Audio Engineering Society

Convention Paper 9373

Presented at the 139th Convention
2015 October 29–November 1 New York, USA

This Convention paper was selected based on a submitted abstract and 750-word precis that have been peer reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This convention paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library, http://www.aes.org/e-lib. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

A Method of Equal Loudness Compensation for Uncalibrated Listening Systems

Oliver Hawker[1] and Yonghao Wang[2]

[1] ohawker@me.com, [2] yonghao.wang@bcu.ac.uk

[1, 2] DMT Lab, Birmingham City University, Birmingham, B4 7XG, UK

ABSTRACT

Equal-loudness contours represent the sound-pressure-level-dependent frequency response of the auditory system, which implies an arbitrary change in the perceived spectral balance of a sound when the sound-pressure-level is modified. The present paper postulates an approximate proportional relationship between loudness and sound-pressure-level, permitting relative loudness modification of an audio signal whilst maintaining a constant spectral balance without an absolute sound-pressure-level reference. A prototype implementation is presented and accessible at [1]. Preliminary listening tests are performed to demonstrate the benefits of the described method.

1. INTRODUCTION

It is widely acknowledged that perception of loudness is nonlinear with signal energy [2], and that the relationship between loudness and signal energy is also dependent on frequency [3]. This suggests that the human auditory system can be modelled as exhibiting a varying frequency response dependent on signal-level, effective upon each component of a complex sound as a function of the intensity of each specific component. This implies an arbitrary perceptual change in spectral balance when the intensity of a sound is changed.

The ubiquitous "volume control" (as commonly found in audio equipment) employs a change of signal level to effect a change in perceived loudness, causing a change in perceived spectral balance, yet remains proliferous due to simplicity and ease of implementation.

In a typical consumer environment, potential permutations of playback hardware, headphones, amplifiers and speakers are numerous. Psychoacoustic processing is often deemed impractical due to the requirement for listening system calibration. A technique capable of spectral balance compensation without calibration could provide a better listening experience in casual, uncontrolled listening contexts as favoured by consumers.

The present paper presents such a technique, which uses a modified psychoacoustic model to apply approximate spectral balance compensation without the need for calibration.

1.1. Background

1.1.1. ISO Equal Loudness Contours

ISO226:2003 [4] (henceforth referred to as "ISO226") is a meta-analysis of equal-loudness data collected by 12 modern studies, which will be recognized as the standard equal-loudness reference throughout the present paper. ISO226 does not guarantee data outside of the following ranges, due to impracticality of measurement and discrepancies between studies:

- 20-90 phon for 20-4000Hz
- 20-80 phon for 5000-12500Hz.

Data outside of the guaranteed ranges will be extrapolated using the ISO226 loudness function if required.

1.1.2. Existing Solutions

Spectral balance compensation in this context has already been explored in the literature, such as the "Tone balance volume control" as detailed in [5]. Here, two biquadratic filters are applied with varying gain, frequency and slope characteristics in proportion to the level of attenuation applied, to apply spectral balance compensation.

Another similar technique is that of "Loudness Domain Signal Processing" [2], capable of performing signal processing operations (such as a change of loudness) in accordance with a psychoacoustic model. The approach employed by this framework differs from that described in the present paper, as it requires an absolute sound-pressure-level reference (calibration of the listening system) [2].

1.2. The "Approximate Spectral Balance Compensation" technique

At the core of the technique presented (henceforth referred to as Approximate Spectral Balance Compensation, or ASBC) is a simplified model of loudness derived from the equal-loudness contours (henceforth referred to as the Linear Model). It proposes that the change in sound pressure level required to effect a desired change in loudness can be approximately expressed as the product of the desired loudness change and a frequency-dependent constant of proportionality. In other words, it postulates a frequency-dependent proportional relationship between change in sound pressure level and change in loudness.

