

# Room Sizing and Optimization at Low Frequencies\*

TREVOR J. COX,<sup>1</sup> *AES Member*, PETER D'ANTONIO,<sup>\*2</sup> *AES Fellow*, AND MARK R. AVIS<sup>1</sup>

<sup>1</sup>*School of Acoustics and Electronic Engineering, University of Salford, Salford, M5 4WT, UK*

<sup>2</sup>*RPG Diffusor Systems, Inc., Upper Marlboro, MD 20774, USA*

Modes in small rooms may lead to uneven frequency responses and extended sound decays at low frequencies. In critical listening environments, this often causes unwanted coloration effects, which can be detrimental to the sound quality. Choosing an appropriately proportioned room, and placing listener and loudspeakers in the right places can reduce the audible effects of modes. A new methodology is detailed for determining the room dimensions for small critical listening spaces as well as the optimum positions for sources and receivers. It is based on numerical optimization of the room geometry and layout to achieve the flattest possible frequency response. The method is contrasted with previous techniques for choosing room dimensions. The variations of the room quality for different room sizes are mapped out. These maps include an allowance for constructional variation, which has not been considered previously.

## 0 INTRODUCTION

The sound that is heard in a critical listening room is determined by the combined effect of the electronics of the audio system and the physical acoustics of the listening environment. The tonal balance and timbre of a sound can vary significantly, depending on the placement of the listener and loudspeaker and the geometry of the room. Indeed, the modal artifacts introduced by the room can be so influential that they dominate the sound. This paper concentrates on the design challenges to minimize these artifacts at low frequencies. Consequently it is concerned with the interaction between sources, listeners, and room modes.

Modes in small rooms often lead to extended sound decays and uneven frequency responses, often referred to as coloration. Problems arise at low frequencies because of the relatively low modal density. Designers try to overcome the problems of modes by choosing an appropriately proportioned room, placing listeners and loudspeakers in suitable positions, and using bass absorbers. This paper is concerned with the first two solutions, room sizing and optimization of loudspeaker and listener location.

---

\*Manuscript received 2001 September 24; revised 2004 January 8 and April 19. Parts of this work were first presented at the 103rd and 110th Conventions of the Audio Engineering Society, 1997 and 2001.

There have been many studies looking at room sizing, and this paper starts by discussing previous work by others who have suggested optimum room ratios or design methodologies.

The determination of appropriate source and receiver locations is often undertaken by trial and error, a laborious and difficult task. This is particularly true of surround-sound systems because so many loudspeakers have to be considered. Therefore a new automated method will be proposed and outlined which is based on numerical optimization. The old and new methods will be compared philosophically. Results in the form of modal responses and transient frequency responses are given to demonstrate the power of the new method. The paper will then investigate in detail the variations of the room spectral responses with the room size and produce plots to enable the choice of appropriate room sizes.

## 1 PREVIOUS WORK

### 1.1 Room Sizing

Many design methodologies and optimum room ratios have been suggested over the years to minimize coloration. Essentially these methods try to avoid degenerate modes, where multiple modal frequencies fall within a small bandwidth, and the corollary of bandwidths with absence of modes. The underlying assumption is that as music is played in the rooms, the absence or boosting of

certain tonal elements will detract from the audio quality. The starting point for these previous methods is usually the equation defining the modal frequencies within a rigid rectangular enclosure:

$$f = \frac{c}{2} \sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2} \quad (1)$$

where  $n_x$ ,  $n_y$ , and  $n_z$  are integers, and  $L_x$ ,  $L_y$ , and  $L_z$  are the length, width, and height of the room. Often the best dimensions are given in terms of the ratios to the smallest room dimension. Previous methods for determining room ratios differ, however, in how they utilize Eq. (1). Bolt [1] produced design charts that enabled him to determine preferred room ratios. His method investigated the average modal spacing to try and achieve evenly spaced modes, the assumption being that if the modal frequencies are spaced evenly, then there will be fewer problems with peaks and dips in the modal response. It is now known, however, that using the average mode spacing is not ideal, and the standard deviation of the mode spacing is a better measure. Ratios of 2:3:5 and 1:2<sup>1/3</sup>:4<sup>1/3</sup> (1:1.26:1.59) were suggested, but Bolt also noted that there is a broad area over which the average modal spacing criterion is acceptable. (Note that this latter ratio appears to be often rounded to the commonly quoted figures of 1:1.25:1.6.)

