Wankling, Matthew and Fazenda, Bruno

The optimization of modal spacing within small rooms

THE OPTIMIZATION OF MODAL SPACING WITHIN SMALL ROOMS

M. Wankling[1], B. Fazenda
[1] University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK

ABSTRACT

The reproduction of audio in small rooms is significantly affected by room modes – resonant frequencies set up as a result of the geometry of the room. These frequency specific modes combine to form a unique frequency response for that room. This frequency response becomes part of the overall reproduction system, and as such, a poor response can have a detrimental effect on the listener's enjoyment of the audio. Researchers have, therefore, consistently looked at how these modes affect our perception, and in particular, attempted to optimize the distribution of the modal frequencies so as to reduce the negative effects. A study is presented which determines an optimal frequency spacing between two adjacent modes which relates to the optimal listening scenario. This result is then discussed in the context of an objective measure which predicts a similar spacing and some remarks on how these results can be incorporated into optimization techniques.

Keywords modal spacing, room acoustics, low frequency

1 INTRODUCTION

The problem of resonant modes in listening spaces has long been acknowledged. Reducing the negative perceptual effects of these modes is fundamental both to room designers aiming for the highest quality of audio reproduction and loudspeaker manufacturers aware that this is one aspect that can severely affect the perceived quality of their product. Due to the relationship of these modes with the physical dimensions of the room, researchers have often looked at optimal room aspect ratios in an attempt to avoid modal degeneracy – multiple modes overlapping at the same frequency. Work of this nature has often concentrated on attempts to control the distribution of all possible modes in a given room [1,2]. More recently, the particular response dependent on source and receiver position has been acknowledged as more representative of the general use of such rooms [3,4]. In any case, the frequency spacing between adjacent modes has been fundamental for all studies of the low frequency modal behavior of these spaces.

Modal spacing has often been used as an objective measure to quantify the quality of reproduction in a listening space. It has been theorized that an increase in room acoustic quality is associated with a greater uniformity of spacing in frequency between adjacent modes. Optimal room ratios such as those published by Louden [1], attempt to optimize this spacing. More recent work by both Cox [4] and Fazenda [3] has also focused on the subject of optimal room ratios and considered objective metrics by which it may be possible to classify the room response.

When considering the effects of modal distribution on the sound quality of a room, it is generally accepted that a flat frequency response is desirable. The presence of peaks and dips modify the overall sound for the listener by altering the amplitude at certain frequencies. Furthermore, the Q-factors (defined as the centre frequency divided by bandwidth) of these peaks and dips are also associated with frequency dependant decay times. In general, the more homogeneous frequency responses (flat), corresponding to lower Q-Factors, are associated with shorter time responses. It follows that if modal frequencies can be arranged to form a more homogeneous response, shorter decay times can be achieved, resulting in improved audio reproduction quality.

This paper examines whether an optimum subjective spacing between resonances can be defined which is associated with the shortest decay time of the system and hence the best perceptual condition. If available, this metric could in turn be incorporated into room design at low frequencies. Conclusions are drawn with reference to the Modulation Transfer Function (MTF), an objective measure which predicts an optimal spacing numerically. Finally, specific ways in which results obtained may be used to optimize rooms for better audio reproduction are discussed and a number of ideas are presented which highlight future experimental work.

3 TEST METHOD

The modal response of a real room is incredibly complex. Therefore, to test for an optimal modal spacing, a simplified case of two single resonances was considered. These two resonances were artificially modeled using an implementation of the complex pressure equation, published by Bolt [5] (equation 1), which has previously been used to successfully model low frequency room responses [3,4,6]. The resulting frequency response was then transformed to the time domain, giving the impulse response of the room in question.

$$P_{\omega}(r) = j\omega\rho Q c^2 \sum_{n} \frac{P_n(r) P_n(r_0)}{X_n(\omega^2 - \omega^2_n - 2j\delta_n\omega_n)} \quad (1)$$

$P_n r$ *and* $P_n(r_0)$ represent modal 'shape functions' dependant on the boundary conditions and source and receiver positions. In this simplified case these were assumed to be one, and the modal array (ω_n) was limited to two frequencies, those of the two spaced resonances.

