MULTI-DOMAIN ANALYSES OF ACOUSTIC ROOM CORRECTION SYSTEMS

P Thakkar  GP Acoustics (UK), Maidstone, Kent, UK

1 INTRODUCTION

At low frequencies, where wavelengths of sound waves are comparable to the physical dimensions of the enclosing rooms, room modes contribute to several audible artefacts due to pressure irregularities and extended decay times at specific frequencies. While careful selection of room dimensions and layouts help minimise their audibility at some listening positions\(^1\)–\(^3\), multiple room modes may still be simultaneously excited to varying degrees. Common acoustic treatments, such as absorbers and diffusors, are often impractical in this frequency range due to their size and operating range limitations. As a result, acoustic technologies, such as room equalisation and room correction, are popular choices for mitigating the detrimental effects of low-frequency room modes in professional and domestic spaces alike.

In this paper, three objective performance metrics are derived to contrast the performance of the two types of technologies in rectangular rooms across the time, frequency, and spatial domains.

2 IDEAL LOW-FREQUENCY ROOM PERFORMANCE METRICS

2.1 Steady-state Performance and Steady-state Deviation (SSD)

Room modes manifest themselves as severe resonances (peaks) and anti-resonances (notches) in the in-room transfer functions of the source loudspeakers. The variations can be as much as 25dB, resulting in significant colouration perceived at the listening positions. The steady-state deviation metric (SSD) in equation 1 is the deviation of the RMS steady-state pressure response, over multiple listening seats, from a flat line coinciding with the mean pressure level within the analysis window of 20 – 100Hz – lower values indicate flatter responses in the range of interest. The assumption is that a flat line response with no peaks or dips represents a room transfer function with no room modes or associated audible artefacts.

\[
SSD = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} (H_{RMS}(n) - H_{average})^2 [dB]}
\]  

\[1\]
The RMS average SPL, $H_{RMS}$, is given by

$$H_{RMS}(n) = 20 \log_{10} \frac{\sqrt{\sum_{r=1}^{R} (|p_{r}(n)|)^2}}{2e^{-5}} [dB]$$

where $p_{r}(n)$ is the pressure response at the $r^{th}$ seating position and $n^{th}$ frequency bin and $R$ is the total number of seating positions.

The RMS average rather than the complex average is used, as the latter would imply superimposition of the room transfer functions, which is not an accurate representation of the perceived room acoustic behaviour.

Finally, the flat line average level is calculated as

$$H_{average} = \frac{1}{N} \sum_{n=1}^{N} H_{rms}(n)$$

### 2.2 Time Domain Performance and Modal Decay Time (MT60)

The occurrence of room modes causes energy to build up between reflecting surfaces. This phenomenon, combined with the ineffective damping of low-frequency sound energy offered by common room treatments and furnishings, causes the stored energy to decay at a significantly slower rate than nearby frequencies. Figure 2 shows an example of this behaviour.

In terms of subjective impact, the extended decay times alter the temporal response of a loudspeaker in a room. For instance, short bass notes, like a kick drum or a bass guitar, may ring on at one or more closely spaced mode frequencies. Furthermore, pitch shifting may occur when a slowly decaying mode replaces a rapidly decaying mode at a nearby frequency, causing the perceived pitch to be different from the original. Therefore, the ideal low-frequency transient response of a loudspeaker-room coupled system should feature a consistent decay across all frequencies in the modal region.

Modal decay time (MT60) is analogous to the reverberation time (T60) used to characterise diffuse sound fields. In the low-frequency region, energy decay is dominated by individual room modes around their centre frequencies. MT60 measures the time taken for a single mode's initial sound pressure level to decay by 60dB.

Figure 2. Waterfall plot of a subwoofer in-room. Note the prolonged ringing at mode frequencies
The typical approach for estimating MT60 involves calculating a Short-Time Fourier Transform (STFT) of the room impulse response to obtain exponential decay envelopes at the individual mode frequencies. Schroeder-integrated energy-time curves are then obtained from the squared decay envelopes. Finally, the MT60 times are derived from the slope of the energy-time decay curves.

