Audio Engineering Society

Convention Paper

Presented at the 117th Convention
2004 October 28–31 San Francisco, CA, USA

Establishing a Reference Playback Level for Video Games

Dr. Mark A. Tuffy

THX Ltd, San Rafael, CA 94903 USA
mtuffy@thx.com

ABSTRACT

Over the last two decades, there have been dramatic advances in video game technology. In this time, audio for games has moved from monophonic beeps to full 5.1 surround sound, utilizing Dolby Digital and DTS. While the games industry has embraced these technologies, there are no standards or guidelines in place to ensure that game audio exploits the potential of this delivery mechanism. As a result, there is still the push toward "louder is better." One element key to moving away from "loud" to "quality" is establishing a reference level for playback. This paper suggests such a reference level and why this would be logical for the games industry to adopt.

1. MULTICHANNEL AUDIO

Traditionally multichannel audio recording and reproduction have their roots in the movie industry. In this industry, multichannel audio consists of many disparate elements, which are combined to produce a soundtrack complementing the artistic ideal of the visuals. Given the diverse nature of these individual elements, the motion picture industry quickly saw there was the need to define procedures for the production and playback of content.

The initial establishment of playback levels for movie content can be found as far back as 1938 with the definition of the "Academy Curve." The movie standards in use today, stem from work by Dolby in the 1970s. This work has progressed to specifications for both recording and playback of movie content [1-3]. These standards allow movie content to be developed and moved from location to location, with the knowledge that its integrity will be maintained, and the artist's vision realized throughout production and post-production.

As home theater systems became more popular in the 1990s and brought multichannel audio into the home, manufacturers and technology companies created playback systems that delivered the content in a manner faithful to the title's original theatrical release. Today, the video games industry is increasingly utilizing the same electronic delivery technologies, providing multichannel surround experiences to video game players. This is not a niche market, with recent surveys of home theater owners showing that close to 40 percent have integrated a video game console—Xbox, Playstation, or GameCube—into their home systems.

Unlike the movie industry, there are no defined standards for the production and playback of game audio. This leads to a wide variation in the perceived quality of game audio.

2. GAME AUDIO PRODUCTION/PLAYBACK

2.1. Similarities/Differences with Other Industries

In many ways, game audio production now mimics the workflow of audio work for the movie industry. In game production an artificial soundtrack is produced by the creation of many different audio assets (similar to movie sound post-production), which are combined to provide a realistic audio counterpart to the visual presentation.

However, unlike movies and broadcast programming, where the intent is for the soundtrack to complement the visual action, game audio plays a more critical role in the interactive experience. It provides the player with information on items that are not necessarily in their field of view. This audio cue may be vital to succeed in the game. For instance, in a racing game, the sound of an approaching vehicle heard from one of the rear speakers (out of view), will affect how the player reacts, and may determine the outcome of the race.

One major difference in making the video game experience immersive is the co-existence of many discrete effects, sounds and music, which trigger interactively in the game. Every time a game is played, the user experiences a new "interactive mix."

2.2. Lack of Standards and the Impact on Content

Given the importance of audio to the game playing experience, the lack of standards in the industry raises many problems. Without any industry definitions for audio studios and how they should be set up (reference levels, relative dialogue levels), the artistic intent is often lost. A game player does not get to experience the full impact of the audio.

In many cases, to try and counteract this problem, the decision is made during production to choose a nominal sound and normalize everything to that level. The dynamic range of the audio is severely compressed, meaning every audio asset in the game can now be heard. In some cases, production of a game may involve the artistic choice to compress audio to achieve a certain effect, but compression is often used to satisfy the mantra of "let's make it loud!" This has a detrimental impact on the artistic intent of each asset and the feel of the game, and leads to experiencing an aural assault.

This practice, combined with the lack of standardization in the games industry, means that audio content level is extremely variable —whether you are comparing different titles on the same platform or titles across multiple platforms. Each switch in title results in the user having to make substantial system volume changes.

