

DSP Compensation Algorithms for Small Loudspeakers

Meir Shashoua
Chief Technical Officer
Waves Audio Ltd., Tel Aviv, Israel
Meir@waves.com

Paul Bundschuh
Vice President of Sales & Marketing
Waves Audio Ltd., Austin, Texas
Paul@waves.com

Abstract

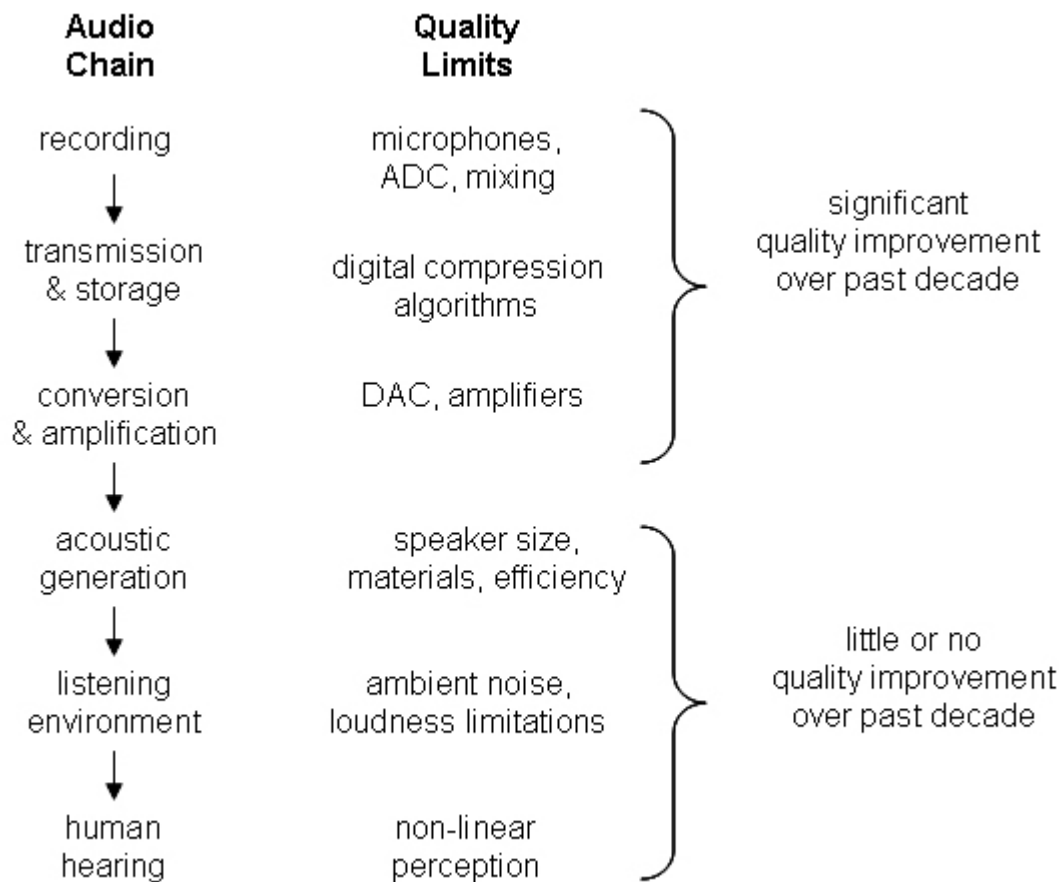
Today it is possible to deliver, store, and amplify a high quality digital audio bitstream in a wide range of consumer electronics products. However, these technical advances have often not improved the perceived audio quality due to the acoustic limitations in small loudspeakers and noisier listening environments. New psychoacoustic loudspeaker compensation algorithms enable dramatic improvements in audio quality by compensating for limitations in frequency response and dynamics in mainstream consumer electronic products including mobile phones, LCD TVs, notebook PCs, and portable audio systems.

1 Audio Limitations in Today's Digital Products

Over the past decade, digital audio technology has seen significant improvements in recording, transmission, and amplification, allowing high quality electronic audio signals to be generated from almost any device. However, these advancements often have not resulted in better audio quality to consumers due to acoustic and other limitations in the audio delivery chain. The acoustic and perception half of the audio delivery chain, as shown in Figure 1, has seen little improvement. In many of today's smaller, thinner, and more power efficient digital products, smaller loudspeakers create more severe acoustic limitations than a decade ago.

Continuing to improve the first part of the audio delivery chain in recording, compression, and amplification is beneficial for consumers with several thousand dollars to invest in audio electronics, speakers, and room acoustics, but it may deliver little, if any, perceptual improvement in mainstream consumer audio products with severely constrained speaker size and power consumption. In contrast, since the acoustic limitation of small speakers are the quality bottleneck, improvements to this part of the audio delivery chain can enable dramatic improvements in audio quality for mainstream consumer electronic applications.

Figure 1. Audio Delivery Chain Limitations



Psychoacoustic audio signal processing technologies can be used to compensate for limitations from the acoustic and perception half of the audio delivery chain. Waves Audio Ltd, the worldwide leader in audio signal processing tools to professionals, offers a suite of DSP algorithms that compensates for acoustic limitations in loudspeakers, high background noise, and non-linear human perception.