Schroeder backward integration for a time response \( h(t) \) is given by

\[
E(t) = \int_0^\infty h^2(t) \, dt
\]

where \( h(t) \) is the exponential decay envelope.

The energy-time decay curve \( ETC(t) \) for \( E(t) \) is given by

\[
ETC(t) = 10 \log E(t)
\]

The slope or the decay rate \( m \) of the energy-time curve \( ETC(t) \) is obtained from a least-squares line fit (linear regression) given by

\[
\min_{m,c} \int_{t_1}^{t_2} \{ETC(t) - [mt + c]\}^2 \, dt
\]

where \( m \) is the optimal decay rate, \( c \) is the y-intercept of the linear regression line and \([t_1, t_2]\) is the numerically determined fitting interval.

Finally, the MT60 is calculated as

\[
MT60 = -\frac{60}{m} \, [s]
\]

Several practical limitations of the line fitting approach, described in equation 6, arise from the fitting interval \([t_1, t_2]\). In practice, the regression line should only be fitted to the exponential decay portion (straight line on a log scale) of the energy-time envelope. Therefore, the initial onset and noise floor bias at the beginning and the end of the envelope should be excluded from the fitting interval, respectively. However, in cases where the useful dynamic range of the measured data is severely limited, alternative fitting approaches such as non-linear decay plus noise models or a dedicated noise compensation stage before ETC calculation may be more appropriate\[4\]–\[6\]. Additionally, the stationary nature of the FFT calculation within individual STFT analysis windows may distort the true decay envelope of the mode peaks to some extent\[7\]. Nevertheless, the proposed method is sufficiently reliable for relative comparisons, such as the ones presented in this paper.

### 2.3 Spatial Domain Performance and Mean Seat-to-seat Variation (MSSV)

Room mode transfer functions are highly variable with source and listener positions throughout the rooms. The variations may be significant with as little as small head movements. Whilst it is reasonable to assume that the source positions may remain fixed in a listening setup, no such assumptions can be made about listener positions. For instance, in a group listening environment such as a home cinema or listening room, listeners less than a metre away from each other may experience vastly different low-frequency performances from the same source loudspeakers. Therefore, to ensure all listeners within a specific listening area perceive the same low-frequency performance, seat-to-seat pressure variations should be minimised as much as possible.

Mean seat-to-seat variation (MSSV) is the average variance of the SPLs observed across a fixed number of seating positions within a listening area, given by
\[ MSSV = \sum_{r=1}^{R} \sum_{n=1}^{N} \frac{(|H(r, n)| - H_{average}(n))^2}{R(N-1)} \text{ [dB]} \] \[ [8] \]

where \( |H(r, n)| \) is the SPL at the \( r \)th seat at frequency bin \( n \). \( H_{average}(n) \) is the average SPL across all \( R \) receivers at frequency bin \( n \) and \( N \) and \( R \) are the total number of frequency bins and receivers, respectively. Note the difference with equation 1, where \( H_{average} \) is calculated over the entire RMS pressure response as opposed to an amplitude average per frequency bin for \( H_{average}(n) \). This metric is also known as the mean spatial variation (MSV)\([8]\).

3 ACOUSTIC ROOM CORRECTION SYSTEMS

3.1 Digital Room Equalisation

For room equalisation systems, the correction approach involves pre-processing the excitation signal to reduce or limit the influence of the physical room acoustics on the direct sound field reaching the listeners. However, due to the high variability of room transfer functions with source and receiver positions, room equalisation systems typically only aim to improve the system performance over a pre-determined listening area. Examples include the parametric equalisation of individual room modes \([4],[5]\) and simple adjustments such as tone and low-frequency tilt controls found on amplifiers and AV receivers.