This effect becomes apparent when game audio is compared to other content that uses the same delivery pipeline (DVD, broadcast and music). For example, as users flip between video games and DVDs, their home entertainment systems will experience dramatic shifts in playback levels. This difference is especially noticeable to the consumer with titles that tie closely into movie or broadcast content licenses, where comparisons can be made between the two.

A key path to start correcting these inconsistencies is to set a standardized reference level for game audio playback. This would not only maintain a more consistent user experience for game players, but it would also allow more artistic exploitation of dynamic range within the audio soundtrack. In order to do this, two main areas have to be defined: the audio recording reference level and the calibration level of the playback system. Therefore, it is important to examine how other similar art forms or industries work, especially when technologies, equipment, and production talent from these industries are being integrated into game development.

3. CURRENT INDUSTRY REFERENCE LEVELS

3.1. Motion Picture

The motion picture industry has the longest experience in multichannel audio production. In movie post-production environments, the equipment being used is designed with standards for recording levels, calibration and playback references. This means a sound mixer can create an atmosphere that audiences will realistically experience when they go to see the movie in a professional cinema or on a multichannel home playback system. This reference level allows the mixer to place both subtle and aggressive audio assets in the same context, and have the user experience them both.

While the technology used for delivery of the audio content has moved from magnetic striping to digital delivery, specifications have been set to allow consistency of content. The foundation of this is the setting of a recording reference level.

3.2. Recording Reference Level

The recording reference level set out by the motion picture industry is defined in [1] as:

$$-20dBFS \quad (1)$$

for digital recording media, with the corresponding analog recording level on magnetic media of:

$$185nWb/m \quad (2)$$

These recording reference levels were set to reflect the increased 20dB of headroom in magnetic media, while maintaining consistency of content. For digital delivery, the 20dB of headroom allotted in this specification allows the inclusion of short dramatic portions of the soundtrack to enforce action on the screen.

(It should also be noted that the recording level for digital television tape recorders [4] is also set at -20dBFS.)

This standard reference also allows for content to be moved between disparate studio locations while ensuring the quality is maintained (something not common in game audio production) and the balance of the mix remains constant.

3.3. Playback System Reference Level

With the reference level fed into the playback system defined in (1), the next step is to determine the physical sound pressure level from the playback chain given this input. In [3], the methodology and calibration of motion picture multichannel sound systems is defined.

- For individual screen channels, the spatially averaged SPL is defined as:

$$85dBc \quad (3)$$

- For surround loudspeakers:
  - Single surround channel

$$85dBc \quad (4)$$

  - Two surround channels

$$82dBc \quad (5)$$

(When fed a simultaneous in-phase signal, these channels will acoustically sum to give the reference level of 85dBc.)

  - For subwoofers:

+10dB of in-band gain (compared to a screen channel) when viewed on a spectrum analyzer.

Given these reference levels, it is important to examine the translation of these to the home.

3.4. Consumer Electronics

The consumer electronics industry uses the same digital delivery pipeline for audio as the motion picture industry (Dolby Digital or DTS). In the home, multichannel playback systems were originally launched as "home theaters." With this, the standards for reference playback of movies in the home were also adopted.

3.5. Electronic Reference Level

The electronic reference level set out for digital delivery of Dolby Digital or DTS content is:

$$-20dBFS \quad (6)$$

This means that any DVD or broadcast content encoded in Dolby Digital or DTS will have a reference level of -20dBFS.

3.6. Playback Reference Level

In the consumer electronics industry, there is an established reference level standard based around the motion picture level of 85dBc. For home systems, though, there is not an array of surround speakers. This means that all the main loudspeakers (Left, Center, Right, Surround Left, Surround Right for a 5.1 system) are all set to:

$$85dBc \quad (7)$$

This reference level can be achieved readily using either test noise embedded in products, or a commercially available calibration disc, along with an inexpensive sound pressure level meter,

The consistency of electrical reference and playback reference level between the professional and consumer electronics worlds means that any variability between content that exists is due to a deliberate artistic intent on the part of that content creator—not by the playback system.

