**Psychological and Physical Properties of Sound

by Hubert Howe**

Sound, and hence music, can be analyzed in two ways: physically, by using instruments to record measurements of its properties, and psychologically, by listening to the sound and ascertaining its properties on the basis of our immediate experience. Unfortunately, there is no one-to-one correlation between the physical and psychological attributes of sound. While measuring the physical properties is usually straightforward, the psychological properties are usually multi-dimensional, meaning that more than one physical property must be taken into account to describe each psychological property.

To this consideration, we must also add the fact that describing sound and music are not the same. Music puts sound into a context, in which our perception of its properties may be influenced by other events that happen in its near temporal vicinity. The same sound in different contexts may not be described in the same manner. Second, most musical sounds are truly complex, meaning that they have many different properties, most of which may be changing at any instant. Finally, we should note that ascertaining the aural properties of music is not the same thing as reacting to it emotionally, or evaluating it.

**Physical Properties of Sound**

1. Frequency: The period of a sound is the duration of one cycle of its motion. Its frequency is the number of cycles that occur within a second. Frequency is measured in Hertz (abbreviated Hz). The period and frequency of a sound are reciprocally related (i.e., period = 1/frequency).

The phase of a sound is the instantaneous amplitude at a given point in time. The entire cycle of a sound wave is divided into 360 equal parts called degrees. Phase is thus a measure of time with respect to frequency or period.

Human beings can perceive frequencies from about 20 Hz up to about 18,000 Hz (18 KHz), the limit varying with the individual (aging causes loss of high frequency detection).

2. Intensity: Sound intensity is a measurement of the amount of power of a sound at a given location. It is measured in decibels (abbreviated dB), which is a measurement of the ratio between the power of a given sound and a reference sound. The formula for calculating the intensity of a sound I is:

I = 10 log (I/I0)

where I0 is the reference sound. In most discussions of intensity, the reference point is taken as the threshold of hearing. The ratio of the threshold to itself is 1, and the intensity level is therefore 0 dB. Above this level, human beings can hear sounds up to a level of about 120 dB before the sound becomes physically painful.

3. Complex Sounds: Most sounds we hear, even those that have only a single perceived pitch, actually consist of several separate components whose amplitudes vary over the course of the duration of the sound. These frequency components may be divided into harmonic and nonharmonic (or inharmonic) partials. All frequency components of a sound constitute its spectrum. Modern studies of timbre further qualify the notion of spectrum to take into account a separate varying amplitude for each component.

Harmonic partials are tones whose frequencies are integral multiples of a fundamental frequency. Tones above the fundamental are called overtones. The fundamental frequency and its harmonic partials constitute a harmonic series or overtone series. Although the frequency between each successive overtone is the same, the musical interval or distance in pitch gets smaller as the series goes higher. It is a remarkable fact of sound perception that an entire collection of harmonic partials is perceived as a single pitch, with the collection itself being taken in as the timbre (see below).

A single tone with no overtones is a sine wave, which is the acoustical manifestation of simple harmonic motion. Each harmonic partial of a complex tone is a sine wave.

Nonharmonic partials are components of a sound whose frequencies are not harmonic partials of the fundamental frequency. Sounds containing nonharmonic partials do not possess a single pitch.

Noise is a sound containing a complex mix of all frequencies simultaneously, which is produced by random vibrations of air particles. White noise, so named by analogy with white light, has an even amplitude for all frequencies. Pink noise is like white noise but has a constant power per octave. (The term "noise" is also used to describe any "unwanted" sound.) Noise is always present in the background of any acoustical environment, which is why sound reproduction devices are measured by their signal-to-noise ratios, or the maximum intensity between the recorded signal and the background noise.

4. Envelope: Envelope is defined as the growth and decay characteristics of some property of sound; thus it has to be qualified by the property to which the envelope pertains. An amplitude envelope is the growth and decay of the amplitude of a sound, a spectral envelope is the growth and decay of its spectrum, etc. While envelope is recognized as a separate property because of the development of synthesizers and electronic music, it is usually not described as a property in books on acoustics.