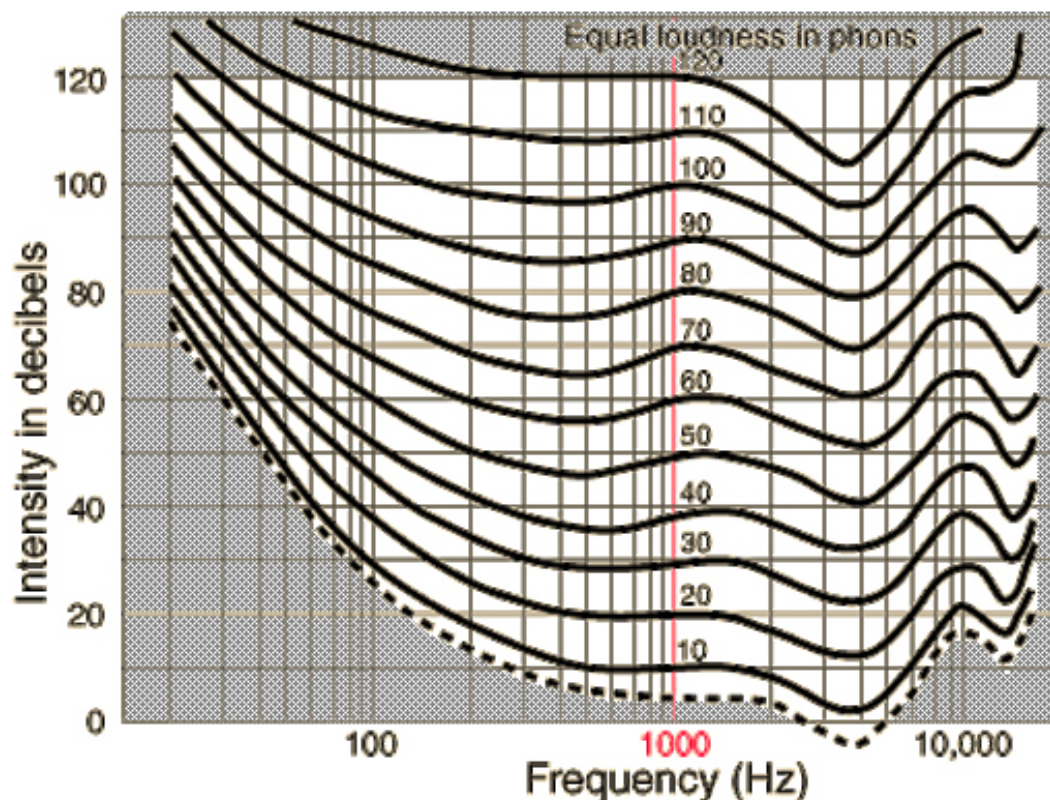
These technologies are delivered in DSP algorithms that optimize the audio signal for the frequency response and dynamics limitations of the reproduction system and acoustic environment. Waves' solutions for these technologies are available in highly customized, low-cost DSP from Waves, licensable algorithms for third party DSP silicon products, or software Audio Processing Objects (APOs) for Microsoft's Vista operating system.

2 Perceptual Audio Limitations

2.1 Capabilities of the Human Hearing System

The final link in the audio delivery chain is that of human perception. Numerous studies have determined that two of the largest factors that determine the perception of quality are wide frequency response and wide dynamic range. The human ear can perceive a wide range of audio frequencies and dynamic range. This is illustrated in Figure 2 as the Audio Perception Window™.

Figure 2. Audio Perception Window™ of Human Hearing
Frequency Response 20 Hz to 20,000 Hz, Dynamic Range ATH to 120 dB



Frequencies below 20 Hz and above 20,000 Hz that are outside our frequency response thresholds are shown in grey as they cannot be perceived.

The Fletcher-Munson equal-loudness contours are also illustrated. The dotted line is the Absolute Threshold of Hearing (ATH). Sounds below this threshold are also shown in grey as they cannot be perceived. A major limitation of the human hearing perception system is the inability to perceive deep, but quiet sounds due to the sharp perceptual roll-off at lower frequencies.

There is no restriction on the maximum volume levels that can be perceived; however, this is limited by the hearing damage created by loud sounds over long term exposure. 120dB SPL is generally considered the loudest sound that should be generated for short periods without risking long term damage.

2.2 Audio Recording, Mixing, and Mastering

After the inherent frequency response and dynamic limitations of the hearing system is considered, it is then logical to consider the first step in the audio delivery chain. How do audio professionals in content creation of movies, music, and video games use frequency response and dynamic range? Professionals in audio content creation and recording optimize their audio content for the frequency response and dynamic range for the expected listening environment. This process of optimizing the audio content is called mastering.

Movie sound tracks are mastered for large, high quality cinemas, and most popular music is optimized for car audio and use in dance clubs.

"A key goal in the creation of a movie's soundtrack, is to deliver to the listener an emotional experience supporting the art of moviemaking and audio story telling. We use a wide palette of sonic tools in the creation of the soundtrack which consists of dialogue, music and soundesign. By utilizing these elements over the entire frequency and dynamics spectrum, available in today's cinema environment, we maximize the listener immersion and emotional involvement," said Scott Gershin, Soundesigner & Sound Supervisor - Soundelux. "The same soundtrack used in the creation of the movie is used to create the TV and DVD versions. Unfortunately, many consumer products fail to match the frequency and dynamics range of the cinema environment, resulting in a substantial loss of the movie's sonic experience and perceived audio quality."