Gilford [2] discusses a methodology whereby the modal frequencies are calculated and listed. The designer then looks for groupings and absences assuming a modal bandwidth of about 20 Hz. The dimensions are adjusted and a recalculation is carried out until a satisfactorily even distribution is achieved. While this is a cumbersome process to undertake by hand, this type of iterative search can now be readily accomplished using computers and numerical optimization techniques. It is this type of computer-controlled optimization that is advanced in this paper as a more efficient method for choosing room dimensions. Furthermore, in addition to the use of numerical optimization to ease the burden of searching, a better basis than modal spacing for evaluating the effects of modes will be detailed. Gilford also states that the 2:3:5 ratio suggested by Bolt is no longer popular and that the axial modes cause the major difficulty in rooms. These points will be returned to later.

Louden [3] calculated the modal distribution for a large number of room ratios and published a list of preferred dimensions based on a single figure of merit. The figure of merit used to judge room ratios is the standard deviation of the intermode spacing, so again this is a regime to achieve evenly spaced modes. The method produces the well-known room ratio of 1:1.4:1.9. Louden undertook the investigation by examining 125 combinations of room ratios at a spacing of 0.1. This type of discretized search can limit the potential solutions found. With the optimized techniques developed since Louden published his work, such as the one used hereafter, the search for the best ratios can be undertaken more efficiently without the need to artificially discretize the ratios tested.

Bonello [4] developed a criterion based on the fact that the modal density should never decrease when going from

the one-third-octave band to the next higher band in frequency. Modes with coincidental frequencies are only tolerated in one-third-octave bands with five or more modes present. Bonello compares his criterion against others used by Knudsen, Olson, and Bolt. Justification for his methodology is drawn from his experience as a consultant for 35 rooms.

Walker [5] develops a low-frequency figure of merit based on the modal frequency spacing. The method leads to a range of practical, near optimum room shapes. In presenting his paper, Walker discussed how the blind application of optimum room ratios does not necessarily lead to the best room, because the room quality is volume dependent. The new method outlined hereafter does not use generalized room ratios, and so avoids this problem.

All the methods mentioned have limitations. Eq. (1) is only applicable for rectangular geometries with rigid surfaces. Absorption has a number of effects and, for instance, may shift the modal frequencies. This will affect the spacing of modes, which are the basis for many of the room ratio criteria. The new method set out in this paper uses a theoretical model which, although not perfect, is a more accurate model of low-frequency room behavior than that given by Eq. (1). Another effect of absorption is that it acts differently on axial, tangential, and oblique modes. For example, axial modes will have the greatest magnitude and least damping. None of the methods mentioned account for this fully, unlike the new method given hereafter, although Gilford, for example, does discuss the prominence of axial modes. A further difficulty with these methods is the choice of the criterion used for the evaluation. For example, Bonello's method makes several assumptions, such as the use of a one-third-octave bandwidth, and that five modes in a bandwidth mask the effects of coincident modes—which are empirical and subjective rather than fundamental in nature. The new method outlined here acts directly on the modal response of the room, so a criterion based on mode spacing is no longer required. Although an evaluation criterion is still required, since this can be based on the modal response of the room, it is much easier to relate to human perception. This is because the mode spacing is one level more removed from the actual signals received by the listener than the modal response.

Standards and recommendations also stipulate good room ratios for activities such as listening tests and broadcasting, and recent versions have drawn on work by Walker [5], [6]. Recommendations of the European Broadcasting Union [7] are discussed by Walker. He states that the aim of the regulations is to avoid worst cases, rather than to provide proscriptive optimum ratios. Consequently the recommendations cover a wide range of room proportions,

$$\frac{1.1L_y}{L_z} \leq \frac{L_x}{L_z} \leq \frac{4.5L_y}{L_z} - 4 \quad (2)$$

$$L_x < 3L_z \quad (3)$$

$$L_y < 3L_z \quad (4)$$

In addition, it is stipulated that ratios of  $L_x$ ,  $L_y$ , and  $L_z$  that are within  $\pm 5\%$  of integer values should also be avoided. The British Standards Institute and the International Electrotechnical Commission [8] give slightly different criteria for Eq. (2),

$$\frac{L_y}{L_z} \leq \frac{L_x}{L_z} \leq \left( 4, \frac{5L_y}{L_z} - 4 \right). \quad (5)$$

The criteria given by Eqs. (3) and (4) are also stipulated along with recommended floor areas. A recommended room size of  $7 \times 5.3 \times 2.7$  m (2.59:1.96:1) is given. Older versions of the standard [9] give different recommendations, with a standard room of  $6.7 \times 4.2 \times 2.8$  m (1.59:1.5:1). These values are also reported in a popular textbook [10].

