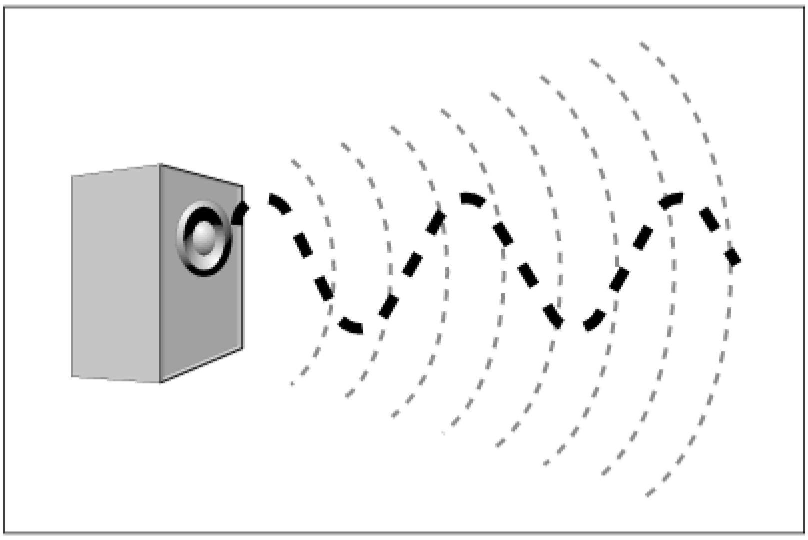
AUDIO FILES

Introduction

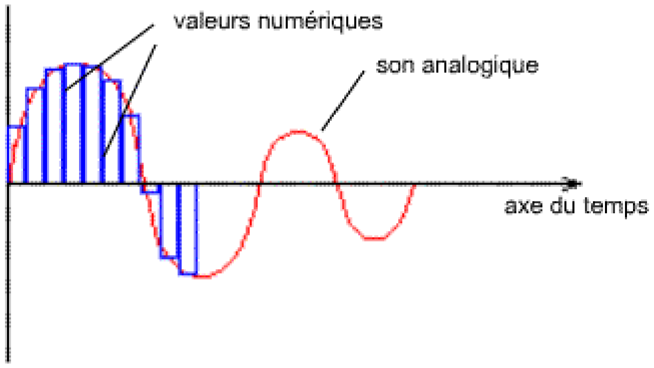
Sound is the sum of the vibrations produced by vocal chords, a loudspeaker, etc…



These vibrations have a frequency, which is measured in Hertz. The human ear acts as a receiver and perceives only the frequencies ranging from 20Hz – 20kHz.

Digital sound

Sound files contain sound in digital format. Real sound is characterized by its amplitude, which varies according to time, and is referred to as analog.
To represent analog sound in digital form, the amplitude of the sound signal is recorded at regular time intervals and a numeric value is assigned to each recording. This is known as **sampling**.



The value of each sample is stored in a file, and the original signal (for example the recorded music) can then be reconstituted based on these values. The higher the number of samples recorded per second, the better the quality of the reconstituted signal.

For example, a sampling frequency of 44100 Hz means that 44100 samples are “taken” per second. The sampling frequency can reach up to 88.2, 96, 176.4, or 192 kHz. This advance in technology allows us to produce sound with more dynamics and more detailed transitories, and therefore get much closer to functioning like the human ear.

Digital sound quality also depends a lot on the precision of the sample values. They can be encoded using 8, 16, 24 or even 32 bits. 8 bit encoding enables the representation of 255 entirely different values (from -127 to +127), and 16 bit-encoding enables values which are a lot more precise (from -32167 to +32167). This is referred to as **sound resolution**.

Sound may also be recorded in Mono or Stereo mode.
It is of course in Stereo mode that sound is best reproduced, but it requires twice as many values.

The more precise the sampling configuration is, the bigger the size of the file containing the sample values will be.

The sound quality represented in the file depends on the sampling frequency, the precision of the sample coding process and whether recording was done in Mono or Stereo mode.

The size of the file also depends on these values.
For example, we calculated that 20 seconds of sound sampled at 44100 Hz, in stereo mode, using 16-bit encoding, corresponds to 3.52 Mo on a hard disk. This sampling configuration is known as “CD quality”.

The coding is very bulky: 1 min of Audio CD = 8,5Mo!
The solution to this is compression!

Sound compression

In order to compress sound, several methods are possible:
- Eliminating the high frequencies which are almost inaudible.
- Removing vibratory noise
- Reducing the sampling frequency (but it is worth noting that this also seriously reduces the resulting sound quality!)
• The solution: VBR (Variable Bit Rate) where the sampling frequency adapts to the sound.

According to the sampling frequency, the sound resolution and whether it is mono or stereo mode, we obtain a rate which is measured in Kbit/s, and represents the degree of sound quality. In VBR, the rate changes constantly.

