

GDC Audio Boot Camp

Glossary of Terms

ADPCM	Adaptive Differential Pulse Code Modulation. A <i>compression</i> algorithm that stores the approximate difference between subsequent <i>samples</i> , typically to represent 16-bit audio data in 4-bits (~4:1 compression).
Amplitude	The magnitude at any given time for a <i>waveform</i> . For audio, typically representative of a sound wave's power, in terms of voltage, <i>decibels [dB]</i> , or <i>sound pressure level [SPL]</i>
Artifact	Human-perceived distortion or inaccuracy in audio signal reproduction typically created through limitations of the recording, storage, or playback processes.
Attenuation	The reduction of an audio signal's <i>amplitude</i> from its authored power level. Equivalent to negative <i>gain</i> . (E.g. a gain of -6 dB is equivalent to 6 dB of attenuation)
Batch processing	Processing a number of files with the same functions/parameters. E.g. adding a "radio effect" to all comms dialogue.
Cent	One hundredth of a <i>semitone</i> , or 1/1200 of an <i>octave</i> .
Clipping	An <i>artifact</i> produced when the recording or playback device is incapable of capturing or creating the peak audio power present in a signal. In the most dramatic scenario, triangle waves are 'clipped' into square waves. Most commonly indicated by multiple sequential <i>full scale</i> audio <i>samples</i> .
Codec	Coder-Decoder. A self-contained module implementing one or more <i>compression</i> algorithms, as well as the ability to decode and playback audio data in these compressed formats.
Compression	1. The process of representing an audio signal with less data than was originally sampled.

2. The process of adjusting the attenuation/gain of a waveform to reduce its *dynamic range*, often to amplify quiet waveforms or reduce the volume of loud ones.

Cross-fade	Fade between two different sound sources. The most common types are linear, logarithmic, and equal power (describing the volume curves used to fade between the two sounds).
Decibel	[dB] A measurement of the ratio of two audio power levels. 1 decibel is the smallest typically human-perceived volume change. $\text{dB} = 20 * \log_{10} \left(\frac{V_1}{V_2} \right)$ $\text{dB} = 20 * \log_{10} \left(\frac{2}{1} \right) \approx 6.02\text{dB}$ - So a sound with twice the power of another sound is roughly 6.02 dB louder.
Decibels full scale	[dBFS] A sound's power relative to a <i>full scale</i> signal. 0 dBFS is a sound representing the <i>full scale</i> level a device can reproduce.
Digital Signal Processing	[DSP] The manipulation of audio data <i>samples</i> in the digital realm.
Doppler	The perceived change in a sound emitter's frequency based on the movement of the emitter and/or the listener.
Dynamic loading	To load/replace resident sound effects "on the fly" - that is, while some sound effects are playing, load other sound effects into memory (perhaps over existing effects). The converse, static loading, has a single 'bank' of sound effects loaded at the beginning of a game or level, which remains resident and unchanging.
Dynamic Range	The difference between the <i>peak</i> audio level and the "noise floor" (e.g. the quietest <i>amplitude</i> an audio signal generates). Humans have a dynamic range of roughly 120-130 dB.

EQ	Equalization. The use of multiple <i>filters</i> at various <i>frequencies</i> to allow the adjustment of signal strength across the spectrum of audible (or even inaudible) frequencies.
FFT	Fast Fourier Transform – a mathematical process by which sound can be converted from the <i>time domain</i> to the <i>frequency domain</i> .
Frequency	The measurement of a repeating signal's duration, measured in cycles per second or <i>Hertz</i>
Frequency domain	The presentation of a sound in terms of its frequency components over a given period of time, in comparison to the <i>time domain</i> (traditional audio <i>waveforms</i>). Some kinds of sound analysis and signal processing are easier to perform in the frequency domain.
Filter, band pass	A filter that passes all frequencies between a low-frequency cut-off point and a high-frequency cut-off point
Filter, high pass	A filter that passes all frequencies above a given cut-off frequency
Filter, low pass	A filter that passes all frequencies below a given cut-off frequency
Full scale	The highest <i>peak</i> level a device is capable of reproducing. The presence of multiple sequential full scale <i>samples</i> often indicates <i>clipping</i> .
Gain	The amplification (positive) or reduction (negative) of an audio signal's <i>amplitude</i> beyond its authored power level. Negative gain is equivalent to <i>attenuation</i> . (E.g. a gain of -6 dB is equivalent to 6 dB of attenuation)
Hertz	[Hz] The unit of <i>frequency</i> for a waveform. 1 Hertz is 1 cycle per second.
Localization	Adapting a product to a specific location or territory. For audio, most typically involves translating text and dialogue.