Modal equalisation is a form of parametric equalisation that operates on individual room modes by pre-filtering the primary signal to reduce the net excitation and increase the consequent decay rate corresponding to each of them. A modal resonance can be represented as a low pass filter in the s-domain, given by

\[ H_{fn}(s) = \frac{1}{1 + \frac{s}{Q_{fn}} + s^2} \] \[ [9] \]

where \( Q_{fn} \) is the unequalised Q factor of the modal resonance and \( s = j \omega \). \( w_{fn} \) is the angular mode frequency \( w_{fn} = 2 \pi f_{n} \).

\( Q_{fn} \) is inversely proportional to the exponential modal decay rate \( m \) in equations 6 and 7. Therefore, by reducing \( Q_{fn} \), a room mode at frequency \( f_{n} \) can be proportionately controlled. The general form of a modal equaliser is given by

\[ H_{equaliser}(s) = \frac{1 + \frac{s}{Q_{fn}} + s^2}{1 + \frac{s}{Q_{target}} + s^2} \] \[ [10] \]

where \( Q_{target} \) is the target Q factor of the mode after equalisation.

In normal use, \( Q_{target} < Q_{fn} \) such that the modal decay rate is increased after equalisation. As such, the equaliser described by equation 10 is a notch filter with peak attenuation \( K \), given by

\[ K = 20 \log_{10} \left( \frac{Q_{target}}{Q_{fn}} \right) \text{ dB} \] \[ [11] \]
Figure 3 illustrates a simple case of a modal equaliser operating on a single mode peak at 100Hz. The ratio of the equalised to unequalised Q factors is 0.3, corresponding to a peak attenuation of -10dB and a 70% reduction of the modal decay time.

This simple scheme can be extended to equalise multiple mode peaks by cascading multiple independent notch filters. Physically, modal equalisation is approximately analogous to adding virtual resonant absorbers to attenuate targeted mode peaks to a fixed degree. Pressure nulls in the room transfer functions are left untreated under this equalisation scheme.

3.2 Physical Room Correction

The room correction approach concerns the entire acoustic environment instead of a targeted listening area. The goal for these systems is to eliminate room modes by complete suppression or cancellation such that they no longer impact the direct sound field produced by the primary systems. Modal suppression and cancellation are achieved by deploying multiple carefully positioned sound sources throughout the acoustic environment[8],[11].

At mode frequencies, the physical distances travelled by the incident waves before reflection are multiples of the half wavelengths resulting in constructive summation and a pattern of pressure nodes and antinodes along the related rectangular room dimensions. By placing sources at pressure nodes of a particular mode, energy transfer and, therefore the net excitation of that mode, are minimised. Naturally, higher-order modes sharing the same nodal line are suppressed simultaneously. Similarly, by placing two identical sources symmetrically about the nodal line, net zero excitations at that mode and related multiples of higher-order modes are achieved due to cancellation. Figure 4 illustrates this concept inside an ideal one-dimensional duct.
In practical application, careful placement schemes featuring two or four identical and identically driven sources can be implemented to achieve meaningful broadband control below 100Hz, as shown in Figure 5 [8], [12].

![Source placement schemes](image.png)

Figure 5. Source placement schemes for modal control. All sources are assumed to be identical and identically driven.

# 4 MODAL DECOMPOSITION ROOM MODEL AND CONTROL CASE

## 4.1 Low-frequency modal decomposition model

A modal decomposition model was chosen over its counterpart, the image source model, primarily for its computation speed. In terms of accuracy, both are comparable up to 125Hz, where the modal density is still relatively low, beyond which both differ significantly from real in-room responses [2].

At low frequencies, the in-room loudspeaker behaviour can be modelled as a linear combination of the room resonances. Equation 12 describes the pressure sound field in a rectangular enclosure with rigid boundaries, a full derivation of this expression is given in Nelson and Elliot [13].

\[
p(\vec{r}) = \sum_{m=0}^{\infty} \frac{\rho_0 \Psi_m(\vec{r})}{V[k_m^2 - k^2]} \sum_{n=1}^{N} \Psi_m(\vec{s}) U_n
\]

[12]