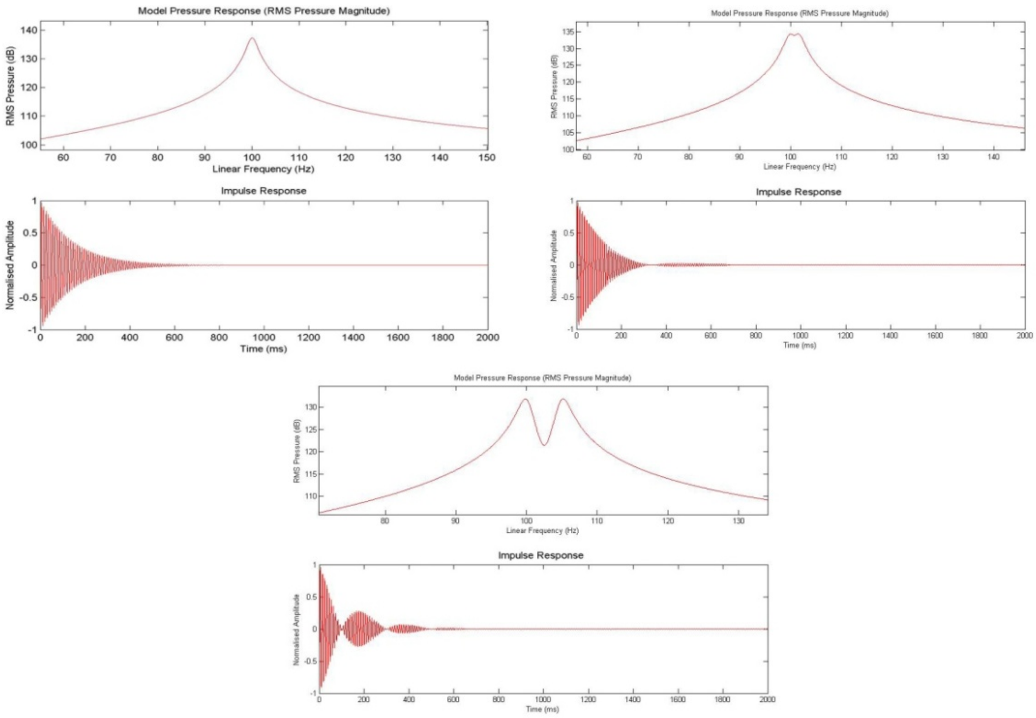

Figure 1: a) 100Hz & 100.1Hz b) 100Hz & 101.5Hz c) 100Hz & 105Hz

Figure 1 represents three scenarios where the response of a system comprises two spaced resonances. A simple visual investigation of the effect of altering the spacing between them reveals a clear reduction in decay time. However, as the second frequency moves away from the first, the magnitude response reveals a large dip and the resulting impulse response begins to show a distinctive amplitude modulation. This is obviously associated with the interaction between the two resonances and at these frequency differences they sound identical to 1st order beats as described in many psychoacoustic textbooks e.g. [7]. When plotted as a logarithmic decay (Figure 2) the beating effects are even clearer.

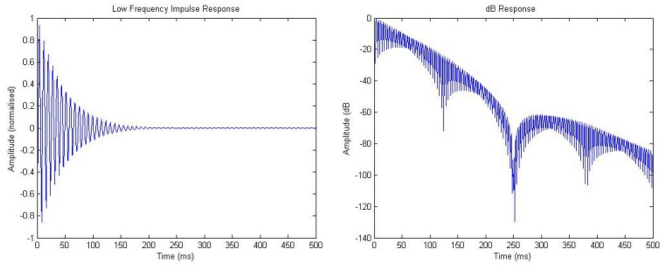

Figure 2: The computed response displayed as a normalized impulse and also in dB

One can make assumptions based upon this visual inspection as to the perceived quality of an audio stimulus when passed through these resonant systems (assuming the audio material were to excite the corresponding frequency range). The shortest decay is clearly preferable, while the introduction of beats will be highly detectable to the listener and perhaps undesirable. The question however remains; at what point along this sliding spacing scale does the optimal compromise between the two degrading effects lie?

To obtain this result, a computer program was written to generate impulses according to the equation described above. Whilst these impulses could be convolved with an input stimulus such as a test tone or musical refrain, to create an auralization of that stimulus in the modeled room, it was decided that the impulse itself should be used as the test stimulus since its effects are more audible and correspond to 'the worst case scenario'.

As previously mentioned, the Q factor of a resonance affects its decay time. Therefore the test was performed using a number of different Q factors (10, 20, 30, and 50) as well as three frequencies at which the first resonance would be placed (63, 125, 250Hz), chosen to represent a broad range typical in listening conditions. The frequency of the second resonance was adjusted by way of a slider on a graphical user interface. Impulses were generated on the fly by the software model each time the slider was moved. During each test, subjects were asked to adjust the slider to the point where the overall decay sounded the shortest. Prior to the test, explanation of the differences in presentation sounds (long decay, shorter decay, and beating effect) were explained, along with images in the time domain. It was also explained that beats were to be considered as part of the overall decay process. No time domain images were displayed during the actual tests to avoid bias.

