Action Plan Background: WAVE

Author: Andrea Goethals
Last Revision Date: 03/29/04

Preface:

The WAVE format complies with the Resource Interchange File Format (RIFF), which is Microsoft's variation on the Interchange File Format (IFF) from Electronic Arts. Like IFF files, RIFF files are composed of a FORM 'chunk' that contains other 'sub-chunks'.

The WAVE format has never been documented in an authoritative stand-alone specification, and therefore has no format 'versions'.

1 General Description

1.1 Format Name: Waveform Audio File Format (WAVE)

1.2 Version: N/A (See *Preface*)

1.3 MIME media type name: audio (See *section 1.4*)

1.4 MIME subtype: There is not an IANA-registered MIME type for WAVE files. Some applications use audio/x-wav, some use audio/wav, and some use both. Obeying the MIME rules, an "x-" prefix for the MIME subtype should be used for formats without registered MIME types, therefore audio/x-wav should be used for WAV files.

1.5 Short Description: a general-purpose audio format

1.6 Common Extensions: wav

1.7 Color depth: N/A

1.8 Color Space: N/A

1.9 Compression: WAVE files can store their sound data in PCM format (equivalent to no compression for digital audio) or in any one of a number of compressed formats.

1.10 Progressive Display: N/A

1.11 Animation: N/A

1.12 Magic number(s): WAVE files are big-endian in some parts of the file, little-endian in other parts. All integers in the file are stored little-endian, everything else is big-endian. The header starts with "RIFF" (0x52, 0x49, 0x46, 0x46), followed by 4 little-endian bytes that tell

how many more bytes follow in the file. After this 4-byte number is “WAVE” (0x57, 0x41, 0x56, 0x45).

1.13 Specification Requirements

- Must have the header described in *Section 1.12 Magic Number(s)*
- A Format chunk is required and must occur before the required Data chunk. All other chunks can occur in any order.
- A Fact chink is required if the waveform data is compressed (non-PCM) or if the data is contained in a wave-list
- Requirements of the optional Cue-Point chunk:
    - each cue point must have a unique name (in the dwName field)
- Each chunk must have an even number of bytes

2 Essential and Distinguishing Characteristics

WAVE is a common general-purpose audio format, similar to the AIFF-C format as it allows for either PCM format or a compressed data format. The data can technically have from 1 to 65,535 channels, although the specification only maps the speaker positions for 1-2 channels[1]. They can technically have an essentially unlimited sampling rate and average bytes per second playback. Of course, applications themselves will only support a certain number of channels, sampling rate, etc. and human hearing has its limits as well. For example, the human hearing range is roughly 20 Hz to 22 kHz. Sampling theorem has shown that a sampling rate of 44 kHz can be adequate for a 22 kHz audio frequency [Pohlmann 2000]. Most digital audio players expect certain values, for example 8.0 kHz, 11.025 kHz, 22.05 kHz, or 44.1 kHz for sampling rate.

The WAVE format supports labeling positions ('cue-points') in the waveform data, associating textual information with these cue-points, specifying the play order of these cue-points, and embedding other textual files to further describe a cue-point directly in the WAVE file. *Table 1* shows the technical metadata that can be extracted from a WAVE file.

1 A more recent Microsoft document describing the WAVEFORMATEXTENSIBLE format (see *Section 4.1 Specification and File Variations*) made the speaker mappings more explicit. An extra field was added to the Format chunk that can be used to map channels to 18 speaker positions. [Microsoft 2003b] [Microsoft 2001]

Technical metadata Element	*Obligation (R = required by spec., S = information given by spec., O = optional but described in spec., X = described in publication external to spec.)*
Number of channels	R
Sample size (bits per sample point)	R for PCM Format
Sampling rate aka Samples per second (Number of sample points that should be played back per second)	R
Average Bytes per second (to play it) For PCM Format: (= Block align * sampling rate)	R
Bytes per sample (aka Block Alignment) (number of Bytes per sample point * number of channels) The number of bytes may include padding that isn't part of the sample.	R
Valid bits per Sample (actual number of bits containing the sound data per sample without any padding)	X [Microsoft 2003b]
Channel Mapping (1-2 channels)	S
Channel Mapping (more than 2 channels)	X [Microsoft 2003b]
Samples per block (number of samples contained in one compressed block)	X [Microsoft 2003b]

Table 1: Technical metadata that can be found in a WAVE file. The WAVE specification sometimes uses the word 'sample' in a way that is equivalent to AIFF's 'sample point' and sometimes uses it in a way that's equivalent to AIFF's 'sample frame'. For clarity, AIFF's terms are used in this table.

