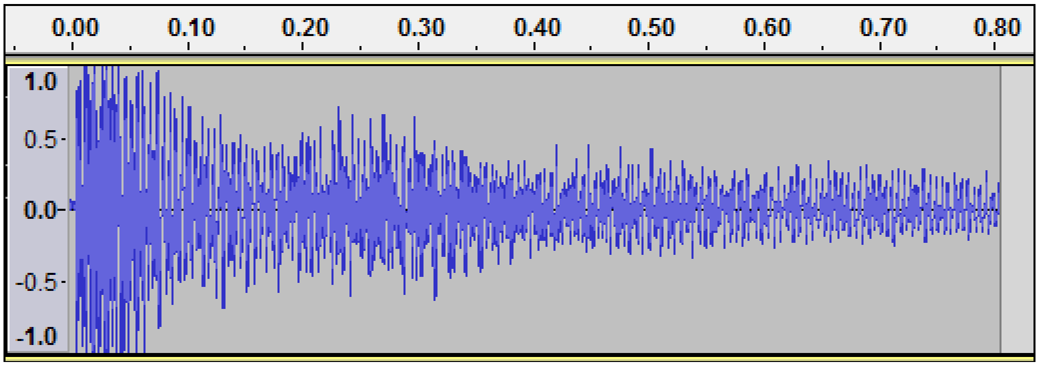
In addition to the two primary oscillators, the SubSynth panel also has two *Low Frequency Oscillators* (LFOs).

A waveform that is 20 Hz or below is not audible to the human ear. These frequencies are below the threshold of hearing. However, a low frequency tone can be used to modulate a higher frequency tone. In the example below, a 20 Hz tone has been created on the first track. When the 20 Hz tone is used to modulate a higher frequency tone (500 Hz in the example below), the result of the modulation can clearly be heard when the two tones are mixed together.

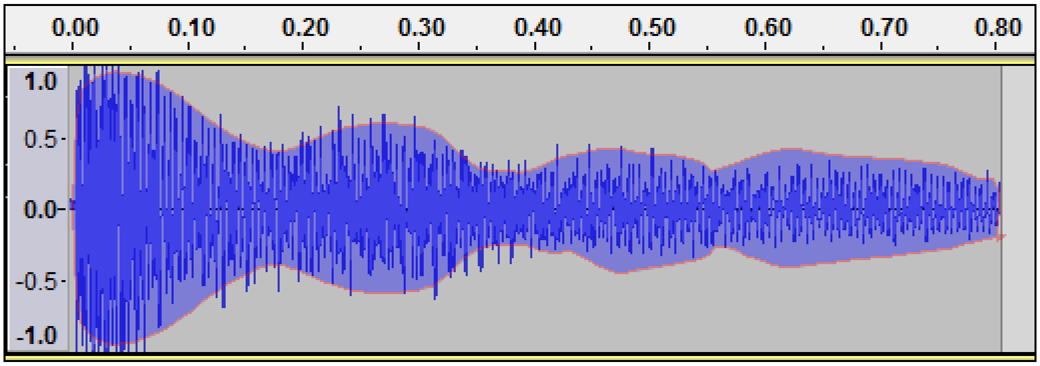


In the early days of hardware synthesizers, special electronics were needed to create low frequency oscillators. Therefore LFOs were installed as special modules that complemented the primary oscillators. Today's software can create a tone of any frequency. However, because of the manner in which music synthesizers evolved historically, LFOs are still included as separate modules in software synthesizers such as Caustic.

All of the examples thus far have involved steady state tones. When an actual guitar string is plucked, the sound does not continue forever. There is an initial spike at the beginning followed by a gradual diminution in sound until the string stops vibrating altogether.



The shape that results when a line is drawn across the tops of the peaks of the waveform is known as the *envelope* of the sound. The envelope of a steady state tone produced by an oscillator is a straight line. The envelope of a naturally occurring sound such as a plucked guitar string has several distinct phases. The moment of impact that occurs when the string is first plucked is known as the *Attack* phase. The gradual diminution in sound that follows is known as the *Decay* phase.



The *Attack* and *Decay* phases of a synthesized sound can be adjusted by *Volume Envelope* module of the SubSynth panel. In contrast to a natural sound, a synthesized sound can be sustained for an indefinite period of time. The *Sustain* adjustment controls how long the sound is sustained. Finally, the *Release* control is used to adjust the point at which the sound is released (i.e., terminated).

The convention of the use of *Attack, Decay, Sustain, Release* (ADSR) controls that shape the envelope of the waveform has now been adopted across most music synthesizers. It is common to see the abbreviation *ADSR* on the control panels of many brands of music synthesizers.