4. PROPOSED REFERENCE LEVEL

Given the continuity of the production world for movies and broadcast, and how this content is replayed in the home, it is important to see how this may affect a reference level standard for game production.

In game development, the same production equipment and codec technology is used as in motion picture creation and playback.. In consumer electronics, the same audio codecs are again being used for multiple forms of media playback in home systems, which are increasingly being used for game playback.

With this in mind, it does not appear appropriate for the games industry to try and produce new reference standards for recording and playback. Rather than promoting continuity between different media, this approach would isolate the games industry. It is therefore proposed that instead the games industry support a reference recorded level of:

$$-20 dBFS \quad (8)$$

this equates to a proposed calibrated reference playback level of:

$$85 dBc \quad (9)$$

The 85dBc reference applies to all of the main loudspeaker channels (Left, Centre, Right, Surround Left, Surround Right), The reference is set by exciting each speaker individually with wideband pink noise and measuring with a sound pressure level, applying a slow "C" weighting response. By contrast, the subwoofer channel should exhibit +10dB of in-band gain as measured with a real-time analyzer.

4.1. Why 85dBc?

The proposed reference level is appropriate for games because it matches the standards already set forth for film and DVD production. In the broadcast industry, the +10dB LFE offset is also an established standard. In addition, adopting this standard will not require a re-design of production tools or consumer electronics components that are used for the current development and playback of games content.

This standard also allows the seamless integration of film and DVD assets into video game production and allows for easy level matching of audio assets gleaned from such content.

Some in the games industry contend that an 85dBc monitoring level is not consistent with the levels a consumer replays content. Their alternative proposal is that the game monitoring playback level should be lower to "hear the content the same way as the consumer." In practice, reducing the reference level would not result in this. In [5] it is described how measurements of 50 typical living rooms produced a background noise level of NC17. This level is considerably lower than the average mix studio used for game production. Reducing the reference level for games would result in consumers hearing content game developers would not hear in their noisier mix studios. Without doubt, this will have an adverse impact on the game.

The best way to ensure this does not happen is to raise the monitoring level to a point where all of the content can be heard during production. This ensures there are no surprises in the audio when played back in the home, even if replayed at a lower level.

5. CONCLUSION

The video game industry, while utilizing some of the practices and technologies from motion picture post-production, is hindered by the lack of standardization. This affects both the production of content and the playback experience for the consumer. To create consistency in the way game audio is created by developers and played back on consumer systems, it is proposed that a recorded reference level of **-20bdBFS** and playback reference level of **85dBc** are adopted by the video game industry.

As video game consoles are increasingly integrated into home theater systems, it is logical to adopt the playback reference already established for content used on these devices and systems. By setting this reference level, the games industry will be able to fully exploit the dynamic range in titles, producing more lifelike and realistic experiences for the gamer.

6. ACKNOWLEDGEMENTS

This work was supported by Steve Martz, Andrew Poulain and Jerry Zernicke of THX Ltd., each of whom provided expertise on motion picture-related matters.

7. REFERENCES

[1] SMPTE EG 9-1995 “Audio Recording Reference Level for Post-Production of Motion-Picture Related Materials”.

[2] SMPTE 202M-1998 “Dubbing Theaters, Review Rooms and Indoor Theaters-B-Chain Electroacoustic Response”.

[3] SMPTE RP 200-2002 “Relative and Absolute Sound Pressure Levels for Motion Picture Sound Systems”.

[4] SMPTE RP 155-1997 “Audio Levels for Digital Audio Records for Digital Television Tape Recorders”

[5] Tomlinson Holman. “5.1 Surround Sound Up and Running” Focal Press 2000