### 1.2 Room Layout Optimization

Room layout optimization concerns the determination of the source and receiver locations in the room. When sources or receivers are moved, the frequency response changes due to the variation in the modal pressure distribution in the room and the changing radiation resistance of the source. By choosing correct positions in the room, it is possible to minimize the audible effects of the modes within a room [11]. Besides considering the modal (steady-state) response, others have considered the effects of first-order reflections from boundaries. In particular, the first-order reflections from the nearest wall boundaries to the source have been considered along with the effect these reflections have on the frequency response (see, for example, [10]). This simple model of room reflection response is an oversimplification in small rooms, as reflections from the other boundaries and higher order reflections arrive quickly, suggesting that the frequency response of the first-order reflections will be masked.

### 1.3 New Method

The new method is based on finding the rooms that have the flattest possible modal frequency response. It uses a computer algorithm to search iteratively for the best room sizes. To determine the best source and receiver location, a similar search is done. However, this time both the modal response and the frequency response are calculated over a shorter time period related to the room transient response. In the following sections details of the new method will be given. The prediction models used will be presented in Section 2, and the optimizing procedure will be discussed in Section 3.

## 2 PREDICTION MODELS

For the purposes of this paper, the modal response of the room is defined as the frequency spectrum received by an omnidirectional microphone when the room is excited by a point source with a flat power spectrum. When considering room sizing, the source and the receiver are placed in opposite corners of the room following normal practice. Two possible models to predict the modal response are considered, a frequency-based modal decomposition model and a time-based image source model.

### 2.1 Modal Decomposition Model

The modal decomposition model used is applicable when boundary impedances are large and real, which corresponds to walls that are nonabsorbing because they are either massive or very stiff. The pressure at  $\bar{r}(x, y, z)$  due to a source at  $\bar{r}_0(x_0, y_0, z_0)$ , at an angular frequency  $\omega$ , is given by [12]

$$p(\bar{r}, \omega) = \sum_{n_x}^{\infty} \sum_{n_y}^{\infty} \sum_{n_z}^{\infty} \frac{A_{\omega_n}(\bar{r}, \bar{r}_0)}{(\omega^2 - \omega_n^2 - j2\omega_n \delta_n \omega)} \quad (6)$$

where

$$A_{\omega_n}(\bar{r}, \bar{r}_0) = j\rho\omega S_0 c^2 p_n(\bar{r}) p_n(\bar{r}_0) \quad (7)$$

$$p_n(\bar{r}) = \cos(k_{n_x} x) \cos(k_{n_y} y) \cos(k_{n_z} z) \quad (8)$$

$$p_n(\bar{r}_0) = \cos(k_{n_x} x_0) \cos(k_{n_y} y_0) \cos(k_{n_z} z_0) \quad (9)$$

$$\frac{\omega_n}{c} = k_n = \sqrt{k_{n_x}^2 + k_{n_y}^2 + k_{n_z}^2} \quad (10)$$

$$\delta_n = \frac{c}{\omega_n} \left( \frac{\varepsilon_{n_x} \bar{\beta}_x}{L_x} + \frac{\varepsilon_{n_y} \bar{\beta}_y}{L_y} + \frac{\varepsilon_{n_z} \bar{\beta}_z}{L_z} \right) \quad (11)$$

with  $\bar{\beta}_x$ ,  $\bar{\beta}_y$ , and  $\bar{\beta}_z$  being the average admittance (the reciprocal of the impedance) of the walls in the  $x$ ,  $y$ , and  $z$  directions and thereby accounting for energy loss and phase change on reflection,

$$\varepsilon_{n_x} = 1 \quad \text{for } n_x = 0 \quad (12)$$

$$= 2 \quad \text{for } n_x > 0 \quad (13)$$

$$k_{n_x} = \frac{n_x \pi}{L_x}. \quad (14)$$

Similar expressions are used for  $\varepsilon_{n_y}$ ,  $\varepsilon_{n_z}$ ,  $k_{n_y}$ , and  $k_{n_z}$ .  $\rho$  is the density of air,  $S_0$  is a constant, and  $c$  the speed of sound.