This simplification allows us to perform loudness adjustment whilst maintaining approximate spectral balance, without knowledge of the absolute loudness or sound pressure level. The constant of proportionality between change in loudness and change in sound pressure level (henceforth referred to as $k(f)$) is calculated to match the response of the ASBC model to that of the equal-loudness contours as closely as possible. A frequency-dependent linear loudness growth function is defined using curve-fitting techniques to match the equal-loudness contours, the gradient component of which $k(f)$ defines the approximate relationship between a change in loudness and a change in sound pressure level.

1.3. Terminology

In the present paper, "spectral balance compensation" and "loudness modification" are used interchangeably depending on the context, to describe the ASBC process. Both refer to a change in loudness whilst maintaining spectral balance.

The term "phon" is also used throughout. The phon is a quantification of the perceived loudness of a steady-state tone. The level of a tone in phon is equal to the signal energy expressed in dBSPL of a 1kHz tone of equal loudness [6]. The phon is used in ISO226 as the unit of loudness.

2. METHODOLOGY

2.1. Process Overview

ASBC attempts to provide a "best fit" loudness modification across all listening systems. Instead of applying a modification based on incoming program material and information regarding the sensitivity of the listening system, it applies a generic modification for required reduction in loudness. It aims to provide the best possible loudness modification across a specified range of potential listening situations.

This is achieved primarily by virtue of a fundamental redefinition of the equal-loudness contours. The ISO226 equal-loudness contours are defined as an arbitrary function, derived to fit experimental data [4]. This function can be approximated as a frequency-dependent linear equation, in order to represent a proportional relationship between change in loudness and change in sound pressure level.

2.2. Frequency-dependent constant of proportionality

The constant of proportionality is determined by fitting a linear function to data from ISO226 using the method of least squares, and taking the gradient per frequency, resulting in the frequency-dependent relationship between changes of loudness and change in sound pressure level, $k(f)$.

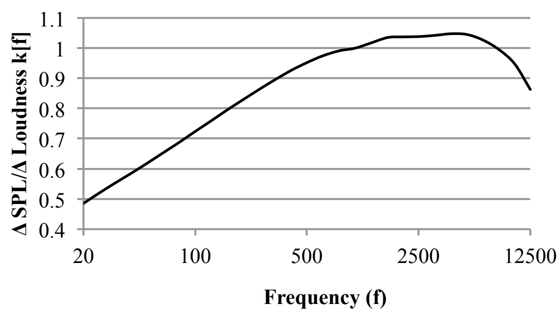

Figure 1 Approximate change in sound-pressure-level required for a loudness change of 1 phon, calculated using least-squares across the 10-120 phon range from ISO226

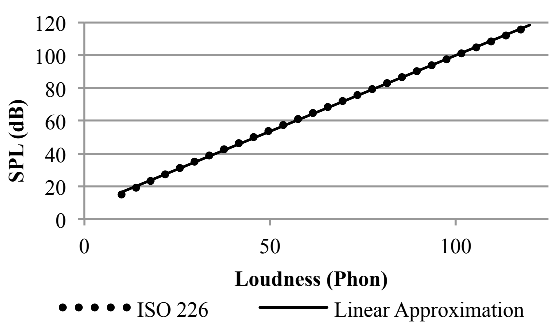

Figure 2 A linear approximation at 400Hz

Shown in Figure 2 is an example linear approximation at 400Hz. As shown, there is little difference between the two functions. For 400Hz, the maximum error is approximately ±1.41dB.)

2.2.1. Applying loudness modification with k(f)

For a required change in loudness, we can consider the required change ΔSPL[f] in dB to be defined as:

$$\Delta SPL(f) = \Delta phon \cdot k(f) \tag{1}$$

where $k[f]$ is a scalar representing the change in sound pressure level required to effect a change of 1 phon.

It may be useful to calculate the change in sound pressure level required to perform spectral balance compensation, after attenuation has been applied.

$$\Delta SPL(f) = (1 - k(f)) \cdot \Delta P \tag{2}$$

where ΔP gives the attenuation that is to be compensated for.

As shown by equations (1) and (2), the change in dB required to move between two loudness values is considered constant regardless of their absolute level, and depends only on the change in loudness required.

