Frequency and Pitch

Musical signals often have a periodic character, that is: we have some wave-shape that is periodically repeated. The interval of repetition (denoted as T, expressed in seconds) is known as the period of the signal and its reciprocal value is called the frequency:

$$f = \frac{1}{T} \tag{1}$$

and is expressed in units of Hertz, Hz - where $1Hz$ means one period per second. Sometimes we also find the unit *cps* which stands for 'cycles per second' and is therefore synonymous with Hz. Typical examples of such periodic signals are the waveforms in synthesizers such as the sawtooth-, pulse-, triangle- and sine-wave. The most basic periodic signal is the sine-wave. Each periodic waveform can be thought of as a sum of scaled and time-shifted sine-waves, that is, we can construct every imaginable waveform by summing weighted and time-shifted sine waves (we can construct an arbitrary waveform from sums of other waveforms as well, but there are some deeper mathematical and perceptual reasons, why the sine was chosen). So let's assume henceforth, that we are talking about sine-waves. We will also use the term pure tone or simply tone for a sine-wave with a given frequency.

Human perception of pitch

Humans are able to perceive sine-waves with frequencies in the range from $20Hz$ to $20000Hz = 20kHz$ (with large individual deviations from that depending on age, exposure to loud signals and other factors). **Like the perception of loudness, the perception of musical pitch is also a logarithmic one** (that's actually a simplification, but we'll stick with that). We perceive the difference between $100Hz$ and $200Hz$ as the same as the difference between $1000Hz$ and $2000Hz$ for example. What really counts, is not the absolute difference but the ratio between the two frequencies. This is the same in both cases: $\frac{2000}{1000} = \frac{200}{100} = 2$. **A frequency ratio of 2 is called an octave** and we say that a tone of $2000Hz$ is an octave above the tone of $1000Hz$, or conversely the tone of $1000Hz$ is an octave below the tone of $2000Hz$.

Subdivision of a logarithmic scale

Now let's see what's necessary to subdivide the octave into smaller intervals, say thirds of octaves. We want a rule with which we can start at some given frequency f_0 (say $f_0 = 100Hz$, for example), apply that rule to f_0 to arrive at f_1 which is a third of an octave higher. If we apply the same rule to f_1, we want to arrive at a frequency f_2 which is again a third of an octave higher than f_1 . And finally, if we apply the same rule to f_2 we want to arrive at f_3 which should once again be a third of an octave higher than f_2 and exactly an octave above f_0. If we define our rule as 'add 33.333...', it would work for $f_0 = 100Hz$ but grossly fail for $f_0 = 1000Hz$. We rather need a multiplicative rule, which multiplies our point of departure f_0 by some constant c and we want a 3-fold iteration of the rule to result in a multiplication by 2. If we iterate a multiplication by some number c three times (each multiplication acts on the result of the previous one), we have - in effect - multiplied our original frequency f_0 by c^3. We want $c^3 = 2$ (three iterations should bring us an octave higher), so we choose $c = \sqrt[3]{2}$. Thus a frequency ratio of $c = \sqrt[3]{2} = 1.25992...$ represents a third of an octave. For graphical equalizers it is often customary to have either octave spaced bands or bands that are spaced by thirds of octaves. If we would have halved the octave, we would have arrived at a frequency-ratio of $c = \sqrt{2} = 1.41421....$

The equal tempered scale and MIDI

To make music, we need to define a set of tuning-frequencies which we want to use. Such a set of frequencies often called a scale. In western music, the most popular scale is the equal tempered scale which is based on starting at some reference-frequency as anchor (chosen to be at $f = 440Hz$), and use frequencies which are at octaves and subdivisions of octaves away from that anchor-frequency. This anchor frequency has been assigned the musical pitch name A4 and the octaves come out as $A1 = 55Hz, A2 = 110Hz, A3 = 220Hz, A4 = 440Hz, A5 = 880Hz, ect.$. **To subdivide the octave, the number 12 was chosen**. Intervals of that subdivision interval are called semitones and a frequency which is a factor $c = \sqrt[12]{2}$ times some other frequency f_0 is said to be a semitone above f_0. Note: in musical terms, there exists the interval of a major and minor third and the term 'major third' happens to coincide with a third of an octave as described above - however this is pure coincidence. To see this, consider the musical fifth: it is halfway between a frequency and its octave (therefore a greater fraction of the octave than a third) in musical terms, but is not a fifth subdivision of an octave (which is a smaller fraction of the octave than a third). In german, we have distinct terms for these different things :-P. The representation of pitch in the MIDI-standard (**M**usical **I**nstrument **D**igital **I**nterface) is directly based on the equal tempered scale and defines 128 (numbered from 0 to 127) possible values for note-pitches. The frequency $440Hz$ is represented in MIDI by the note-number 69, and each whole number codes one semitone, such that the MIDI-representation of $A3$ (which is an octave below $A4$) is $69 - 12 = 57$. In general, if we denote the pitch in terms of a MIDI note number as p and the frequency in Hz as before with f, the formulas for converting back and forth between pitch and frequency become:

$$f = 440 \cdot 2^{\frac{p-69}{12}} \qquad \Leftrightarrow \qquad p = 12 \cdot \log_2(f/440) + 69 \tag{2}$$

and these formulas work not only for the discrete set of frequencies which is representable by the integer MIDI note values, but for any.