### 2.2 Image Source Model

The image source model is a fast prediction model for a cuboid room. The image solution of a cuboid enclosure gives an exact solution of the wave equation if the walls of the room are nonabsorbing. The energy impulse response is given by

$$E(t, \bar{r}, \bar{r}_0) = \sum_{n_x}^{\infty} \sum_{n_y}^{\infty} \sum_{n_z}^{\infty} \sum_{i=1}^2 R_{n,i}^2 \frac{1}{d_{n,i}^2} \quad (15)$$

$$d_{n,i} = \sqrt{d_{n_x,i}^2 + d_{n_y,i}^2 + d_{n_z,i}^2} \quad (16)$$

$$d_{n_x,1} = (L_x - x_0) + (-1)^{n_x} x + L_x(n_x - 1) + L_x[(-1)^{n_x+1} + 1]/2 \quad (17)$$

$$d_{n_x,2} = x_0 + (-1)^{n_x+1} x + L_x(n_x - 1) + L_x[(-1)^{n_x} + 1]/2 \quad (18)$$

Similar expressions are used for the distances in the  $y$  and  $z$  directions. The surface reflection factors are given as

$$R_{n,i} = R_{x,i}R_{y,i}R_{z,i} \quad (19)$$

$$R_{x,i} = R_{x,i}^{|l_{x,i}|} R_{x,\text{mod}(i,2)+1}^{|l_{x,i}-1|} \quad (20)$$

where  $R_{x,1}$  and  $R_{x,2}$  are the surface reflection factors for the front and rear walls, respectively, and similar expressions are used for the distances in the  $y$  and  $z$  directions. Reflection factors are approximated to be purely real. Once the energy impulse response is obtained, it is Fourier transformed to form the modal frequency response.

### 2.3 Transient Response

The modal response represents the steady-state reaction of the room to sound. Music is a complex mix of both steady-state and transient signals, and so the modal response describes only part of the subjective listening experience. For example, Olive et al. [13] showed that the first part of the decay of a high- $Q$  mode is most noticeable, as the later parts of the decay are often masked by a following musical note. Consequently some measure of the early arriving sound field should be considered in addition to the modal response. To do this the first part of the impulse response is taken, a half-cosine window applied, and the windowed impulse response Fourier transformed to give a transient frequency response. The half-cosine window is used to reduce truncation effects. For the results presented in this paper, a time period of 64 ms was chosen because it approximately relates to the integration time of the ear for the detection of reflections [14]. It can be argued, however, that for the perception of loudness at low frequencies a longer time period might be more applicable.

### 2.4 Prediction Model Critique

The modal decomposition and the image source models both offer a better representation of the sound field in the space than the simple modal frequency equation, Eq. (1), used for previous room-sizing methodologies. This is primarily because the modal decomposition and image source models allow for absorption, but also because it is

possible to calculate a quantity—the modal response—that is easier to relate to the listener experience. Both models, however, are not completely accurate. Fig. 1 compares measurements in a listening room to the modal decomposition and image source models. The listening room has dimension  $6.9 \times 4.6 \times 2.8$  m. All the walls were smooth plastered concrete, except for the back wall, front wall, and ceiling, which contained areas of diffusers, and the floor, which was covered with carpeting. Normalization of the loudspeaker sound power was carried out by measuring the cone acceleration using an accelerometer attached near the center of the loudspeaker cone. If the cone radiates as a piston at low frequencies, the free-field pressure should be omnidirectional and proportional to the cone acceleration.

Below 125 Hz good agreement is shown between models and measurement. The agreement diverges somewhat above 125 Hz. There could be many sources of error, most likely the improper modeling of frequency-dependent absorption and the influence of the large-scale diffusers present in the room. A slightly better agreement can be achieved [15] by taking more terms in Eqs. (6) and (15). The models deliberately used a reduced number of terms (15) in the infinite sums to enable calculations to be quick enough for subsequent optimization. The accuracy of the predictions is similar to that observed by others [16].