3 Usefulness

3.1 Version Duration: The WAVE format was documented in the first version of [Microsoft 1991], which was issued in November, 1990, which makes it 13 years, 4 months old.

3.2 History of Prior Versions Duration: There have never been any 'versions' of the WAVE format but Microsoft has changed the WAVE specification over time (see *Section 4.1: Specification and File Variations*).

3.3 Expected Newer Versions: None expected.

3.4 Existence of Publicly Available Complete Specifications:

The earliest specification of the WAVE format was part of a larger document [Microsoft 1991], and was vague about WAVE format requirements. That document is still available from third party websites, but not Microsoft's website. The WAVE format has since been documented in a limited way on web pages on the *MSDN Library* website [Microsoft 2003a], [Microsoft 2003b]. Web pages do not have the same level of permanence as versioned, published

specifications. In addition, these web pages state at the top of each page "This is preliminary documentation and subject to change."

There is a lot of non-authoritative, incomplete and sometimes factually wrong information about the WAVE format on the Internet. It is not uncommon for the WAVE documentation that shows up on many websites to have an unspecified author and date. It is also common to read WAVE format information on the Internet that mentions that other WAVE format documentation was incorrect. In one widely-published document the formulas for computing required WAVE file fields were wrong. [Sonic Spot] notes that WAVE readers have to expect wrong values for these fields along with other common WAVE file anomalies. These errors can be contributed in part to a lack of authoritative documentation on the WAVE format especially during times when many WAVE file writers and readers were being written.

3.5 Specifications-controlling Body: Microsoft Corporation

3.6 Related Legal Issues: There are no known legal problems specific to the WAVE format.

3.7 Application and Platform Support: The WAVE format, using PCM or a few other common compression algorithms, is widely supported by most digital audio players on all major operating systems.

3.8 Limitations: Like the AIFF/AIFC format, the flexibility of this format as far as the sampling rate, bit depth etc. that it can support can be a disadvantage. It is possible to create technically valid WAVE files that no player can interpret. This may be why Microsoft introduced the newer more restrictive format - see *Section 4.1 Specification and File Variations*.

3.9 Perceived Popularity: The WAVE format is still a very popular format, especially as a high-resolution format that can be converted to a smaller compressed data format (MP3, etc.) for transfer across a network or to a portable device.

4 Related Formats

4.1 Specification and File Variations:

WAVEFORMATEX [Microsoft 1994]
[Microsoft 1994] describes the 'general extended wave format' structure that should be used to describe all non-PCM format WAVE files instead of the structure described in [Microsoft 1991]. It added one extra field to the Format chunk (See *Figure 1*).

WAVEFORMATEXTENSIBLE [Microsoft 2003b]
Microsoft realized that the WAVE format was "too under-specified to be used for high bit-depth samples or multiple channel streams" [Microsoft 2001]. They decided that the original WAVE format and the extended wave format (described above) should be made obsolete. This newer

format does not have a different version number but can be detected by its having extra fields in the Format chunk (See *Figure 1*). This format puts restrictions on existing fields as well as adding more fields. For example, the BitsPerSample value must be a multiple of 8 in this format. The previously vague elements of the WAVE format are better defined in this document. The problem is that there is so much existing documentation on the WAVE format, as well as existing WAVE file writers that have already been created using the old documentation. One result is that WAVE file readers have to expect many variations on the WAVE format.