Where \( p \), the complex pressure amplitude at the receiver position \( \vec{r} \) within the enclosure volume \( V \), is given by a linear combination of the enclosure eigenfunctions \( \Psi_m(\vec{r}) \). \( \omega \) is the angular frequency of excitation, \( \rho_0 \) is the ambient fluid density, \( k \) is the wavenumber (given by \( \frac{\omega}{c_0} \) where \( c_0 \) is the speed of sound) and \( k_m \) is the \( m^{th} \) eigenfrequency expressed as a wavenumber. The magnitude of excitation of each eigenfunction is determined by the summation on the right of the expression, where \( U_n \) describes the volume velocity of the source \( n \) and \( \Psi_m(\vec{s}) \) is the value of the \( m^{th} \) eigenfunction at source position \( \vec{s} \). Note that for this cuboid case, \( m \) is short for \( n_x n_y n_z \). The sources were assumed to be ideal monopoles with a uniform free-field pressure sensitivity of 90dB across all frequencies.

The eigenfunction \( \Psi_m \) for a cuboid room is given by

\[
\Psi_m = \sqrt{\epsilon_n} \cos \left( \frac{n_x \pi x}{L_x} \right) \cos \left( \frac{n_y \pi y}{L_y} \right) \cos \left( \frac{n_z \pi z}{L_z} \right)
\]

[13]

The Eigen wavenumbers \( k_m \) and scaling factors \( \epsilon_n \) are found by equations 14 and 15, respectively.
\[ k_m = \sqrt{\left(\frac{n_x \pi}{L_x}\right)^2 + \left(\frac{n_y \pi}{L_y}\right)^2 + \left(\frac{n_z \pi}{L_z}\right)^2} \]  

[14]

The scaling factors \( \epsilon_n \) ensure that the volume integral of each eigenfunction is normalised to the room volume. These are

\[ \epsilon_{n_x} \epsilon_{n_y} \epsilon_{n_z} = \begin{cases} 1, & n = 0 \\ 2, & n \geq 1 \end{cases} \]  

[15]

In addition, real room boundaries have some absorption or damping characteristics. To include this and to ensure that the solution for pressure is not infinite at mode frequencies, equation 12 needs to be adjusted as shown in equation 16.

\[ p(\vec{r}) = \sum_{m=0}^{\infty} \rho_0 \Psi_m(\vec{r}) \sqrt{k_m^2 - k^2 + \frac{Q_m k}{Q_m}} \sum_{n=1}^{N} \Psi_m(\vec{s}) U_n \]  

[16]

Where \( Q_m \) represents the Q factor of the \( m^{th} \) eigenfunction or mode and \( j \) is the complex operator \( \sqrt{-1} \). For lightly damped enclosures, \( Q_m \) is relatively large. For the case studies presented in this paper, this was set to 20 for all modes up to 100Hz. The frequency limits for the numerical room model and calculation of the metrics described in section 2 are 20-100Hz.

### 4.2 Control case

This paper analyses room correction systems for stereo performance in a lightly damped rectangular room of dimensions 7.0 x 5.0 x 2.7m. The primary stereo pair are positioned at 1.0 x 1.0 x 1.2 m and 1.0 x 4.0 x 1.2 m as the L and R channels, respectively. A targeted listening area of 6 seats is centred around the typical stereo listener position derived from an equilateral triangle arrangement, as shown in Figure 6. The seated ear height is assumed to be 1.2m, as per BS1116-3[14]. The sources and listening seats are positioned to represent a typical home cinema setup with two rows of seating.

The RMS average of the pressure responses obtained using equation 16 and the corresponding aggregate time-frequency behaviour across all six seats are shown in Figures 7 and 8, respectively. The control room acoustic parameters and the corresponding performance metrics are shown in Table 1.
5 RESULTS

5.1 Digital Room Equalisation

A modal equaliser was set up to attenuate all modes in Table 1 by 6dB, which corresponds to halving their respective pressure amplitudes. Six independent notch filters, given by equation 10, were cascaded before applying the result to the excitation signals for the primary L and R channels. The results are summarised in Table 2 and Figure 9.