2.3. Linear Model

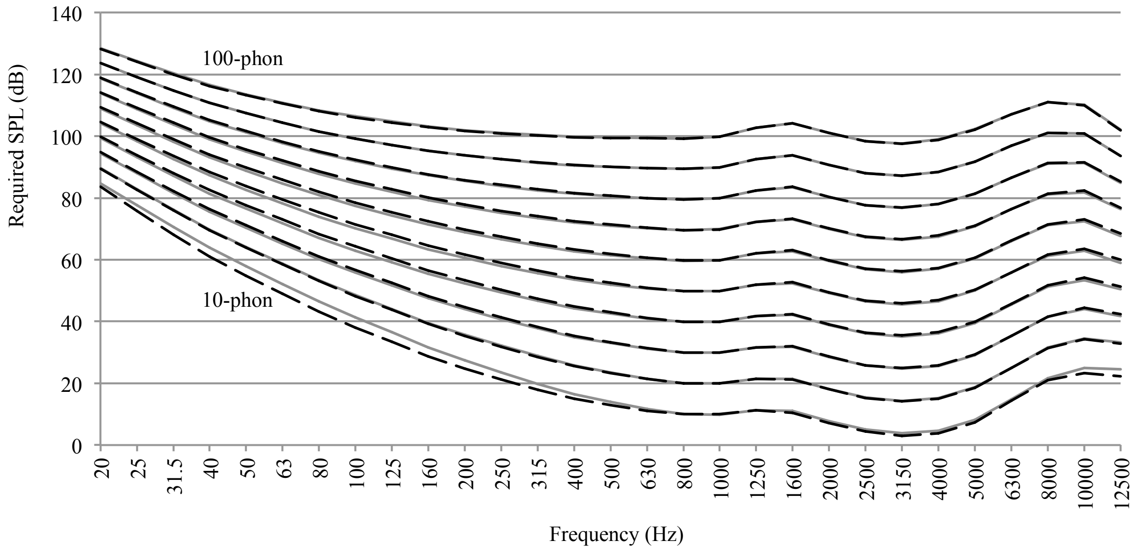

Figure 3 Comparison between ISO226 equal-loudness contours and the "Linear Model" contours effective when using $k(f)$. *(Solid Line: Linear Model, Dashed Line: ISO226)*

2.3.1. Deviation from ISO Contours

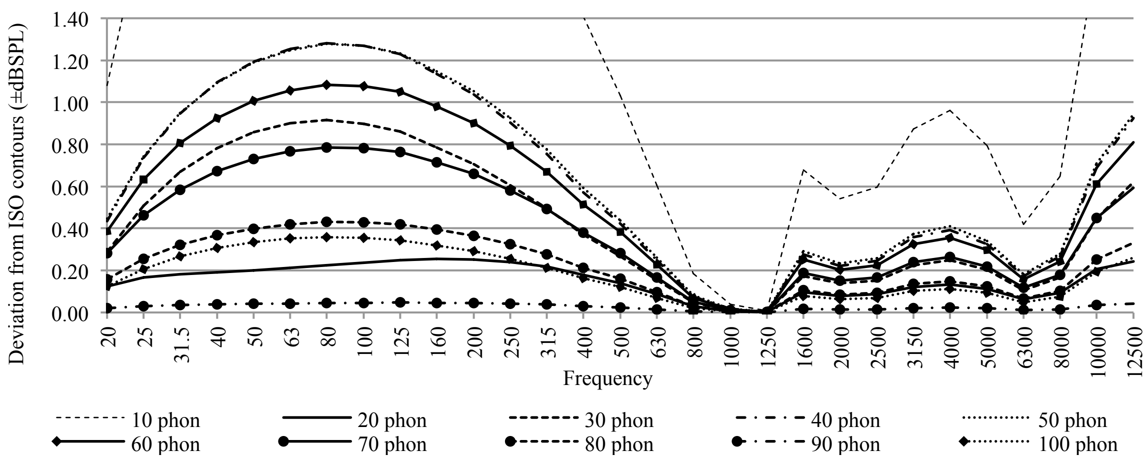

Figure 4 Maximum error introduced by the Linear Model as a function of frequency *(10-phon is omitted for clarity. Peak error at 10-phon is 3.27dB).*