Although the method for choosing room dimensions is based on a better prediction model than previous methods, there is room for further refinement. There are some basic problems with both the modal decomposition and the image source models, and currently there are no established solutions to deal with these difficulties. For example, while absorption coefficients for surfaces are widely available, the surface impedances, which include both phase and magnitude information, are not. Indeed, given that room surfaces at low frequencies will often not behave as isolated locally reacting surfaces, defining a surface impedance can be problematic. Consequently, for this work an assumption of no phase change on reflection has to be made, which means that the models are more accurate for walls that are nonabsorbing. It might be envisaged that a

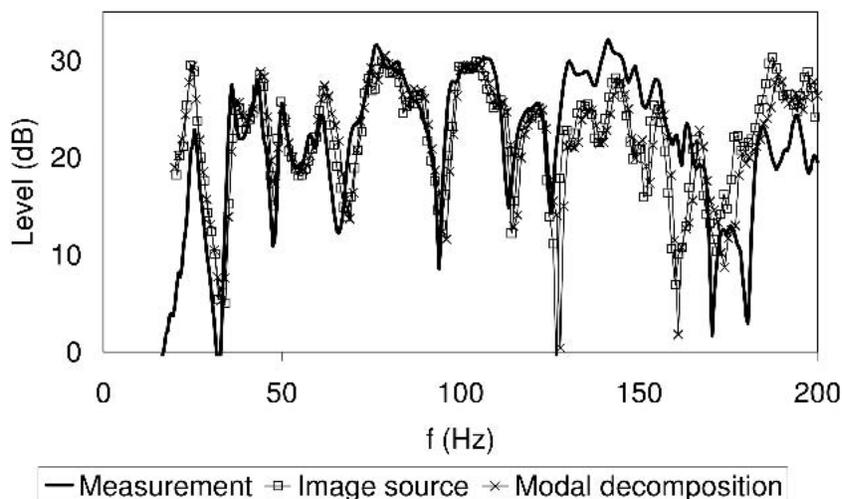


Fig. 1. Comparison of image source model, modal decomposition model, and measurement in a listening room.

finite-element model could overcome some of these difficulties, but currently the calculation time would be too long for efficient optimization. During an optimization process, many hundreds or thousands of room configurations have to be modeled, and consequently the prediction time for a single calculation must be kept small.

For the results presented here, the image source model was favored over the modal decomposition model. This is because the image source model is faster. For the modal decomposition model all modes within the frequency range of interest must be considered, plus corrections for the residual contributions of modes with peak frequencies outside this range [17]. In the image source model all images contribute to the impulse response in a cuboid room. Consequently using the image source model reduces the optimization time. Furthermore, the methodology of room layout optimization uses a transient response in the room that can best be obtained from a time-based calculation. The relationship between the modal decomposition and the image solutions for a lossless room has been derived and shown to be equivalent for nonabsorbing boundaries [18].

### 3 OPTIMIZATION PROCEDURE

Numerical optimization techniques are commonly used to find the best designs for a wide variety of engineering problems. In the context of this paper, a computer algorithm is used to search for the best room dimensions and locations for sources and receivers. To simplify the explanation, first consider the case of finding the best room dimensions by searching for the flattest modal response. This is done with a source in one corner of the room and a receiver in the opposite corner. The iterative procedure is illustrated in Fig. 2. The user inputs the minimum and maximum values for the width, length, and height, and the algorithm finds the best dimensions within these limits. The routine then predicts the modal response of the room and rates the quality of the spectra using a single figure of merit (cost parameter). A completely random search of all possible room dimensions is too time consuming, and so one of the many search algorithms that have been developed for general engineering problems was used—in this case the simplex method [19]. This is not the fastest procedure, but it is robust and does not require knowledge of the figure of merit’s derivative.

In developing a single figure of merit, it is necessary to consider what would be the best modal response. It is assumed that the flattest modal response corresponds to the ideal. This is done even though a perfectly flat response can never be achieved, as in the sparse modal region there will always be minima and maxima in the frequency response. The cost parameter used is the sum of the squared deviation of the modal response from a least squares straight line drawn through the spectra. If the modal response level of the  $n$ th frequency is  $L_{p,n}$  then the cost parameter  $\epsilon$  is

$$\epsilon = \sum_{n=1}^N (L_{p,n} - mf_n + c)^2 \tag{21}$$

where  $m$  and  $c$  are the gradient and intercept of the best-fit line and the sum is carried out over  $n$  frequencies  $f_n$ . This is illustrated in Fig. 3. Consequently this is a least squares minimization criterion, which is commonly used in engineering. The deviation from a best-fit line rather than the mean is used because it is assumed that a slow variation in the spectrum can be removed by simple equalization, and what is important is to reduce large local variations. Before calculating Eq. (21), some smoothing over a few adjacent frequency bins is used. This is done to reduce the risk of the optimization routine finding a solution that is overly sensitive to the exact room dimensions. Furthermore, in prediction models very exact minima can be

