

INTRODUCING STEM, A NEW MULTI-CHANNEL AUDIO FORMAT

Mickael Le Goff, Chad Carrier, Stewart Walker

Native Instruments GmbH

{mickael.legoff, chad.carrier, stewart.walker}@native-instruments.de

ABSTRACT

We present STEM, an open standard for stem content in the form of a single Stem file. Motivated by the need of DJs to manipulate individual sound sources independently, the Stem file contains 4 stereo channels containing for example drums, vocals, bass and melodies, along with a mixed-down version of the original music. The file has backward compatibility with 2-channel players and a JSON format is used to attach additional metadata. This document outlines the structure of the Stem File, an open format which can be adopted freely in any music software, for playback or creation.

1. INTRODUCTION

In the age of modern digital DJing, it is a common practice to mix more than two stereo tracks simultaneously. With access to multi-track recordings, DJs can create new types of mixes and manipulate the different sound sources independently instead of using rough three band filtering. However, no simple solution that embeds multi-track audio into one single file had a commercial success so far, leading to a small amount of multi-track content available on the market. Native Instruments recently created the STEM file format [1], an open standard for multichannel audio data which has been adopted by numerous music labels and online music stores.

Similar to the Mpeg-A IMAF file format [2], Stem files use the .mp4 container format to store the four individual stems of a track within a single file. Using the file extension .stem.mp4, this single file can be managed just like an mp3 file. A standard master version of the track is also included in the file and can even be played in stereo with any compatible audio player following standard mp4 specs, like iTunes for example. To play with a track's individual stems, Stem-compatible software or hardware is required. However, as an open audio file format, Stems can be implemented into software or hardware of any kind. Developers can access the specifications and libraries required for integrating Stems into their music production and performance tools, or media players.

The Stem files contain audio usually encoded using the AAC format in order to be free from any licensing fees. If an optimal sound quality is required, Stem can also be encoded with the lossless audio compression format ALAC at the cost of an increased file size.

In order to ensure that the audio quality of the stem playback matches the original stereo master as closely as possible, a DSP chain composed of a compressor and a limiter needs to be added to the playback engine. The settings of the DSP effects can be modified within the Stem Creator when creating the Stem file and are stored as metadata in the file itself. The DSP effect is provided as a pre-compiled library which can be included in your own projects.

As a single file, the Stem files facilitate tasks such as file management, distribution, and sale of multichannel audio data. By providing more value for the end user, Stems can create additional revenue streams for the music producers and labels. It would also give the possibility of online music stores to grow their business. Several online stores and more than a hundred labels already create professional Stem content and close to a thousand files are already available. A free Stem Creator tool has also been developed to give the possibility to easily create your own Stem file. We expect that additional royalty-free content will be accessible on online sharing platforms.

This new format could make multi-track database easier to handle for the MIR community. Ground truth annotation can be embedded directly into the file as JSON, making it easier to share between researchers. It could also help improving audio analysis by running MIR algorithms on the relevant sound source (e.g. tempo detection on the drum channel, key on the melodic channel) or could be useful in the field of source separation for direct comparison between results and original source.

2. FILE STRUCTURE

The structure of the Stem file follows the specification for MP4 containers given by ISO/IEC 14496-12 [3] and ISO/IEC 14496-14 [4]. The Stem file is formed as a hierarchical structure of objects called boxes (sometimes called atoms). The box type is given using a 4-byte character code. The Stem file format mostly relies upon al-



© Mickael Le Goff, Chad Carrier, Stewart Walker.

Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** Mickael Le Goff, Chad Carrier, Stewart Walker. "Introducing STEM, a new multi-channel audio format", Extended Abstracts for the Late-Breaking Demo Session of the 16th International Society for Music Information Retrieval Conference, 2015.

ready existing box types and introduces only one new box called stem, which is marked blue in the following diagram.

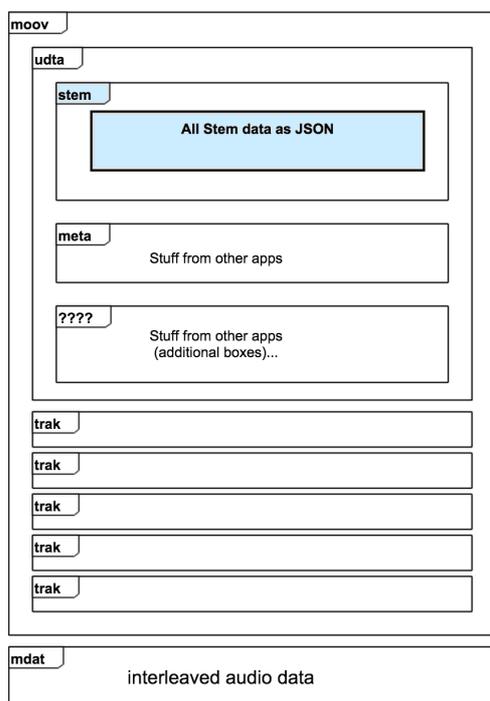


Figure 1. Detailed STEM file format structure.

The outermost box is the file type box `ftyp` followed by the movie metadata box `moov`. The stem metadata is stored in a `stem` box inside the user data box `udta`. While it would make sense to place the box outside of the user data, placing it in `udta` makes it more accessible through existing APIs such as the AVFoundation on iOS. This compatibility to existing APIs is vital, as it allows Stem files to be read and played which are stored in the iTunes Library on an iOS device, to which developers have only limited access through said APIs.

3. STEM CREATION

To facilitate the process of creating Stem files, Native Instruments released a simple standalone application called Stem Creator. The Figure 2 illustrates the interface for the application. Users can import five stereo tracks, add labels and export them in a `stem.mp4` file. The Stem Master Dynamics DSP has also been added to the Stem Creator, allowing the user to adjust the compression and limiting settings with a real time preview and add the parameters values to the Stem file. If adding more information is needed, such as extra metadata into the JSON file, a command line tool ‘`ni-stem`’ will be available soon. This tool gives access to more options for file creation such as the file path to a specific JSON file that will be embedded into the Stem file.

4. MASTERING STAGE

When an artist creates a track, it typically goes through a final mastering stage where an engineer applies additional EQ, compression, and other techniques to make the music sound balanced, controlled, and often with maximum perceived volume. This mastering process is the final stage of the music creation, but Stem files allow a user to change the audio mix of a track which happens before the mastering stage. Therefore, the normal process of mastering will not work for Stem files and part of it needs to be replicated in real-time during playback.

This is the purpose of the Stem Master Dynamics DSP which is a high-quality stereo Compressor followed by a Limiter. The settings of these dynamic processors can be adjusted in the Stem Creator and are embedded into the Stem file. If the Stem player implemented the Stem Master Dynamics in the signal chain, the parameters can be read from the Stem File and applied to the DSP for optimal audio playback which responds correctly to any changes made to the stem mix by the user.



Figure 2. Stem Creator interface.

5. REFERENCES

- [1] <http://www.stems-music.com/>
- [2] I. Jang, P. Kudumakis, M. Sandler, K. Kang: “The MPEG Interactive Music Application Format Standard”, *IEEE Signal Processing Magazine*, pp. 150–154, Vol.28, Issue 1, Jan 2011.
- [3] ISO/IEC 14496-12:2012 – Information Technology – Coding of audio-visual objects – Part 12: *ISO base media file format*.
- [4] ISO/IEC 14496-14:2003 – Information Technology – Coding of audio-visual objects – Part 14: *MP4 file format*.