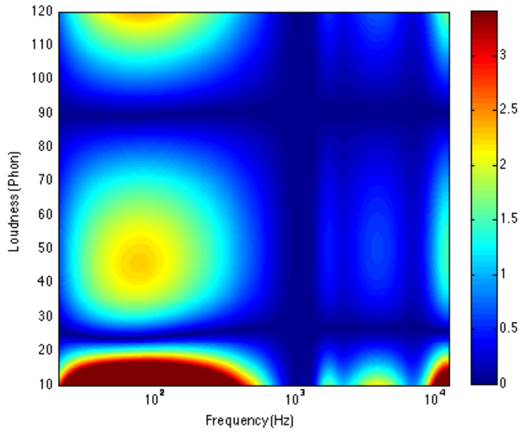

Figure 5 Error introduced by the Linear Model as a function of frequency and loudness

There is 0dB of error at 1kHz, as the phon is normalized against dBSPL at this frequency. The highest deviation is around the area approaching the threshold of hearing.

2.3.2. Probable Listening Scenarios

Disregarding improbable listening scenarios when calculating the linear approximation can further reduce error. For example, playback at the threshold of hearing is unlikely in a consumer environment, as is playback at the threshold of feeling, so the response to loudness at these levels need not be accounted for in the linearization. For example, a study from the European Commission [7] found the maximum sound pressure level for personal music players to range between 80-115dBSPL, suggesting 97.5dBSPL mean level. For a modification designed to apply up to a given maximum attenuation (e.g. 70dB), the following weighting scheme between 27.5dBSPL and 97.5dBSPL could be applied to reduce error in the useable range.

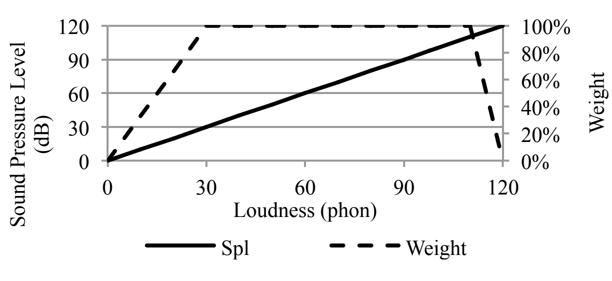

Figure 6 Example weighting scheme

Care must be exercised to avoid over-specialization, as wide range of loudness levels must remain accounted for. In order to successfully optimise the model using a weighting scheme, extensive and reliable data regarding listening habits is required, to minimize the risk of increasing error due to an inappropriate specialization.

3. LISTENING TESTS

3.1. Test Protocol

Preliminary double-blind listening tests were conducted on a random stratification of listeners, including critical and non-critical listeners, as well as a mix of age, gender and hearing ability. Tests were conducted using an iOS-based DSP platform, in an ad-hoc fashion. Test subjects were presented with 3 repeats of a set of 5 audio samples, each with an uncompensated attenuated variant and a variant having undergone ASBC loudness reduction. Test subjects were able to play and replay the samples for comparison at their leisure.

Samples were selected from a range of genres, and consisted of sections from the following recordings:

1. “Up, Up and Away”, The 5th Dimension, 1967
2. “Lovely Day”, Bill Withers, 1977
3. “Barcarolle in F# Major (Op. 60)”, Chopin, performed by Idil Biret, 2000
4. "Shining Star”, Earth, Wind & Fire, 1975
5. “Gaslighting Abbie”, Steely Dan, 2000

4-second excerpts were used, chosen for their broad spectral content. Loudness reduction was varied between 12 and 30 phon.