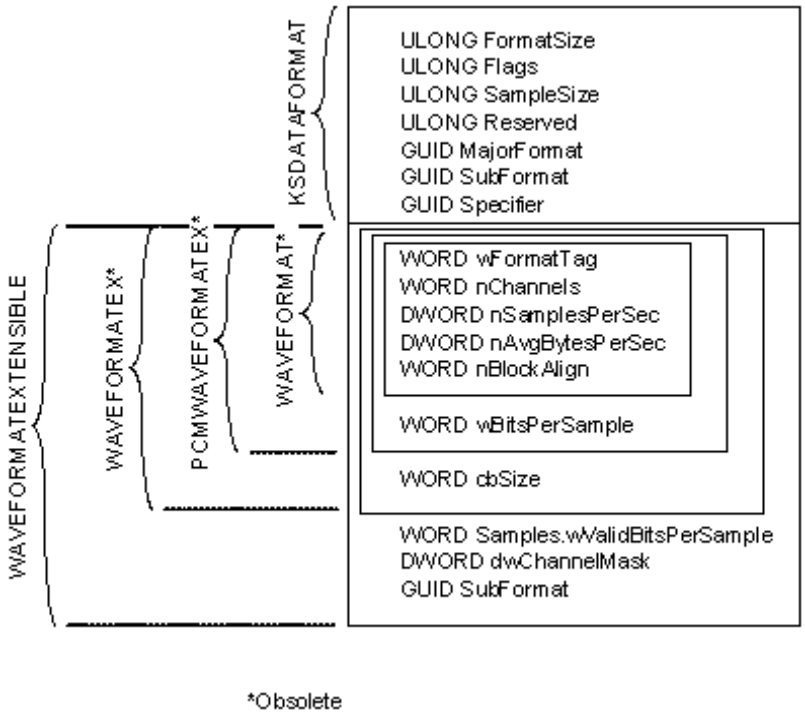

Figure 1: The WAVEFORMATEXTENSIBLE is a superset of the "obsolete" formats: WAVEFORMATEX, PCMWAVEFORMATEX, and WAVEFORMAT. The original specification [Microsoft 1991] describes what is here referred to as PCMWAVEFORMATEX and WAVEFORMAT. [Microsoft 1994] describes what is here referred to as WAVEFORMATEX. This figure appears in [Microsoft 2003c] .

RIFX, Microsoft Corporation [Microsoft 1991]
A RIFX file is the same as a RIFF file except that the first four bytes are "RIFX" instead of "RIFF", and a RIFX uses big-endian format for all of its byte ordering.

WAVE Files have a field in the Format chunk that indicates the WAVE 'format category' of the file which generally means the compression algorithm used on the sound data, if any. [SDL_Site 2001] and [Microsoft 1994] list 38 different WAVE format types. Each of these types can have different fields in the Format chunk.

Canonical WAVE Format [Weber b]
Before the Win32 API was published by Microsoft, a description of the WAVE format was circulated and used to write WAVE file creators and players. This format is simpler than Microsoft's WAVE format. It only has 2 chunks: the Format and the Data chunks, in that order.

The data is in PCM format, always. This format is compatible with players that can play Microsoft's WAVE format, but players built specifically for the Canonical WAVE format would have trouble with Microsoft's WAVE format.

BWF Format (http://www.ebu.ch/departments/technical/pmc/pmc_bwf.html)
The BWF (Broadcast Wave Format), developed by the European Broadcasting Union, is based on the WAVE format. While a WAVE file can use any one of a number of compression formats, BWF files can only use either PCM format or MPEG compression. A BWF file has at least one extra chunk that a WAVE file does not - the Broadcast Audio Extension chunk. This chunk contains material that broadcasters would exchange with each other, like a textual description of the sound file. The EBU came up with a core metadata set for radio archives that is based on the Dublin Core metadata. This metadata can be stored in the <axml> chunk that was added to the BWF specification in 2003. A BWF is compliant with the WAVE format and uses the WAVE file extension (wav) so WAVE players can play BWF files (but can not parse any added metadata).

5 Summary and Conclusions

The WAVE format is an old but still very popular digital audio file format, especially for PCM sound data. Files in the WAVE file format can be somewhat varied in structure, in part due to the conflicting WAVE format information on the Internet.