Soundelux is a leading post-production sound company. Since its launch in 1982, it has worked with every motion picture studio in Hollywood, bringing the most comprehensive level of audio expertise and service to the content community.

2.3 Cinema and High Quality Reproduction Environments

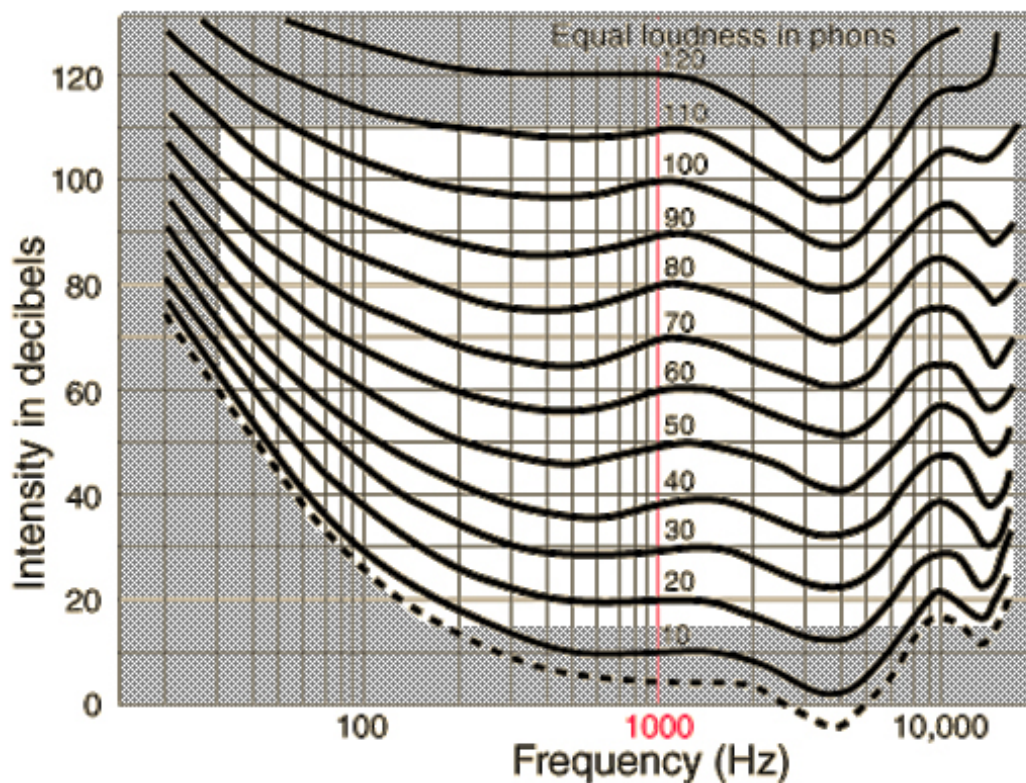
With the exception of listening to a symphony in a top concert hall, today's cinemas provide most consumers their best acoustic environment. High quality cinemas utilize a large set of speakers including subwoofer arrays with carefully designed room acoustics. This allows a broad frequency response and high SPL (Sound Pressure Level or volume). These cinemas are also designed with good sound isolation allowing quiet passages to be heard clearly without masking from background noise.

Audio systems found in concert halls, large cinemas, and high-tier home theater systems generally do a satisfactory job of reproducing audio across both the wide frequency range and dynamic range, and as a result, can deliver the audio quality and emotional impact desired by the audio content creator.

As result, movie soundtracks and high quality cinemas are designed together to deliver almost the entire wide frequency response and dynamic range of human perception as shown in Figure 3.