Eleven subjects were tested, in quiet studio conditions, with samples auditioned over a pair of Sennheiser HD-650 headphones. Each subject was given time to practice before the test commenced. The presentation levels of the three frequencies were weighted to ensure that the perceived level of each sample was the same - samples were presented according to the 90dB equal loudness contour [8].

4 EXPERIMENTAL RESULTS

Results are shown and statistical analysis has been carried out to show the significance of each result.

Figure 3 shows the mean spacing identified by the 11 subjects. A simple visual inspection reveals clear trends. As the Q factor increases, the optimal spacing needed to provide the shortest decay reduces. This is to be expected – as Q increases, the resonant peaks become sharper and a greater definition between individual frequencies is detectable. They must therefore be closer together to 'flatten' the overall response. When comparing the test frequencies, Figure 3 clearly shows that higher frequencies require a greater spacing. It should be noted here that this is in direct contradiction to the natural decrease of modal spacing in rooms as frequency increases. Furthermore, the level of uncertainty, shown by the standard deviation error bars also increases with frequency indicating that an optimal spacing becomes less meaningful as frequency increases.

Analysis of variance was carried out to ascertain the level of significance across the variable parameters. Table 1 shows that both the Q factor and modal frequency are highly significant, i.e. $p<0.01$, which indicates the success of systematic testing.

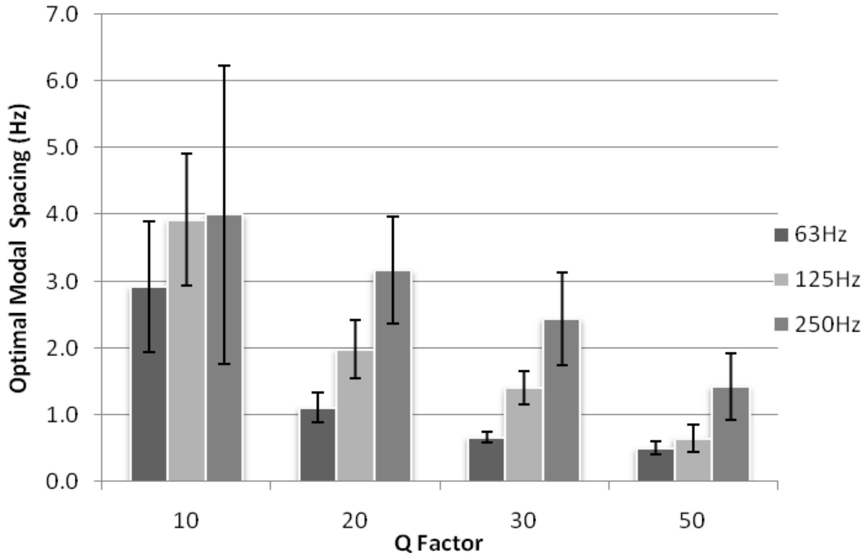

Figure 3:Mean Spacing across Q Factor and Frequency

Experimental Factor	***p***
Q	0.00
Frequency	0.00

Table 1: Anova Test

Although both factors are highly significant, it is useful at this point to wrap them into a single factor - that of modal bandwidth. Frequency, Q and bandwidth are related according to the equation:

$$Bw = \frac{f}{Q} \qquad (2)$$

Table 2 considers each of the 12 test scenarios in ascending bandwidth. The results again show a clear trend of increasing spacing (mean) with bandwidth.

BW	1.26	2.10	2.50	3.15	4.17	5.00	6.25	6.30	8.33	12.50	12.50	25.00
Q	50	30	50	20	30	50	20	10	30	20	10	10
Freq	63	63	125	63	125	250	125	63	250	250	125	250
Mean	0.5036	0.6643	0.6458	1.1079	1.4075	1.4284	1.9860	2.9183	2.4411	3.1664	3.9237	4.0013
St.Dev	0.0959	0.0866	0.1998	0.2220	0.2512	0.5007	0.4355	0.9729	0.6961	0.8013	0.9843	2.2361

Table 2: Mean Subjective Optimal Spacing presented in ascending Bandwidth

Figure 4 shows optimal spacing as a percentage of the modal bandwidth. This figure reveals that, for Q's of 20, 30 and 50, regardless of frequency or Q, the optimal spacing lies between 25 and 40%. At lower Q's, the standard deviation becomes higher (see Table 3) and results are less reliable. These results were confirmed by comments from subjects who each stated that the shortest impulses were significantly harder to judge than those of longer length.