Test subjects were asked to identify which variant was “clearer”. This ambiguous criterion was deliberate; more specific adjectives risked introducing bias through forcing the subject to look for a particular characteristic, rather than relying on their own subjective interpretation of “clarity”.

A total of 27 listeners were tested, each with 15 audio samples, totalling 405 comparisons. Half of the tests were conducted using a pair of Sennheiser HD650 headphones, and half with Beyerdynamic DT990.

3.2. Test Results

Overall, listeners preferred ASBC for 87.8% of the tests conducted. As shown in Figure 7, the vast majority of listeners preferred the compensated variant to the uncompensated variant.

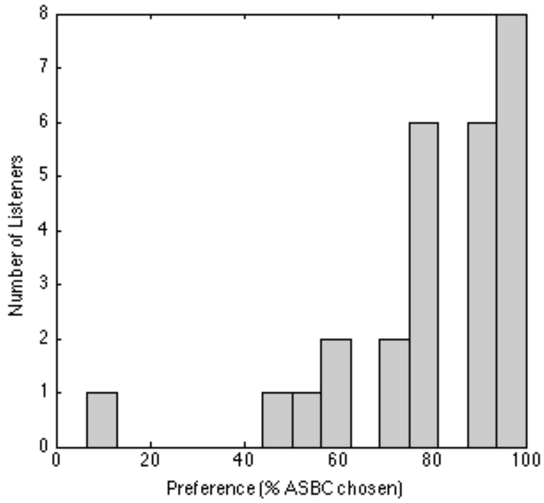

Figure 7 User preference of ASBC variant over uncompensated variant

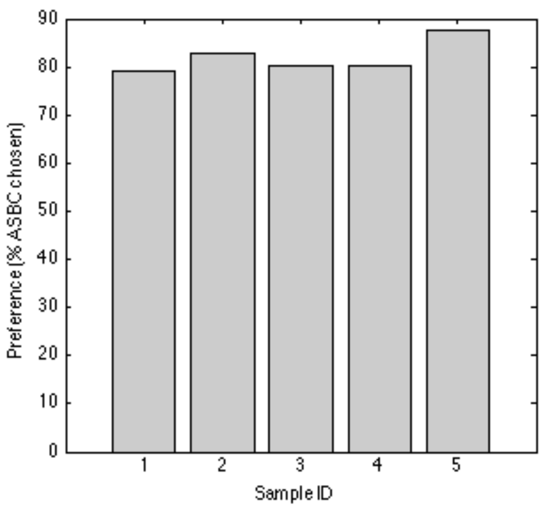

Figure 8 User preference of ASBC per audio sample

As shown, there was no statistically significant correlation between preference and audio sample, suggesting that the technique may work well across different genres. There was also no correlation between headphone type and preference.

4. DISCUSSION

4.1. Implications of a Linear Approximation

4.1.1. Flexibility

The primary implication of a linear approximation is, as stated earlier, the lack of requirement for knowledge of the absolute listening level, and subsequently for knowledge of the listening system. It postulates that the sensitivity of the listening system can be ignored in most cases, as applied loudness modification would be approximately the same regardless of the absolute sound pressure level being produced.

An issue faced by similar psychoacoustic processing systems is that of non-flat frequency response; a listening system must be calibrated in terms of frequency response as well as overall sensitivity for a valid compensation. However, as a non-flat frequency response can be modelled as an artificial change in absolute sound pressure level of specific frequency ranges, it can be safely disregarded, as the per-spectral-component loudness modification would still be the same regardless of the per-spectral-component sensitivity of the listening system.

4.1.2. Programme-independence

If the signal modification required for a given change in loudness is the same regardless of absolute sound pressure level, it can be argued that knowledge of the spectral content of incoming audio need not be accounted for when applying loudness modification. ASBC does not account for the composition of a complex sound or the level at which each component requires modification, because the modification is considered to be the same regardless of its level.