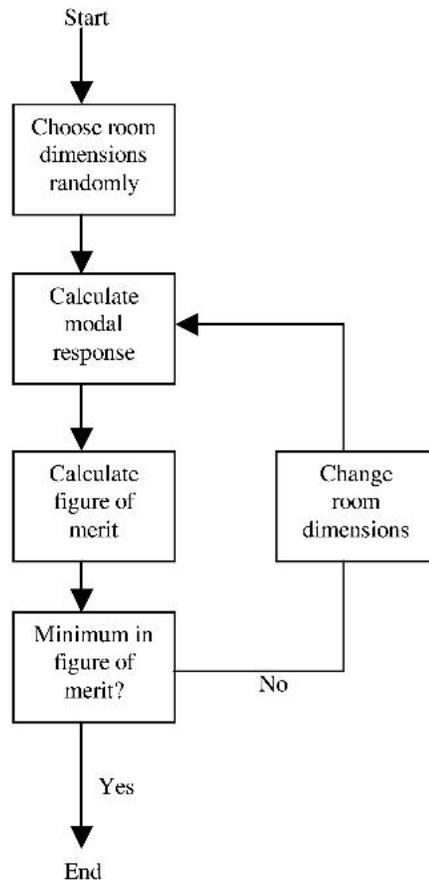


Fig. 2. Optimization procedure for room sizing.

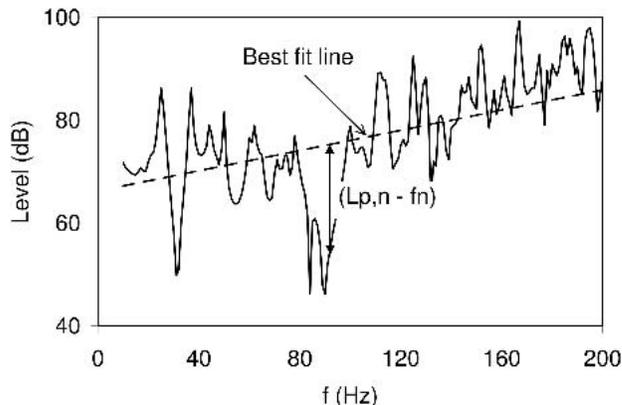


Fig. 3. Use of best fit-line to obtain figure of merit (no spectral smoothing).

found which would never be replicated in real measurements; the smoothing helps mitigate against this. The peaks in a modal response are generally more of a problem than the dips [13]. It is possible to give more weighting to the peaks by altering the merit factor. The squared deviations above and below the best-fit line can be calculated, and these can be added with a weighting factor so that a greater contribution comes from the deviations above the best-fit line. In the present work, however, equal weight is given to the peaks and dips.

When optimizing the room layout it is necessary to consider both the steady-state modal and the transient responses. These are both calculated with the desired source and receiver positions, that is, the modal response is not calculated between the corners of the room. The figure of merit must be a single cost parameter, and so it is necessary to combine the cost parameters for both the modal and the transient responses calculated using Eq. (21). For this a simple average is taken. Complications in the room layout optimization arise because the loudspeakers and the listener positions are interdependent. Consequently, defining the search limits for loudspeakers and listeners can result in highly nonlinear constraints being applied to the optimization procedure. It is necessary to define the search regions for the source and the receivers, and in our implementation these regions are defined as cubes. It would be possible to allow all the sources and receivers to vary independently, but in reality some constraints must be applied. For example, it is necessary to maintain the angles subtended by the front loudspeakers to the listeners within reasonable limits for correct stereo reproduction. These constraints are applied by brute force. If the simplex routine attempts to place the front stereo loudspeakers at an inappropriate angle, the sources are moved to the nearest point satisfying the angular constraints for stereo within the search cube defined by the user. Further complications occur because in most listening situations certain loudspeaker positions are determined by others. For example, in a simple stereo pair, both loudspeakers are related by a mirror symmetry about the plane passing through the listener. Consequently there is only one independent source location that defines a stereo pair, since the other is de-

pendent on the independent loudspeaker and listener positions. For surround-sound formats, similar interrelations exist, which can be exploited.