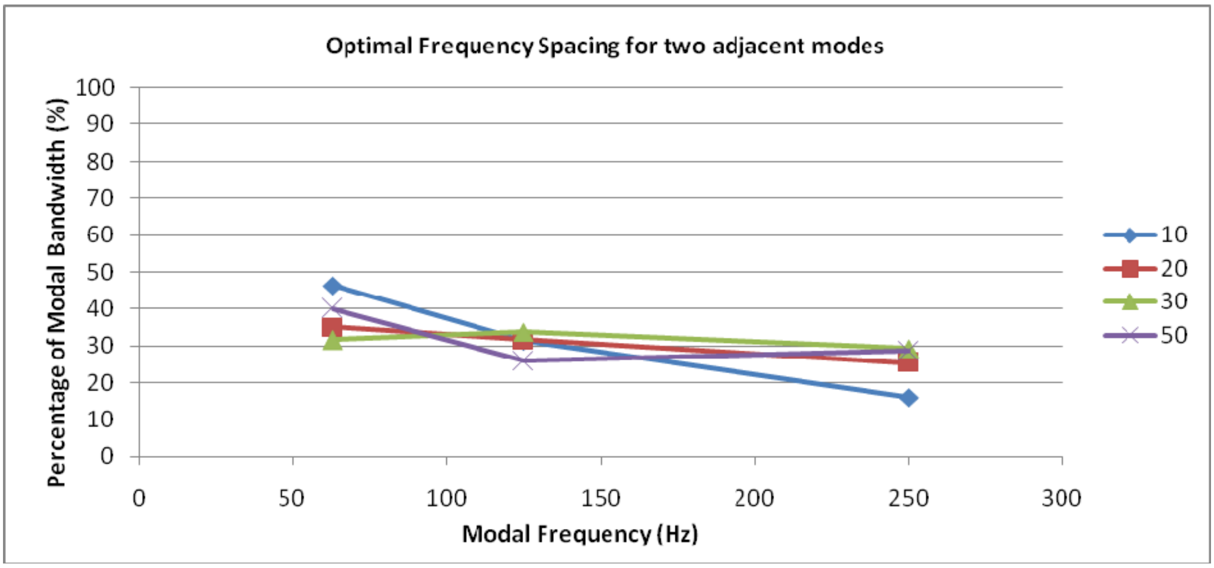

Figure 4: Optimal Spacing across ascending bandwidth for the four different Q Factors tested

We can see that there is a clear indication that the optimal spacing and therefore audio reproduction quality may be related to a percentage of the modal bandwidth. Optimization of rooms in this frequency range may be based upon this percentage, although care must be taken in doing so. One reason for this is that this test only takes two resonances into account. A real room's response becomes more complex with many resonances interacting not only due to frequency but also phase. Without such a simplified system of two carefully spaced modes of identical amplitude and phase, a simple examination of the time domain response becomes increasingly difficult. Thus, a computational method for predicting the same result is desirable. One such method which may be useful is the Modulation Transfer Function.

.

5 THE MODULATION TRANSFER FUNCTION

The Modulation Transfer Function (MTF), originally developed in the field of optics as a quantifier of lens image resolution, has also been shown to correlate well with audio reproduction quality [9-11]. It measures the system's ability to preserve amplitude modulations of a signal over a set frequency range. The modulation frequencies are defined as representative of audio signals and in particular those found in speech where this technique is applied to define a speech transmission index. The function takes the input response of the system and calculates a figure of merit between 0 and 1 with the top of the scale corresponding to an exact copy of the input signal.

The same variables, frequency and Q, were examined. Measurements were carried out at the same three test frequencies, 63, 125 and 250Hz and Q's of 10, 20, 30, 40 and 50. Figure 5 shows an example of the MTF mapping across a range of modal spacing for a number of modal Q-factor values.

Table 3 shows the optimal spacing, that is, the spacing at which the highest figure of merit is obtained, as calculated by the MTF metric at each frequency and for increasing values of Q-factor.