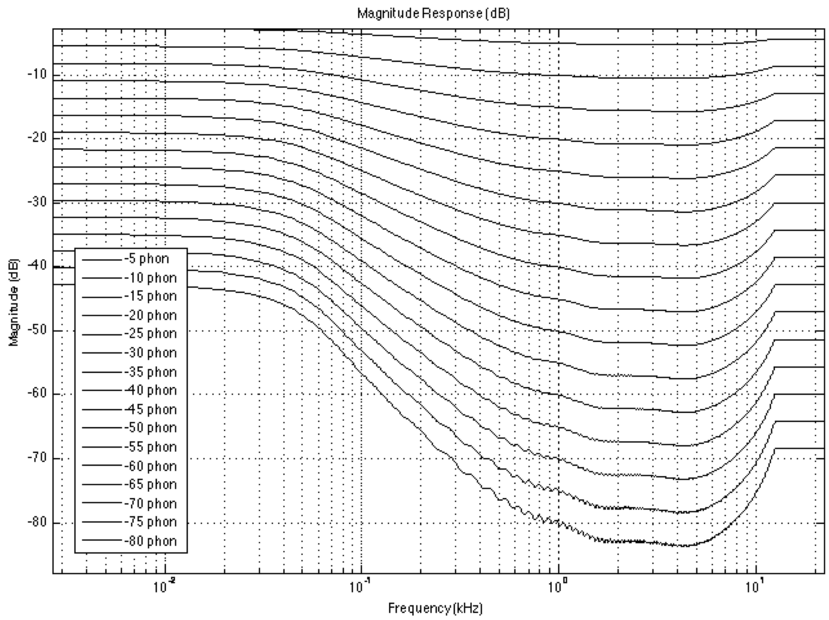

Figure 9 Magnitude response of filters required to perform a loudness reduction and spectral balance compensation for various reductions of loudness

4.2. Real-time implementation

Constant of proportionality $k(f)$ provides a specification for the change in sound pressure level required in order to effect a change in loudness of 1 phon. Consequently, a filter can be designed with characteristics corresponding to the product of said specification and any desired change in loudness. Figure 9 shows the magnitude responses for filters created for different changes in loudness, in -5 phon steps from 0 to -80 phon. The level of spectral balance compensation must change in proportion to the attenuation, which is reflected in the behaviour of the filters.

4.2.1. Real-time design issues

A real-time implementation of ASBC must be capable of high resolution at low frequencies to achieve the required filter specification for large changes of loudness. It must also either 1) Store a discrete set of pre-calculated compensation curves and associated filter coefficients, or 2) Calculate new curves and coefficients “on the fly” for continuous operation.

4.2.2. Current exemplary embodiment

The existing real-time implementation (as used in the listening tests) contains a discrete set of filters designed to perform increasing reductions in loudness in intervals of 1 phon. The filters are 1024th order FIR filters, since any lower order could not fully represent the low-frequency slope required for high levels of loudness reduction. The filters are convolved with incoming audio, producing the processed output. Convolution is performed in the frequency domain using the overlap-add method and vector operations in the Accelerate framework [8], to achieve high levels of CPU and power efficiency.

Use of such a high-order can be computationally expensive, especially for platforms without the vector operations in the Accelerate framework. There is also a high latency introduced (1.5x the length of the filter, in this case 1536 samples).

4.3. Practical Considerations

4.3.1. Relationship with Volume Control

A typical volume control scales an input signal between maximum volume and total attenuation of a signal. Such a process will often employ a logarithmic gain function, to mimic the logarithmic loudness response of the human auditory system. In a scenario where ASBC replaces the volume control, the mapping between loudness change and volume control position must be examined. Specifically, it must be determined at what volume level a target platform considers “flat” with no loudness change. For standard volume controls, this point is 0dB, or maximum volume, so it stands to reason that ASBC should function in the same way. This may vary with the target platform, and the intended use of a particular implementation.

4.3.2. Limitations of ISO226

The ISO standard equal-loudness-contour data is only valid up to 12.5kHz, leaving the 12.5kHz to 20kHz range of human hearing unaccounted for. We are unable to identify literature that can reliably define the loudness growth function at these frequencies; therefore ASBC continues to apply the 12.5kHz modification to all frequencies above 12.5kHz. Due to the logarithmic nature of pitch perception, this range only accounts for about half an octave, or approximately 6%, of the frequency range of human hearing [9].