5. Modulation: Modulation is defined as the periodic change of some property of sound. Thus, like envelope, it must be qualified by the property being modulated. The most common types of modulation are frequency modulation (abbreviated FM), amplitude modulation (abbreviated AM), timbre modulation, and location modulation. FM and AM are also known, respectively, as vibrato and tremolo.

Modulation always involves two signals: the carrier, or "original" signal before the modulation occurs, and the modulator (sometimes called the "program" signal), which changes the carrier signal. For clarity, it should also be noted that some books describe modulation as any type of change of a sound, thus including both periodic and aperiodic signals. It is more useful to refer to aperiodic change as random modulation.

There are always three characteristics involved in modulation: the speed or rate, the amount, and the shape of modulation. These characteristics are determined by the frequency, amplitude and waveshape of the modulator.

Modulation is the primary manner in which subsonic frequencies (i.e., tones below the lower threshold of frequency discrimination) occur in music. When modulation frequencies approach and exceed the lower frequency threshold, they begin to interact with the other audible frequencies in the tone, producing complex tones called sidebands. Control and manipulation of these sidebands is the manner in which FM synthesis (embodied in the Yamaha DX-7 and other instruments) occurs.

5. Reverberation: Reverberation is the cumulative effect that occurs when a sound is played in an acoustical space, where the sound that reaches the listener is a mixture of direct sound and sounds reflected off the walls, floor and ceiling, which arrive at the listener's ears at slightly delayed intervals. When a sound wave strikes a physical surface, some portion of the sounds is reflected away from the surface and some is absorbed by it. Researchers have determined that there are optimal reverberant characteristics for musical spaces, and this is therefore an important subject for architects and engineers.

While reverberation is associated with many different psychological properties, it is important to point out that it can effect the spectrum of the sound, since the reflected sound can resonate some of the partials of the tones.

**Psychological Properties of Sound**

1. Pitch: For many reasons, pitch is probably the most important characteristic in music. First, pitch can be broken into several distinct properties. Second, listeners are sensitive to the smallest changes in pitch. Finally, pitch organization is the primary topic of music theory.

There are at least three components of pitch that are basic to practically all music. The most basic is the "higher than" relationship, by which any tone can be described as higher, lower or the same as another. Second is the notion of octave equivalence, whereby tones an octave apart possess an "identity" with each other not shared with other pitches. Finally, two or more pitches sounding together create an "interval" or "chord" which has an additional similarity not shared by other intervals or chords.

Of all psychological and physical properties of sound, pitch is the one that practically has a one-to-one correlation with the physical property of frequency. The exception concerns only very low frequencies, which may sound flat. Whereas frequency is measured in Hz, pitch is measured by the identification of tones in the equally-tempered scale, or other musical scales.

The subject of tuning systems is very important for most music, but today we rarely discuss it since we assume equal temperament for most music. This is probably an error. The 12-tone equal-tempered scale took centuries of music history to evolve. Pianos are tuned even today in something more approaching meantone temperament, and organs are often tuned by historical methods. Live music is constantly adjusted by ear. Another topic for further investigation is the subject of intonation and pitch deviations, about which little research has been done.

2. Loudness: As pitch corresponds to frequency, loudness corresponds to intensity; but in this case, psychoacoustical research has shown that loudness is a function of both intensity and frequency. The accepted explanation for this is that the human ear itself possesses a resonance at about 3,000 Hz. Since this is rather high in musical terms, tones must be boosted progressively as they move lower to produce the perception of equal loudnesses.

3. Timbre: Timbre (pronounced "tam-ber") is defined in the literature as the property that enables a listener to identify the instrument playing the sound. It is also described as the "tone quality", the psychological property corresponding to the spectrum of a sound.

The traditional definition both helps and confuses the issue. It helps, because it identifies timbre as multi-dimensional (many different properties help identify the instrument, not just the spectrum); but it confuses because there are no easy generalizations about the similarities between one tone on an instrument and another. Especially when we accept the spectrum of a sound as multiple components, each with a separate envelope, we see that sounds are too complex for this property to be a single item. Also, in electronic music, identifying the "instrument" is not a purposeful activity, since this is often just something like a synthesizer.

There are two basic similarities by which the timbres of different pitches may be compared. The most direct is similarity of waveshape. Two sounds that have the same spectrum would thus be equal according to this definition, regardless of their pitches. This is the property that exists in common between all members of the clarinet family, since the shape of the instrument resonates only odd-numbered partials.