Frequency (Hz)	Q=10	Q=20	Q=30	Q=40	Q=50
63	8.5	5.3	4.1	3.5	3.3
125	12.6	8.4	6.5	5.3	4.6
250	21.6	12.6	9.9	8.4	7.4

Table 3: Optimal Spacing as Predicted by MTF (Hz)

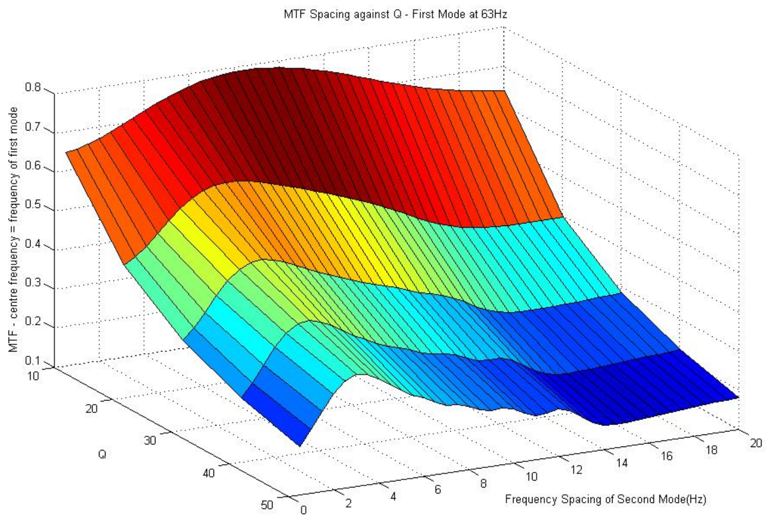

Figure 5: Example of MTF scores across spacing at different Q's - frequency of first resonance 63Hz

Comparison between subjective test results, in Table 2, and those predicted by the MTF, reveals that although they differ significantly in value, the same trend is clearly apparent – an increasing optimal spacing with frequency and decreasing optimal spacing with Q. Therefore it would seem that an adjustment of the MTF metric, or indeed, a metric with better correlation to perception could accurately predict the subjective optimal spacing between the two resonances. Refinements to the metric may well achieve this in the future. Any objective metric should ideally be able to predict subjective ratings when using modal scenarios of greater complexity than the two single resonances used in this test. This would further increase the potential for optimization.

6 OPTIMIZATION

These results pose intriguing questions in the field of room optimization. Whilst the MTF in its current state cannot be used as an accurate subjective measure of a room, the listening test results may be taken more readily. The optimal modal spacing as a function of modal bandwidth is of particular interest. With this ideal scenario lying between 25 and 40%, this figure could potentially be averaged and approximated at 33%. This single optimal figure then opens up a number of potential optimization techniques.

Firstly, careful design of room dimensions may lead to modes with a closer relationship to this ideal spacing. Although it is unlikely that this will prove consistent across the frequency range, this research has also highlighted the fact that it is only the lowest modes which can be said to have an optimal spacing with any confidence. Therefore if this type of optimization was carried out only in the lowest frequency region, or at least weighted towards it, a subjective improvement could be achieved.

Furthermore, we can see that the Q factor has an important part to play. If a listening space has already been constructed, with modes set at specific frequencies due to the geometry of the room, a computer algorithm may be able to predict the best Q factor for the current spacing and room absorption modified to produce these Q's, thereby improving sound reproduction quality.

Finally, the active addition of modal frequencies, either through the reproducing system or a separate installation, becomes an attractive option as specific frequencies based on these results could be added. If correct modal additions do indeed aid the reduction of decay times, it may well be necessary to add only a few, based around the most prominent modes at the lowest frequencies (see Fig. 6), reducing the decay to a level just below the threshold of human detection. This decay threshold is to be investigated as part of the ongoing research project.

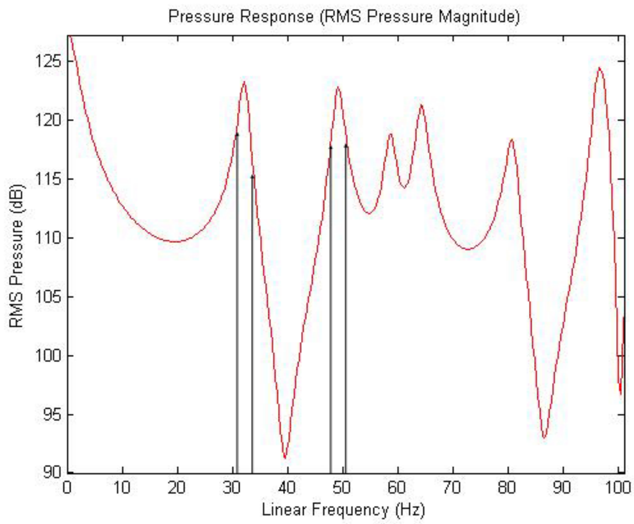

Figure 6: An original room response with prominent modes at around 32Hz and 50Hz. The black lines represent modes to be added artificially, at spacing's based upon the subjective results, with the aim of reducing the decay to a level just below the threshold of detection.