Typical values
• A CD: 1440 Kbit/s
• An MP3 file with minimum compression: 320 Kbit/s
• Acceptable quality in MP3 format: 128 or 192 Kbit/s
• An MP3 file with maximum compression: 64 Kbit/s
• Telephone sound quality: 32 Kbit/s

AUDIO FORMATS

There are about 50 different audio formats in the world, which are more or less widely used.

For a given format the files exist in several compression rates or data formats with different sampling frequencies, resulting in very different levels of sound quality and file size. Generally, for a given codec, the lower the compression rate, the bulkier the file, but the closer the compressed file gets to the non-compressed file in terms of sound quality. By contrast, the higher the rate of compression, the more the sound loses in quality (this is only relevant for codecs with losses).

The most frequent audio file extensions

WAV: Waveform Audio Format

This is a “basic” audio format developed by Microsoft. The coding and decoding are immediate since there is no compression. Sound may be Mono or Stereo.
WAV files offer unmatchable sound quality and are compatible with all audio players but file size is very large. They are not suitable for distribution over the Internet.

CDA: Audio CD

It is not a directly readable file as such but a pointer towards a track on the CD which is in WAV format. Software is required to extract the WAV file.

FLAC: Free Lossless Audio Codec

FLAC is a free codec for lossless audio data compression. Contrary to codecs such as MP3 or Vorbis, it leaves the audio source totally intact.
FLAC is appropriate for all types of audio data storage and supports metadata, cover art, as well as fast search. FLAC is free and open source; there are no royalties to pay and it is well supported by other software.

ALAC: Apple Lossless Audio Codec

ALAC is a lossless coding format which was released quite late in 2004. The format was developed by Apple in order to close a loophole in the MPEG-4 standard, as the section dedicated to lossless encoding has still not been finalized. ALAC is a proprietary format which mainly concerns iTunes and iPod users.

AIFF: Audio Interchange File Format

AIFF is a digital audio file format developed by Apple for storing sounds on their brand computers.
AIFF files have the file name extensions .aif, or .aiff.
The data is encoded in non-compressed PCM big-endian. Thus, a CD Audio track, coded in 16 bits, stereo 44.1 kHz will have a bitrate of 1411.2 Kbit/s. There is also a compressed variant of the format, known as AIFF-C or AIFC. It can achieve a compression rate of 1:6.

MP3: MPEG1 audio layer 3

The MPEG-1/2 Audio Layer 3, more commonly referred to as MP3, is the audio specification of the MPEG-1 standard, from the Moving Picture Experts Group (MPEG). It is an algorithm used for audio compression (see codec) and is capable of drastically reducing the amount of data necessary in order to reproduce the audio message, but for the listener it would appear to be a non-compressed reproduction of the original audio source. In other words, there is a significant loss in terms of sound quality but it is acceptable to the human ear.
At the moment, this is the most widely used format and is compatible with almost all existing software.

Real audio

The RealAudio format is very old and allows distribution of music on the internet using the streaming technique. This proprietary audio format was developed by the company RealNetworks and is readable with RealPlayer, and it can be created with RealProducer.

WMA: Windows Media Audio

WMA is a proprietary audio compression format developed by Microsoft. The WMA format offers the possibility of protecting the output files against illegal copying directly at the encoding stage, using the Digital Rights Management technique (DRM).

The format exists in four forms :
- *WMA Standard*: it was the first version to be released and the most widely used on the Internet. It is the only one (at the present time) to be readable on many portable digital music players.
- WMA Pro: in theory, it offers better quality but is a lot less widespread.
- WMA Lossless: it offers identical sound quality to the original source and is a serious rival to other lossless formats such as FLAC.
- WMA Voice: is dedicated to low bit-rate voice encoding (below or equal to 20 Kbit/s).

WMA is an alternative to MP3, is more flexible but less widespread and is only compatible with Microsoft software.

Ogg: Ogg Vorbis

Vorbis is a digital audio compression and decompression algorithm (codec), is not patented, is open and free to use and is more effective in terms of quality and compression rates than the MP3 format, but is less popular than it. It remains compatible with few applications at the present time.

AAC: Advanced Audio Coding

AAC is an audio compression algorithm with data information loss aiming at offering a better quality/binary rate ratio than the older MP3 format. For this reason, t was chosen by different companies such as Apple or RealNetworks.
It is a direct rival format to WMA, mainly used for iPod and iTunes.

Format	**Size**	**Quality**
AAC	3,99 Mo	128 kbps
Vorbis	3,75 Mo	128 kbps
WMA	3,9 Mo	128 kbps
MP3	5,2 Mo	128 kbps
AIFF	43,5 Mo	1411 kbps (CD)
Apple Lossless	25,4 Mo	841 kbps
FLAC	26 Mo	CD
WAV	42,50 Mo	CD

Table comparing the size and quality of the main audio formats used