More important, however, is comparing tones on the basis of their formants. A formant is a fixed frequency area in which the loudest partials occur, regardless of the partial numbers. (For example, a formant at 1000 Hz would resonate the tenth partial of a 100 Hz tone, but the fourth partial of a 250 Hz tone, and the 25th partial of a 40 Hz tone.) Formants are produced by filters that resonate a particular part of the frequency continuum. It is on the basis of formant similarities that listeners identify vowels in speech, and this is probably our most innate concept of timbre. (Most vowels have three or more different formants, but usually one is most prominent.)

4. Sound Location: The location from which a sound emanates is a property always present in any sound. This is an obvious fact, and it is usually not structured in live music. In electronic music multitrack playback systems allow composers to make use of it.

5. Envelope: Envelope is a psychological property as well as a physical property, but very few studies have been made of perceptual envelopes. Musical terminology is full of terms that describe different types of articulations, such as legato, pizzicato, staccato, sforzando, etc. There are many perceptible gradations between the shortest pizzicato-like percussive tone to the legato-like organ tone, and electronic music practitioners can experiment with different values to determine what may be useful.

6. Properties associated with delays: The reverberant characteristics of musical spaces give rise to a number of identifiable properties that can be perceived by listeners. Effects devices allow the delay times to be manipulated in many ways, producing effects that go beyond what is possible in acoustic spaces.

Delay devices (which are usually digital, so these are often called "digital delays") allow a sound to be delayed for a specific duration and then mixed in with the original sound, with volume controls on both signals. Delays are usually measured in milliseconds (abbreviated ms), and are variable from about 1 to 500 or more ms (half a second). A variety of effects exist in a continuum from very short to very long delays. There are two "magic" numbers that help separate these effects. At about 40 ms, the delayed signal begins to become distinct, and at about 100 ms, it can be heard as a separate tone.

Thickening occurs below 40 ms, as the sound begins to increase in "fullness". At about 40 ms, the effect is described more as doubling, where two distinct voices can be heard. Between 40 and 50 ms the delayed sound starts to break away from the original sound and be perceived as an echo. Above 100 ms the effect is described as "slapback" echo, where the original signal has a distinct "answer" in the reflection. A continuum of different effects can be perceived from the shortest to the longest delay between these values.

Chorus effect is defined as the properties that exist when a tone is played on two or more instruments, which produces something different from merely intensifying the amplitude. It is produced by deviations in intonation, envelope, modulation, and rhythm, since two human instruments will never play precisely together. Many effects devices include a type of chorus effect produced by delaying the original sound by the "doubling" amount of time.

Effects devices also contain a number settings that are probably not distinct perceptual properties, but which are similar to those mentioned above. Reverberation itself, usually produced by a complex mixture of short delays, is both a coloring and "smearing" property, since it has filtering effects on the sound and a variety of the thickening-doubling-echo properties. Different reverberation settings allow experimentation with these values. Flanging is created by modulating the delay time, and produces a kind of "whooshing" effect.

7. Other Properties: There are many additional properties of sound that can be identified, and some that have been researched. All that is needed to determine whether some property exists is, first, that someone formulate an idea that something exists, and then to test the idea by listening for it in music. As time goes on, undoubtedly new ideas will be postulated and tested.

**Basic Synthesis Controls**

The designers of electronic music synthesizers have generally attempted to control the properties of sound through methods that access the psychological and physical properties of sound described above, rather than duplicating the designs of acoustic instruments. On virtually any synthesizer, a wide range of sounds is made possible by controlling the different devices over all their values.

1. Oscillators (sometimes abbreviated "VCO", for "voltage-controlled oscillator"): The tones produced by synthesizers are generated by oscillators (periodic waveshape generators), noise sources, or sampled instrumental tones. The basic signals are usually unvarying, the changes being produced by other devices in the patch. Harmonic generators, which produce each partial of a tone separately, are achieved by using groups of oscillators tuned individually to the partials.