Table 1. Control room acoustic parameters and performance metrics

<table>
<thead>
<tr>
<th>Mode frequency (Hz)</th>
<th>Mode order (L W H)</th>
<th>MT60 (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.5</td>
<td>1 0 0</td>
<td>1.91</td>
</tr>
<tr>
<td>49.0</td>
<td>2 0 0</td>
<td>0.96</td>
</tr>
<tr>
<td>68.1</td>
<td>1 0 1</td>
<td>0.69</td>
</tr>
<tr>
<td>68.6</td>
<td>0 2 0</td>
<td>0.71</td>
</tr>
<tr>
<td>72.8</td>
<td>1 2 0</td>
<td>0.82</td>
</tr>
<tr>
<td>84.3</td>
<td>2 2 0</td>
<td>0.56</td>
</tr>
<tr>
<td>SSD (dB)</td>
<td></td>
<td>7.32</td>
</tr>
<tr>
<td>MSSV (dB²)</td>
<td></td>
<td>14.27</td>
</tr>
</tbody>
</table>

Clear improvements in modal decay times are achieved across all modes. However, the steady-state and spatial performances are unaffected. This result is expected as equalisation of mode peaks alone is insufficient to achieve a flat in-room response where pressure nulls below 100Hz still dominate the flatness metric. For this reason, commercial modal equalisation systems often opt for an additional inverse-filtering based magnitude equalisation stage to improve the overall flatness (SSD) of the equalised responses at the selected seats. As for the spatial performance or MSSV, equalising the excitation signal reduces the net excitation at the targeted modes, however, this reduction is uniform across all seats and consequently, does not improve the underlying seat-to-seat variation.

Table 2. Control room acoustic parameters and performance metrics after room equalisation

<table>
<thead>
<tr>
<th>Mode frequency (Hz)</th>
<th>Mode order (L W H)</th>
<th>MT60 (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.5</td>
<td>1 0 0</td>
<td>0.96</td>
</tr>
<tr>
<td>49.0</td>
<td>2 0 0</td>
<td>0.48</td>
</tr>
<tr>
<td>68.1</td>
<td>1 0 1</td>
<td>0.54</td>
</tr>
<tr>
<td>68.6</td>
<td>0 2 0</td>
<td>0.53</td>
</tr>
<tr>
<td>72.8</td>
<td>1 2 0</td>
<td>0.63</td>
</tr>
<tr>
<td>84.3</td>
<td>2 2 0</td>
<td>0.23</td>
</tr>
<tr>
<td>SSD (dB)</td>
<td></td>
<td>7.41</td>
</tr>
<tr>
<td>MSSV (dB²)</td>
<td></td>
<td>14.27</td>
</tr>
</tbody>
</table>

Figure 8. Spectrogram plot for the RMS pressure average in Fig 7 (Magnitude normalised and IR zero padded by 0.15s)

Figure 9. Spectrogram plot for the RMS pressure average after digital room equalisation (Magnitude normalised and IR zero padded by 0.15s)
5.2 Physical Room Correction

The room correction system was based on the four midpoints scheme illustrated in Figure 5. Four identical and identically driven monopole sources were positioned halfway up the height and at the midpoints of the four enclosing room boundaries. The primary L/R channels were high-pass filtered to form a 2.4 channel bass-managed system. The results are summarised in Table 3 and Figure 10.

<table>
<thead>
<tr>
<th>Mode frequency (Hz)</th>
<th>Mode order (L WH)</th>
<th>MT60 (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.5</td>
<td>1 0 0</td>
<td>0.12</td>
</tr>
<tr>
<td>49.0</td>
<td>2 0 0</td>
<td>0.11</td>
</tr>
<tr>
<td>68.1</td>
<td>1 0 1</td>
<td>0.13</td>
</tr>
<tr>
<td>68.6</td>
<td>0 2 0</td>
<td>0.15</td>
</tr>
<tr>
<td>72.8</td>
<td>1 2 0</td>
<td>0.39</td>
</tr>
<tr>
<td>84.3</td>
<td>2 2 0</td>
<td>0.56</td>
</tr>
<tr>
<td>SSD (dB)</td>
<td></td>
<td>5.23</td>
</tr>
<tr>
<td>MSSV (dB²)</td>
<td></td>
<td>4.20</td>
</tr>
</tbody>
</table>