MP3	MPEG-1 audio layer 3. A <i>compression</i> algorithm using human perceptual characteristics. Rate and artifacting can vary based on content and algorithm aggressiveness, but are typically significantly better than ADPCM.
Nyquist Limit	One half of the <i>sampling rate</i> ; this is the highest <i>frequency</i> that can be accurately stored and recreated.
Octave	A doubling or halving of a given <i>frequency</i> . For instance, if the note A is used as a reference pitch at 440 Hz, then 55, 110, 220 will be A's in lower octaves, and 880 Hz, 1760 Hz, etc. will be A's in higher octaves.
Pan	The positioning of a sound between several speakers; between the left and right speakers in a stereo environment, or between the four, five, or seven directional channels in a multichannel environment.
Peak	The maximum <i>amplitude</i> (typically in decibels) of a given <i>waveform</i> .
Perceptual codec	A <i>compression</i> algorithm that takes advantage of human hearing characteristics (for instance, sensitivities to certain frequencies) to attempt to minimize audible <i>artifacts</i> . MP3, WMA, and Ogg Vorbis are all examples of perceptual codecs.
Pitch	A subjective term for the perceived frequency of a tone.
Pitch Shifting	Adjusting the <i>pitch</i> of a playing sound, either to simulate <i>Doppler</i> , or for another effect (such as basic variation). Typically performed via <i>sampling rate</i> conversion when working in the <i>time domain</i> .
Quantization	The process of limiting an infinite range of values to a discrete set of values (based on the smallest measurable 'quantum'). In audio <i>waveform</i> sampling, this occurs in both the <i>frequency domain</i> (<i>sampling rate</i>) and the <i>amplitude domain</i> (the number of bits used to store values, for instance 8 or 16 bits). Quantization may introduce <i>artifacts</i> .

Resident Sound Effect	A sound that is stored in RAM and can be accessed instantly. Also known as “in-memory”.
Rolloff	How a sound emitter’s volume is <i>attenuated</i> with distance.
Root Mean Square	[RMS] A method for measuring perceived loudness of a sound, based on the square root of the average of the squares of a sound’s <i>amplitude</i> over time.

$$\text{RMS} = \sqrt{\frac{\sum_{t=1}^n (\text{Amplitude}(t))^2}{n}}$$

Sample	The measurement of the instantaneous <i>amplitude</i> of a <i>waveform</i> at a given point in time. The process by which analog audio signals are converted to digital form.
Sampling Rate	The <i>frequency</i> with which <i>samples</i> are captured or reproduced by a system, typically measured in <i>Hertz</i> [Hz] or kilohertz [kHz]. For instance, 44.1 kHz means 44,100 samples per second.
Semitone	One twelfth of an <i>octave</i> . In equal temperament systems, equivalent to a “half step”.
Seek-time	The time it takes for the physical disk head (CD, DVD, hard disk, etc.) to find the beginning of a file on disk and begin to read it. Also known as seek “latency”.
Sound Pressure Level	[SPL] A measure of perceived <i>volume</i> based on the force of disturbance a sound causes in a medium, typically in <i>decibels</i> . 0 dB SPL is the average human hearing threshold; 120-130 dB SPL is typically the pain threshold.
Spot Effect	A one-off sound effects that are triggered and just play once until a similar event causes them to be re-triggered. E.g. gun-shots, explosions, ball kicks, menu sounds, etc. Also known as “One-shot” or “Fire and Forget” effects.

Stream	A sound file that is stored on disk (not in memory) and played by seeking for data, which is loaded in small segments into memory.
Time Domain	The presentation of sound data relative to the passage of time (in comparison to the <i>frequency domain</i> , where sounds are presented in terms of their frequency data over a given period of time.
Time Invariant	Processing on a sound that does not impact the speed at which it is played back. In the <i>time domain</i> , basic pitch shifting is typically time-dependent (e.g. pitch shifting a wave up causes it to play faster). In the <i>frequency domain</i> , pitch shifting can often be performed without altering playback time.
Volume	Human perception of <i>sound pressure level</i> , influenced by <i>frequency</i> and <i>amplitude</i> among other factors.
Waveform	Representation of a sound wave in terms of <i>amplitude</i> vs. time
WMA	Windows Media Audio. A <i>compression</i> algorithm using human perceptual characteristics. Rate and artifacting can vary based on content and algorithm aggressiveness, but are typically significantly better than ADPCM.