Pitch on synthesizers is controlled by key selection, detuning controls, and the master tuning adjustment. Most keyboards are equal tempered, but some allow individual tuning of keys to produce pure, historical, non-Western or microtonal scales. Pitch is modified by LFOs (for vibrato), pitch envelope generators, the pitch bender, and in special cases by key pressure or controllers like the modulation wheel.

2. Amplifiers (VCAs): An amplifier is not an envelope generator but a means of applying a control signal to the amplitude of a sound. Amplitude is affected by: (1) velocity, usually with a sensitivity adjustment, (2) the volume slider, and (3) pedals (but note that the sustain pedal affects duration, not amplitude). Other methods that may be used for individual sounds include the breath controller or key pressure.

3. Envelope Generators (EGs): An envelope generator produces a control signal which can be used as an envelope. They usually work in conjunction with the keyboard, as follows: when the key is depressed, the attack begins, and the envelope varies in accordance with preset values until the key is released, when the final decay of the sound begins (note that the sound may last longer than the key remains depressed). Envelopes always apply to amplitude, and they are sometimes used for pitch, filtering, and panning (location) as well. Historically, envelopes can be divided into the following four categories:

Attack-Release (AR) envelopes are the simplest, which consist simply of an attack, sustaining period, and decay. All systems can produce these with a minimal setting of values. Certain percussive and plucked sounds can have an even simpler envelope, namely one that consists only of an attack and decay with no sustaining period.

ADSR envelopes ("ADSR" stands for "attack-decay-sustain-release") allow an "accent" at the beginning of the sound, by allowing the attack to reach a higher value than the sustaining level, with the initial decay beginning immediately after the attack and the final decay after the key release. ADSR envelopes have now generally been replaced by:

Multi-position contour generators: These allow the four (or more) points that exist in an envelope to be set to any values, as long as they begin and end on zero, so that crescendos and diminuendos can occur at any point in the envelope. While these are definitely more flexible than an ADSR envelope, they are still used pretty much the same way.

Non-amplitude envelopes are similar to contours, but they don't have to begin and end at zero, since other properties (such as pitch or filtering) may need to be controlled in opposite ways. (If an amplitude envelope doesn't begin or end at zero, the sound will "bleed through".)

The attack times of envelopes are usually much shorter than decay times; this is both a fact of instrumental sounds as well as an indication that the same values used for attacks and decays will not produce the same durations. Since the release time begins after the key is released, too-long decays can cause voice stealing as well as a "smearing" of the sounds. EGs can produce crescendo-diminuendo as well as articulative properties; the disadvantage is that the total EG is usually applied to all notes played with the specific patch. Sometimes (but rarely) EG values can be affected by attack or release key velocities.

4. Filters (VCFs): Filtering is basic to analog synthesis, but is often not found on FM or phase modulation synthesizers (these have other methods of producing dynamically varying timbres). While high pass, low pass and band pass filters can all be found on synthesizers, most now use a type of low pass filter than can be converted to a band pass filter by means of a "resonance" function, which emphasizes the resonance at the cutoff frequency. If a filter is truly "VC" (voltage controlled), then it can be varied by an applied signal (such as an EG, key pressure or a controller) to produce dynamically varying sounds such as "wa-wa" and "yeow". A filter "tracking" control causes the frequency response to follow the keyboard, sometimes in a non-linear manner.

5. Low Frequency Oscillator (LFO): An LFO produces a low frequency signal that can be used to control vibrato, tremolo, and sometimes filtering. The rates usually vary from 0 Hz up to above 20 Hz, and sometimes more. (Also, some synthesizers allow an audio oscillator to go down to these rates and be used for these purposes.)

LFOs usually have four controls: the speed (frequency), amount (amplitude) for pitch (vibrato) or amplitude (tremolo) of the sound, the waveshape (usually sine, triangle, sawtooth and square waves are possible, sometimes more), and a delay before the introduction of the LFO (from zero to 1 or more seconds). Even though both vibrato and tremolo can both be affected by the same signal, this is rarely desirable.

Sometimes the amplitudes of the LFO can be controlled by a real-time device such as the modulation wheel, a pedal, a slider, or a breath controller, allowing the performer to "bring in" the vibrato or tremolo selectively. These usually just control the amplitude and not the other aspects of the LFO, and the actions of the performer do not affect the LFO other than to vary the amplitude of the applied signal.