## 4 TEST BED

### 4.1 Room Sizing

The optimizer was run for a wide variety of room sizes:  $7 \text{ m} \leq L_x \leq 11 \text{ m}$ ,  $4 \text{ m} \leq L_y \leq 8 \text{ m}$ , and  $3 \text{ m} \leq L_z \leq 5 \text{ m}$ . A large number of solutions were gathered (200) to enable the performance of the optimization to be tested and to undertake a statistical analysis of the solutions found. In most multidimensional optimization running the procedure repeatedly from random starting positions will give different “optimum” solutions. This happens because the optimizing algorithm will get stuck in a local minimum that is not the numerically best solution (the “global minimum”) available. It has been found that in room sizing, the difference in the modal response between the global minimum and other good local minima is negligible, however. Consequently, when used as a design tool, far fewer solutions need be calculated than might be thought necessary, and the best used with a good degree of confidence.

A frequency range of 20–200 Hz was chosen, as above 200 Hz the flatness of the modal response was not particularly sensitive to changes in dimension. As might be expected, the gains to be made in avoiding degenerate modes are at lower frequencies, where the modes are relatively sparse. In addition, the accuracy of the prediction models decreases with increasing frequency. The frequency range for optimization may also be guided by the Schroeder frequency. For the results shown here, the absorption coefficients were chosen to be 0.12 for all walls.

#### 4.1.1 Results

To compare with previous work, an optimized solution whose volume was roughly the same as that used by Loudon was chosen. This is to enable the fairest possible comparison.

Fig. 4 shows the optimized modal response (1:1.55:1.85) compared to one of the ratios suggested by Bolt (2:3:5). In addition, the modal spectrum for the worst di-

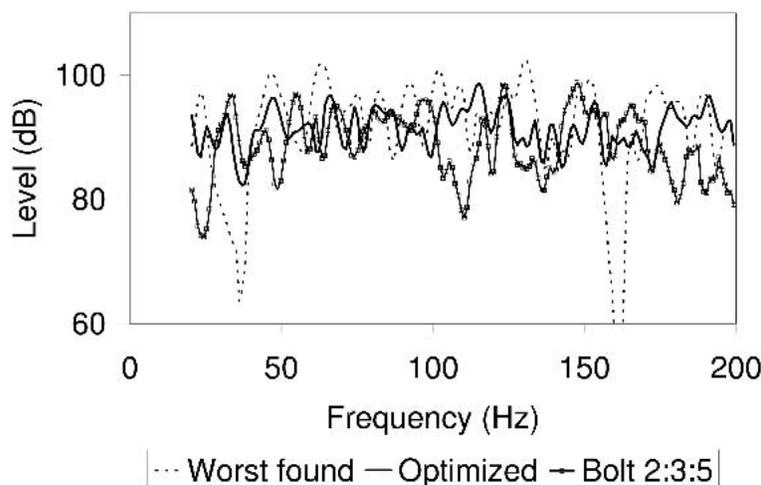


Fig. 4. Modal response for three room dimensions, including Bolt’s 2:3:5 ratio.

mensions found during the search is shown to give a sense of the range of spectra that can be achieved (1:1.07:1.87). As expected, a completely flat spectrum is not achieved with optimization. A clear improvement on the Bolt 2:3:5 room is seen, however. The 2:3:5 room suffers from significant dips, an example of which can be seen at 110 Hz.

The best ratio found by Loudon (1:1.4:1.9) is compared to the optimized response in Fig. 5. Improvement on the Loudon ratio is achieved, although the improvement is less marked than with 2:3:5. Bolt also suggested the ratio 1:1.25:1.6, which meets Bonello’s criteria as well. Fig. 6 shows the spectra compared to the optimized solution. The modal response spectrum achieved by optimization is clearly flatter.