Figure 3. Audio Perception Window in Cinema

Frequency Response 30 Hz to 20,000 Hz, Dynamic Range 15 to 110 dB



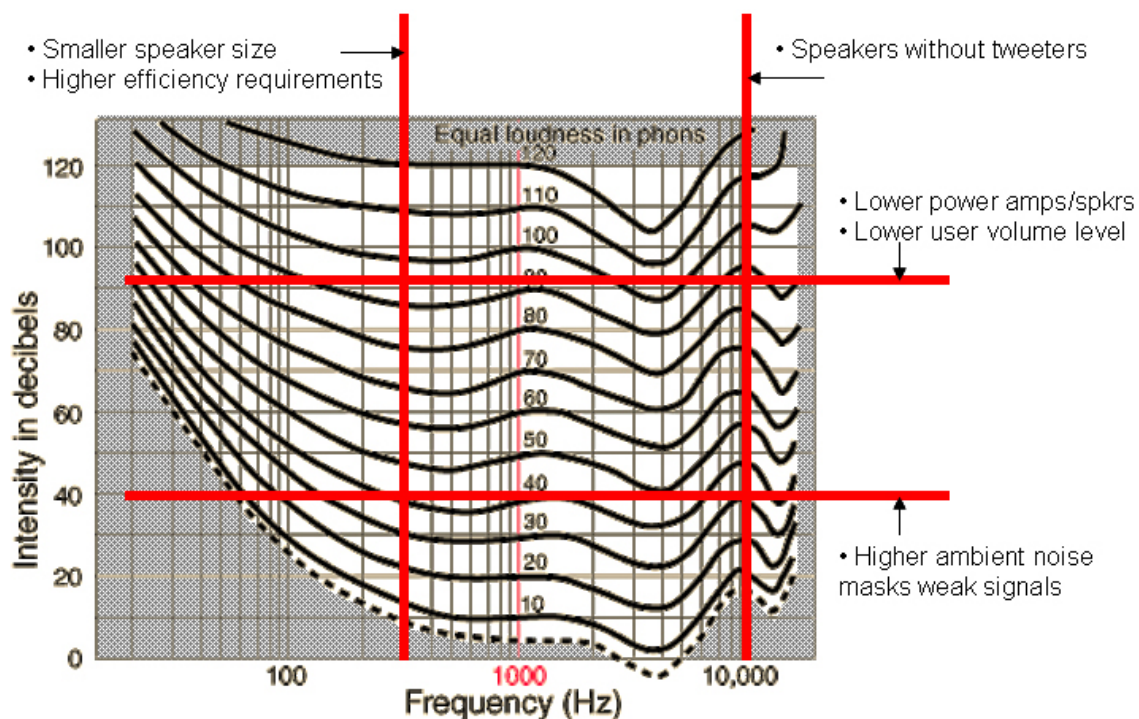
Unfortunately, to deliver these results requires several thousands of dollars in speakers, electronics, and room acoustics. This is not practical for a wide range of mainstream audio applications including mobile phones, LCD televisions, notebook PCs, and portable speaker systems.

3 Loudspeaker Design Effects on the Audio Perception Window

3.1 Design Limitations

Several factors limit the ability to deliver wide frequency response and dynamic range in consumer electronic products. These effects close the size of the audio perception in each dimension as illustrated in Figure 4. Together these issues severely constrain the frequency response and dynamic range that deteriorates the perceived audio quality.

Figure 4. Loudspeaker Design Effects



- **Bass Frequency Response** is limited by speaker transducer size, enclosure volume, design, speaker efficiency, and amplifier power.
- **High Frequency (Treble) Frequency Response** is limited by transducer weight and sensitivity, usually a problem in one-way speaker systems without special tweeters.
- **Maximum Signal Strength** is limited by the maximum amplifier power, speaker efficiency, speaker excursion, and the user volume level.
- **Minimum Perceived Signal Strength** is limited by the ambient or background noise of the listening environment.

3.2 Audio Reproduction Limitations in LCD TVs

Televisions often play movies and other audio content optimized for wide frequency range and dynamics, yet unless consumers utilize a high quality external home theater system, they can not achieve even close to cinema audio performance. It is ironic that with today's market drive for more expensive HDTVs, the audio reproduction capabilities have not generally improved.

In fact, the frequency response and dynamics limitations with today's popular flat screen LCD TV models are more severe than the "obsolete" analog CRT TV models they replace. For example, CRT televisions will usually play frequencies at least an octave lower than the LCD TVs of the same screen size. The ability of these TVs to receive and decode digital audio transmissions does nothing to resolve these fundamental issues in audio reproduction.

Figure 5 Audio Perception Window in LCD TVs
Frequency Response: 200 Hz to 14k Hz
Dynamic Range: 90dB to 30dB (ambient noise at home)

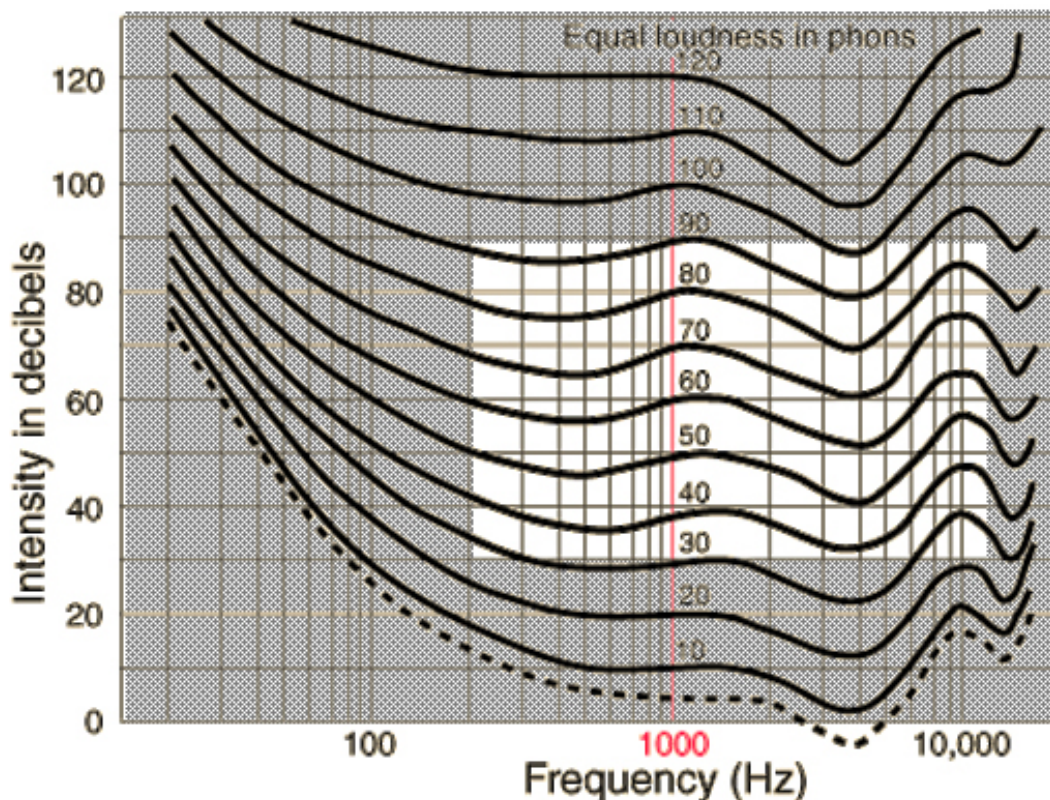


Figure 5 provides the frequency response and dynamic capability of a typical LCD TV model. Some of the major user limitations will be:

- The lowest frequencies with most of the emotional content are lost when audio is played from the built-in TV speakers.
- The reduced dynamic range means that the quietest passages may be lost to background noise, while at the same time, the loudest passages can be too loud especially for listening at night.

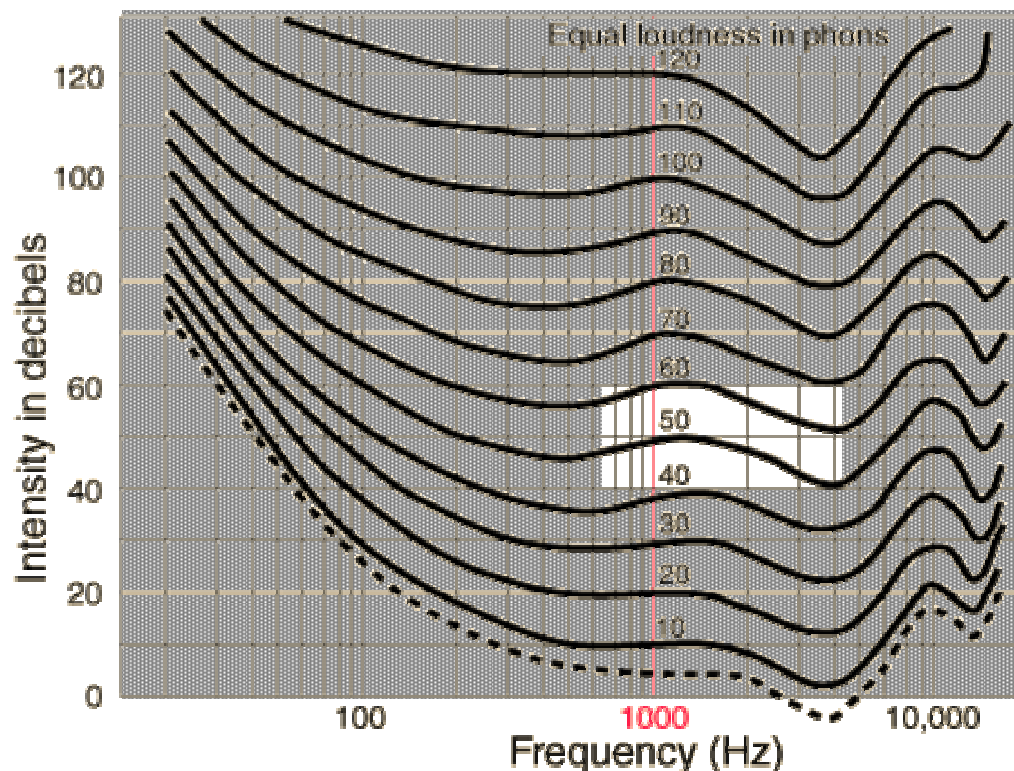
3.3 Audio Reproduction Limitations in Wireless Handset Speakerphones

Another interesting product to examine is the speakerphone found in today's wireless handset or mobile phone. Although the speakerphone is used for handsfree calling, these tiny speakers are increasingly expected to reproduce a variety of audio sources including personalized music ringtones, videoconferencing, web surfing, and even movie content.

The ability to listen to any wide dynamic range of content over such systems can stress the user's patience between audio clipping and not hearing the content at all. The reason for these problems can be shown in Figure 6, where the Audio Window has shrunk dramatically.

Figure 6. Audio Perception Window for Wireless Handset Speakerphones

Frequency Response: 600 Hz to 3.2k Hz
Dynamic Range: 60dB to 40dB (ambient noise in office)



4 Frequency Response Compensation Algorithms

4.1 Bass Boost Functions for Consumer Electronics

The frequency response of small loudspeakers is severely reduced in bass frequency response. The loudspeaker frequency response is relatively flat in its passband and then drops rapidly at frequencies below its roll-off frequency (f_3). The acoustic efficiency of audio frequencies below the roll-off frequency are only a small fraction of that in the passband with the remaining energy dissipated as heat which can damage the loudspeaker.

The rate of the loudspeaker roll-off is -12dB/octave for sealed enclosure designs and -24dB/octave for ported enclosure designs. As result, the acoustic energy output of an audio signal one octave below the roll-off is $\frac{1}{4}$ that of the passband for sealed enclosure speakers or $\frac{1}{8}$ that of more commonly used ported enclosure speakers.

The industry has used a variety of bass boost or specialized EQ methods to address this need for decades. These methods all seek to push the loudspeaker electro-mechanics harder by increasing the signal and energy levels to compensate for this sharp loss in efficiency.