Table 3. Control room acoustic parameters and performance metrics after physical room correction

As seen above, the performance of the control room is significantly better across all three domains. This improvement is attributed to two key features. First, sources placed at the mid-points of the length and height dimensions coincide with the pressure nodes of the fundamental and all odd-order axial modes along those dimensions, resulting in complete modal suppression at those frequencies (24.5Hz and 68.1Hz). Second, by virtue of symmetrical placement of sources within the room, excitation generated at even-order axial modes by one pair of laterally opposing sources is balanced and cancelled by the second pair of sources radiating with inverted polarities. This form of complete modal cancellation occurs at 49Hz, 68.6Hz and 72.8Hz. At 84.3Hz, all four sources coincide with pressure antinodes of the same polarity, resulting in significant excitation of the mode which decays at the same rate as the control case. Note that the MT60 times for the suppressed modes in Table 3 are largely determined by the STFT time-frequency resolution parameters.

Finally, the overall bass level of the corrected system is significantly higher than the control system due to the room being driven by four sources instead of two. This can be addressed by adjusting the pre-amp gain for each of the four sources.

Modal suppression and cancellation achieved by the physical room correction scheme eliminates all pressure

Figure 10. Spectrogram plot for the RMS pressure average after physical room correction (Magnitude normalised and IR zero padded by 0.15s)

Figure 11. Post-correction RMS pressure averages across all six seats

Vol. 43. Pt. 2. 2021
undulations (peaks and nulls) caused by room modes up to 84.3Hz across all six seats. This results in a reduction in the steady-state deviation metric. Furthermore, by suppressing the modes themselves, the same improvement is achieved across the entire volume of the room, i.e., lower mean seat-to-seat variation (MSSV). The RMS pressure averages for the three cases are shown in Figure 11.

6 DISCUSSION

The previous sections summarise the motivation and performance of the two acoustic room correction systems across three domains of interest: time, frequency, and space.

By simulating representative examples, it has been shown that modal equalisation systems can provide meaningful improvements in the time domain by pre-filtering the primary excitation signals of a typical stereo setup. However, while subsequent magnitude equalisation stages may further optimise their steady-state performance, equalising the primary signals alone is insufficient to improve their spatial domain performance and consistency.

On the other hand, the room correction system achieved modal suppression and/or cancellation at most of the targeted frequencies. Consequently, the corrected sound field was free of resonant behaviour across all three domains. The critical limitation of the multi-source placement scheme is the lack of robustness with changes in the relative positions of the constituent sources. Minor deviation from their optimised positions may imbalance the suppression and cancellation mechanisms such that the net excitations at the affected modes are no longer close to zero. The residual modal excitation at such modes would then deteriorate the performance of the correction system across all three metrics. Furthermore, such correction schemes require multiple identical sources, making them expensive and space-consuming, especially in small domestic listening spaces.

7 CONCLUSION

This paper objectively contrasts the acoustic mechanisms underpinning room equalisation and room correction systems. Three performance metrics have been derived; modal decay times (MT60), steady-state deviation (SSD), and mean seat-to-seat variation (MSSV) to characterise the performance of the correction systems in the time, frequency, and spatial domains, respectively. Only the room correction system with additional sources for spatial control of the resultant sound field has been shown to offer meaningful improvements across all three domains.

Finally, while the chosen room model is only valid for rectangular rooms, the metrics presented can readily be used to assess the performance of non-rectangular rooms, provided the full complex room transfer functions are available by alternative means.

8 ACKNOWLEDGEMENTS

The author would like to thank his friends and colleagues at KEF R&D for their help and support.

9 REFERENCES


Vol. 43. Pt. 2. 2021


R. Wilson, M. D. Capp, and J. R. Stuart, ‘The Loudspeaker-Room Interface-Controlling Excitation of Room Modes’, May 2003.