4.3.3. Sensitivity Approximation

Although ASBC purports to apply to any listening system sensitivity, such an assumption is only valid within limits. For example, a secondary gain stage such as a preamp between ASBC and the listening system could force a user to reduce the signal level of a playback device and unknowingly apply ASBC spectral balance compensation, resulting in an inappropriately high level of spectral balance compensation for a “normal” eventual sound pressure level.

4.3.4. Fade-To-Silence

A volume control with a logarithmic gain function must eventually perform a step from passing audio to complete silence. Generally, this is done after the audio has been attenuated beyond the threshold of hearing, so there is no audible “disappearance” of the audio.

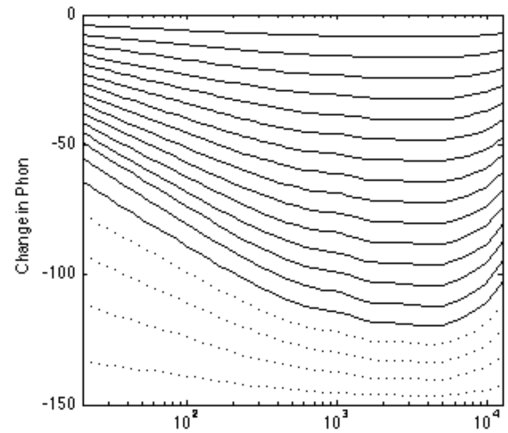

Figure 10 A transition between a compensated signal and a flat frequency response *(Solid Line: Compensated Signal, Dotted Line: Transition to flat frequency response)*

Inherent in ASBC is a large low-frequency boost, which implies that a much higher attenuation would be required for the entire spectrum to be below the threshold of hearing. Either 1) the range of the volume control could be extended to be operational to the point where the low frequencies are also below the threshold of hearing, or 2) a transition between a compensated signal and a flat frequency response could be implemented around the threshold of hearing, to ensure all frequencies are below the threshold of hearing before switching to complete silence.

4.4. Summary and Future Work

We have presented a method of preserving approximate spectral balance during loudness modification in an uncalibrated listening environment, including the calculation of the frequency-dependent constant of proportionality $k(f)$. Proposed applications include portable music players, live mixing, and use as a mastering tool. Future work includes fulfilling the requirement for an efficient, low-latency platform-independent implementation, and specialisation of the model for specific listening environments.

4.5. Acknowledgements

The authors would like to thank the staff and listening test participants at Birmingham City University, Raymarine Ltd and Japanese Knotweed Ltd, as well as family and friends for contributions to the project.

4.6. References

[1] Hawker, O. et al., ASBC source code repository, *https://github.com/ohawker/ASBC* (2015)

[2] A. Seefeldt, “Loudness Domain Signal Processing”, *Audio Engineering Soc. Convention 123,* (2007 Oct)

[3] H. Fletcher, M. Munson, “Loudness, Its Definition, Measurement and Calculation”, *Bell System Technical Journal,* Volume V, (1933 Aug)

[4] International Organization for Standardization, “Normal equal-loudness-level contours”, *ISO226:2003*, (2003 Aug)

[5] THX Ltd., “Tone Balance Volume Control”, *US Patent 8,509,457*, (2013 Aug)

[6] M. Florentine et al., “Loudness”, *Springer Handbook of Auditory Research, Volume 37*, (2011)

[7] SCENIHR, “Potential health risks of exposure to noise from personal music players and mobile phones including a music playing function”, *European commission*, (2008)

[8] Apple Inc., “Accelerate Framework Reference”, *https://developer.apple.com/library/mac/documentation/Accelerate/Reference/AccelerateFWRef/*, (2011)

[9] Klapuri, A. Davy, M. “Signal Processing Methods for Music Transcription”, *Springer*, (2006)