These traditional bass boost or EQ techniques all suffer from the same inherent problems:

- Bass boosts require high levels of system gain or headroom, increasing requirements and costs in peak amplifier power and speaker excursion.
- Bass boosts reduce overall speaker efficiency and increase system power consumption.
- Bass boosts destroy the frequency balance of the music or audio from the artist. It also generates an unnatural boomy sound.

4.2 Waves MaxxBass® psychoacoustic bass extension

In response to this industry need, Waves developed and patented the MaxxBass® psychoacoustic bass extension technology. MaxxBass differs dramatically from traditional bass boost technologies in that it does not seek to push the loudspeaker harder where it is least efficient. Instead of increasing the gain at frequencies far below the roll-off frequency, MaxxBass removes these potentially damaging and inefficient audio signals and instead creates higher frequency harmonics of these signals which are perceived as the original bass signals. Since MaxxBass utilizes a deep understanding of psychoacoustics, how we perceive frequencies, the effect is completely natural and has been used for years by top mixing and mastering engineers.

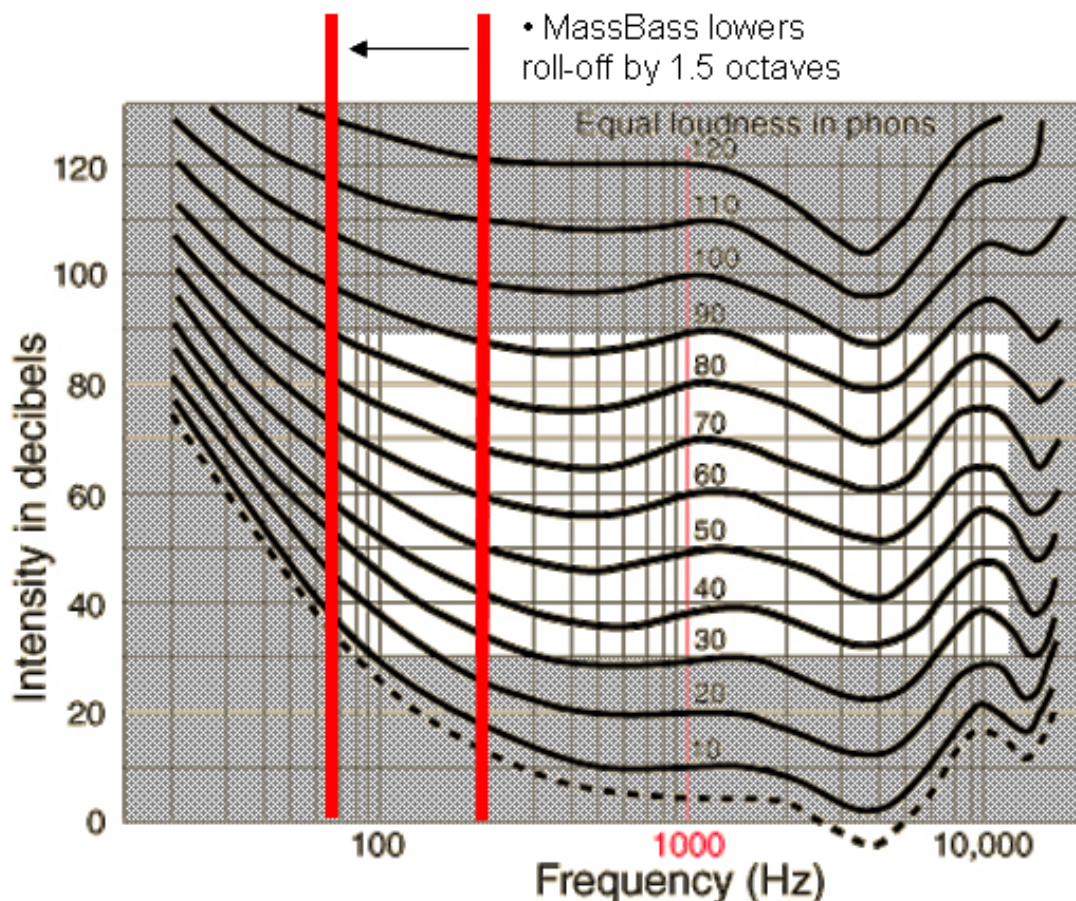
Since the speaker is not forced to operate below its efficient operating range, the overall system efficiency is improved. As result, MaxxBass has been widely accepted in battery operated audio systems as the only effective technology for improving bass response.

MaxxBass uses patented psychoacoustics to extend the perceived bass frequency response by about 1.5 octaves. This allows the audio perception window to open significantly to support lower frequencies. This improves the emotional impact of audio delivered through the bass frequencies.

4.3 MaxxBass Improvement in LCD TVs

Figure 7 shows the relative improvement in the LCD TV audio perception window using MaxxBass. The perceived low frequency roll-off can be shifted from about 200Hz to about 65Hz. This gives the listener a much improved listening experience by delivering a wide range of bass frequencies that would not otherwise be reproduced.

Figure 7. MaxxBass Improvement in LCD TVs



MaxxBass also compensates for the low frequency roll-off in hearing perception. The balance of deep bass to higher frequency stays consistent across all volume levels, which obsoletes the traditional loudness function. This allows the user to hear the bass frequency content even at low volumes, which is not normally possible.