The WAVE format is being used as a 'master' audio format for preservation purposes by some archives and archive projects, while others are using the AIFF format. Both of these files support the PCM format which is what is of importance to the archives. As [Fleischman 1998] notes:

> *... it is the "PCM-ness" of the file that is important, not the "WAVE-ness." This idea can be compared to our use of TIFF files for image masters, where it is the uncompressed bit-mapping that is more important than the "TIFF-ness" of the file.*

The BWF format (see *Section 4.1 Specification and File Variations*) might be an attractive master format for archives. It can use the PCM format and supports the inclusion of a core metadata set in accordance with the Dublin Core. It is not as well-known as WAVE and AIFF, but it may become more popular, at least with archives.

Other possible successors to the WAVE format are discussed in the FCLA AIFF/AIFC Action Plan Background Report. These include lossless compressed data formats and PCM alternatives. All of these technologies should be monitored until a clear successor to the WAVE format emerges.

6 References

[Bourke 1996] Bourke, Paul. “WAVE sound file format”, Web page, November 2001.

http://astronomy.swin.edu.au/~pbourke/dataformats/wave/

[Fleischman 1998] Fleischman, Eric. "WAVE and AVI Codec Registries", RFC: 2361, Microsoft Corporation, June 1998.
ftp://ftp.isi.edu/in-notes/rfc2361.txt

[Fleischhauer 2003] Fleischhauer, Carl. "Audio and Video Preservation Reformatting: A Library of Congress Perspective", Paper delivered at the Preservation Conference: Digital Technology vs. Analog Technology, National Archives at College Park, March 27, 2003.
http://www.archives.gov/preservation/conferences/papers_2003/fleischauer.html

[Kabal 2003] Kabal, Peter, "Wave File Format Specifications", Web page, McGill University, August 5, 2003.
http://www.tsp.ece.mcgill.ca/MMSP/Documents/AudioFormats/WAVE/WAVE.html

[Microsoft 1991] Microsoft Coproration. "Multimedia Programming Interface and Data Specification 1.0", Issued as a joint design by IBM Corporation and Microsoft Corporation, August 1991.

[Microsoft 1994] Microsoft Coproration. "Microsoft Multimedia Standards Update: New Multimedia Data Types and Data Techniques", Revision 3.0, April 15, 1994.

[Microsoft 2001] Microsoft Corporation. "Multiple Channel Audio Data and WAVE Files", Web page on the MSDN website, December 4, 2001.

[Microsoft 2003a] Microsoft Corporation. "WAVEFORMATEX", Web page on the MSDN Library website, June 2003.
http://msdn.microsoft.com/library/default.asp?url=/library/en-us/stream/hh/stream/aud-prop_5s4z.asp

[Microsoft 2003b] Microsoft Corporation. "WAVEFORMATEXTENSIBLE", Web page on the MSDN Library website, June 2003.
http://msdn.microsoft.com/library/default.asp?url=/library/en-us/stream/hh/stream/aud-prop_7wz7.asp

[Microsoft 2003c] Microsoft Corporation. "Data Formats for Wave Audio Streams", Web page on the MSDN Library website, June 2003.
http://msdn.microsoft.com/library/default.asp?url=/library/en-us/stream/hh/stream/aud-design_23l3.asp

[Palacios 2001] Palacios, Jesus Olivan. "wav files", Section in *An Introduction to the treatment of neurophysiological signal using Scilab,* version 0.02, May 2001.
http://www.neurotraces.com/scilab/scilab2/node24.html

[Pohlmann 2000] Pohlmann, Ken C. *Principles of Digital Audio*, Fourth Edition, McGraw-Hill, 2000.

[SDL_Site 2001] Author unknown, "New WAVE Types", SDL_Site documentation, icculus.org Web site, October 2001.
http://www.icculus.org/SDL_sound/downloads/external_documentation/wavecomp.htm

[Sonic Spot] Sonic Spot. "Wave Files", Sonic Spot Web site, date unknown.
http://www.sonicspot.com/guide/wavefiles.html

[Weber a] Weber. Timothy John. "The WAVE File Format: Answers to Common Questions", Web page, date unknown.
http://www.lightlink.com/tjweber/StripWav/WAVE.html

[Weber b] Weber. Timothy John. "The Canonical WAVE File Format", Web page, date unknown.
http://www.lightlink.com/tjweber/StripWav/Canon.html