7 CONCLUSION

In this investigation, it is clear that, when using a simplified scenario of two single resonances, the decay time imposed by the response of the system can be optimized by an ideal spacing of their centre frequencies. As the bandwidth of each resonance increases, so does the optimal spacing. When specified in terms of percentage of modal bandwidth, the optimal spacing lies between 25% and 40%, regardless of frequency and Q (with exception to a Q value of 10). At the present time, these results cannot be replicated using an objective measure such as the MTF.

A smaller spacing than optimal leads to longer but homogenous resonant decays. This has been shown to be problematic for sound reproduction [3,13]. However, with larger spacing than optimal, the two peaks begin to separate, leading away from a flat frequency response, and beats become identifiable. The relative importance of these two factors (long single decays vs. perception of beats) has not been measured and it stands out as an interesting avenue for future research.

The subjective results reveal that at these low frequencies, a much closer spacing is needed than is usually achieved by room design. The reliability of subjects responses also show that modal spacing is important at the lowest modes but its significance decreases with increasing frequency, giving weight to the argument that it is at these lowest frequencies that modal optimization should be focused. At 250Hz, the differences in spacing were very difficult to perceive. Furthermore, at the lowest tested Q value of 10, spacing differences were also difficult to perceive. This result is in agreement with previous research which suggests a threshold for detection of changes in modal Q-factor at around Q=16 [12].

Finally, these results open up further research avenues. For example, will the masking effects of a musical stimulus cause a difference in result, or will the same detection of the shortest decay and onset of beats remain? Further work currently being undertaken also looks at the effects of multiple modes rather than the simple pair used in this test.

8 REFERENCES

[1] M.M. Louden, "Dimension-Ratios of Rectangular Rooms with Good Distribution of Eigentones," *Acustica*, vol. 24, 1971, pp. 101-04.

[2] O.J. Bonello, "A New Criterion for the Distribution of Normal Room Modes," *J. Audio Eng. Soc*, vol. 19, 1981, pp. 597-606.

[3] B. Fazenda, M.R. Avis, and W.J. Davies, "Perception of Modal Distribution Metrics in Critical Listening Spaces-Dependence on Room Aspect Ratios," *J. Audio Eng. Soc*, vol. 53, Dec. 2005, pp. 1128-1141.

[4] T.J. Cox, P. D'Antonio, and M.R. Avis, "Room Sizing and Optimization at Low Frequencies," *J. Audio Eng. Soc*, vol. 52, Jun. 2004, pp. 640-651.

[5] P.M.C. Morse, *Vibration and sound*, McGraw-Hill New York, 1948.

[6] N. Stefanakis, J. Sarris, and G. Cambourakis, "Source Placement for Equalization in Small Enclosures," *J. Audio Eng. Soc*, vol. 56, May. 2008, p. 357.

[7] B.C. Moore, *An Introduction to the Psychology of Hearing*, Academic Press Inc, 1997.

[8] D.W. Robinson and R.S. Dadson, "A re-determination of the equal-loudness relations for pure tones," *British Journal of Applied Physics*, vol. 7, 1956, pp. 166-181.

[9] T. Houtgast and H.J.M. Steeneken, "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria ," *J. Acoust. Soc. Am.*, vol. 77, Mar. 1985, pp. 1069-1077.

[10] L. Harris, K. Holland, and P. Newell, "Subjective assessment of the modulation transfer function as a means for quantifying low-frequency sound quality," *Proc. Inst. Acoust.*, vol. 28, 2006.

[11] B. Fazenda, K. Holland, and P. Newell, "Modulation transfer function as a measure of room low frequency performance," *Proc. Inst. Acoust. (UK)*, vol. 28, 2006.

[12] M. Avis, B.M. Fazenda, and W.J. Davies, "Thresholds of detection for changes to the Q factor of low-frequency modes in listening environments," *J. Audio Eng. Soc*, vol. 55, Aug. 2007, pp. 611-622.

[13] F.E. Toole, "Loudspeakers and Rooms for Sound Reproduction—A Scientific Review," *J. Audio Eng. Soc*, vol. 54, 2006, pp. 451-476.