5 Dynamic Range Compensation Algorithms

5.1 Automatic Gain Control Functions in Consumer Electronics

The problem of limited dynamic range in consumer products is also well known. The traditional industry response to this problem has been to utilize Automatic Gain Control (AGC) functions. For example, mobile phones use AGCs to provide a more consistent signal level and volume output to the user. In advanced television sets, an AGC is often labeled as smart or intelligent volume control to reduce volume variation.

An AGC measures the average signal strength over a relatively long sample period of 100 milliseconds or longer. It compares the average signal strength to the desired average signal strength and increases or decreases the system gain slowly until they match.

These traditional AGC are relatively simple to implement and provide some improvement in clarity; however they all suffer from the same inherent problems in that they can not respond to rapid changes in signal level.

- Strong audio peaks clip and cause distortion.
- Rapid drops in level cause the signal to be lost and to be temporally unintelligible.
- If the AGC is used too aggressively, the audio signal sounds “lifeless” since too much of dynamics is lost.

5.2 Compressors and Limiters as Professional Dynamics Tools

Professionals involved in recording, mixing and master music have many of the same inherent problems with dynamic range; however, they don't use crude automatic gain control functions, but prefer more flexible and powerful professional signal processing tools such as compressors and limiters.

AGCs work on the assumption that the audio signal is a relatively steady state and only the gain or volume needs to be adjusted. However, audio signals as described in section 2 usually have wide dynamic range. Compressors and

limiters allow professional audio engineers to carefully optimize the dynamic response of the audio signal without the problems inherent in AGCs.

The key to accomplishing these tasks for professional audio engineers is to use transparent signal processing tools. Transparent means that the processed audio signal does not sound like it has been processed. Often this means that the tools need to create these audio changes without altering critical perceptual features of the signal such as attack time.

Waves introduced its L1 Ultramaximizer in 1994 as the industry's first transparent peak limiter. This groundbreaking signal processing tool allowed the processed audio with higher output levels to sound exactly like the unprocessed audio with greater dynamic range. Since this time, Waves' dynamics tools including limiters and compressors have become the recording standard worldwide and are used on virtually all major movie soundtracks, music, and videogames released worldwide.

5.3 Waves MaxxVolume™ Dynamic Range Compensation

Volume is the most important user control in an audio system, and until now, this basic control has not taken advantage of the advancements in audio signal processing to optimize for a system's dynamic range. Waves MaxxVolume is an entirely new signal processing technology for a wide variety of consumer audio products with limited dynamic range. MaxxVolume is a patented technology that utilizes the same dynamic technologies found in Waves award-winning professional audio tools.

MaxxVolume combines the following audio functions:

- A transparent limiter function allows the volume level to be increased far beyond normal levels without clipping or distortion based on Waves famous **L1 Ultramaximizer** processor. This function allows the maximum output (volume) levels to be increased up to +12dB higher without distortion. This increases output 4x without increasing the cost of amplifiers or speakers in the product.
- A compression function is used to improve clarity and intelligibility in noisy environments based on Waves award-winning **Renaissance Compressor**. This enables dramatically clearer audio in noisy background environments or when the volume levels are significantly reduced.
- A noise gate function is used to eliminate noise from the sound card, CD player, and background noise in the original recorded sound.
- MaxxVolume also supports an optional auto leveler function using a more traditional AGC approach. This function enables consistent volume or

listening levels, despite significant level changes in the audio input. It is particularly beneficial to users watching TV or listening to radio since it reduces annoying listening level variations from commercials that require continuous volume adjustments.

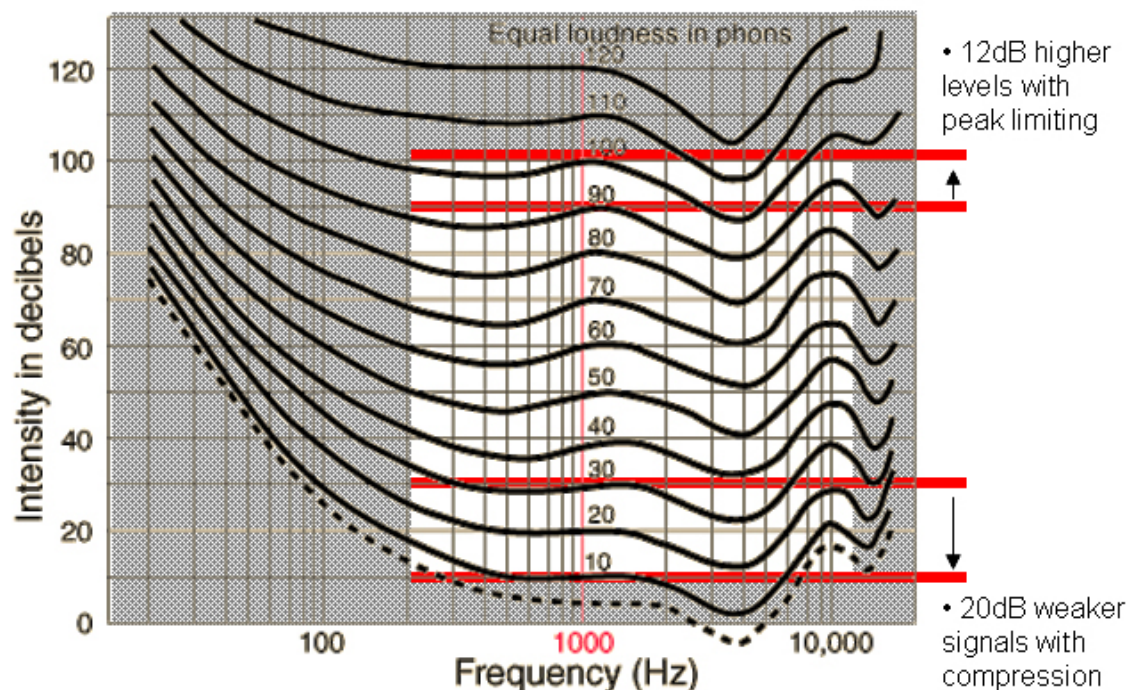
5.4 MaxxVolume Improvements in LCD TVs

The transparent peak limiter in MaxxVolume opens the Audio Perception Window by enabling a +12dB increase in peak signal levels to be generated without distortion or more expensive amplifiers and speakers.

The low level compressor enables a weaker signal to be perceived that would otherwise be lost in the ambient noise of the listening environment. This is particularly problematic when wide dynamic range content, such as movies, is listened to at lower volume levels. In this situation about a +20dB improvement in perceiving lower signal can be obtained.

The overall improvement in the Audio Perception Window can be seen in Figure 8.

Figure 8. MaxxVolume Improvements in LCD TVs

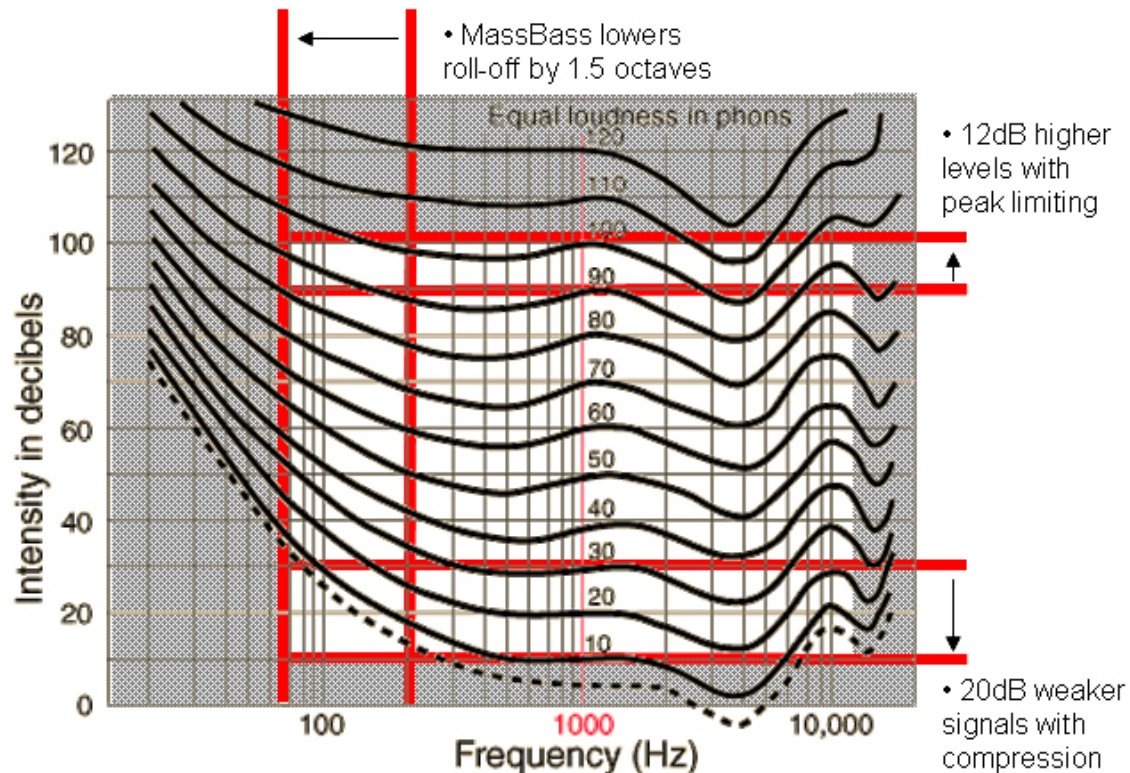


6 Summary

Frequency Response and Dynamic range compensation technologies can increase the effective range of operation of small loudspeakers systems. Together, these technologies dramatically improve the Audio Perception Window enabling dramatic improvements in perceived audio quality.

An example of the improvement in the Audio Perception Window for an LCD TV application is shown in Figure 10. Similar performance benefits can be obtained for a wide range of consumer audio products including speakerphones in wireless handsets, notebook PCs, and portable audio systems.

Figure 10. MaxxBass and MaxxVolume Improvements in LCD TVs



Besides enabling dramatic audio quality improvements, OEMs can also use these capabilities to improve their product designs in other ways. This includes lowering costs of amplifiers and speakers, reducing system power consumption, and enabling more ergonomic product design through smaller loudspeaker systems.

#

Maxx, MaxxBass, MaxxVolume, and Audio Perception Window are trademarks or registered trademarks of Waves Audio Ltd. in the U.S. and other countries.