The optimized solution was also compared to the regulations and standards mentioned earlier. All of the ratios by Bolt, Bonello, and Loudon presented pass the EBU and IEC regulations, as does the optimized solution. Only the worst case fails to meet the regulations. The standards appear to achieve their remit of not being overly proscriptive while avoiding the worst cases. A comparison with the preferred standard room sizes given in the standards and regulations was also undertaken. Fig. 7 compares the op-

timized solution to the old and new IEC regulations. The new standard room and the optimized solution are very similar in performance. While the cost parameter is better for the optimized solution (2.2) than for the new standard room (2.5), this does not translate into an obvious improvement in the spectrum. (This gives a little evidence of the sensitivity of the cost parameter; and the difference limen appears to be greater than 0.3.) The old standard room, however, is far from optimum, indicating a wise revision of the standard.

Finally the optimized solution is compared to the “golden ratio” (1:1.618:2.618) in Fig. 8. The golden ratio is often quoted in the audio press, and so is of interest. It was tested for this reason, despite the fact that the rationale behind the golden ratio for room dimensions does not appear particularly compelling from a scientific point of view. It can be seen that the optimized solution has a more even modal response and so is better.

**4.2 Room Layout Optimization**

Fig. 9 shows the best and worst spectra found for a stereo loudspeaker position optimization. For both the transient and the modal responses there is an improvement

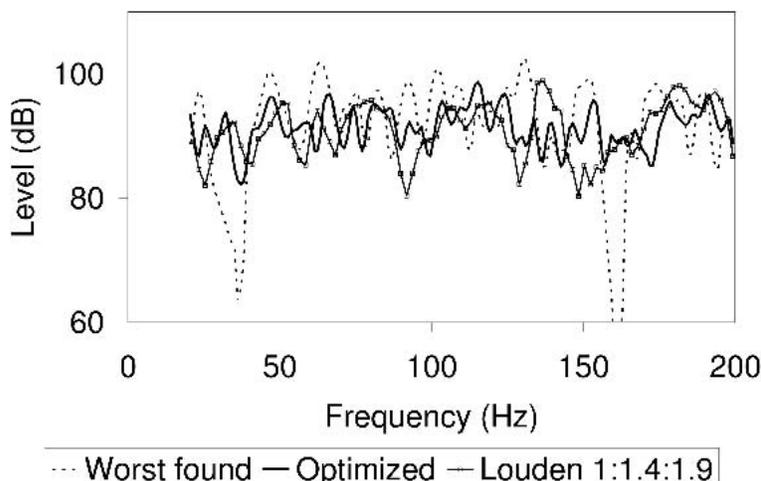


Fig. 5. Modal response for three rooms, including best ratio found by Loudon (1:1.4:1.9).

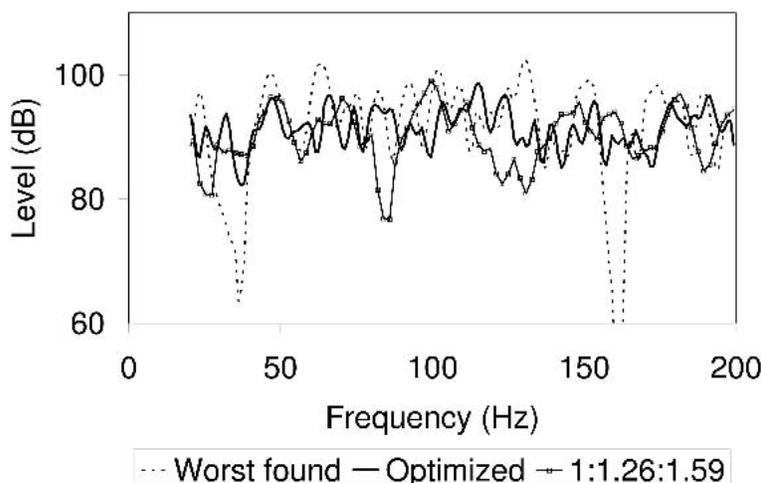


Fig. 6. Three modal spectra, including 1:1.26:1.59.

in the flatness of the frequency response. These are typical of the results found. Unfortunately a comparison with other work is difficult because of the lack of previous literature.

### 5 DISCUSSION

The new method produces as good or better room dimensions than those based on previous work. The new

method has been shown to be an efficient way of finding optimum dimensions and loudspeaker and listener positions. The modal response in a room is complex, and there does not appear to be one set of magical dimensions or positions that significantly surpass all others in performance. There may be a numerically global minimum, but many of the local minima are actually equivalent in terms of the quality of the frequency response achieved.

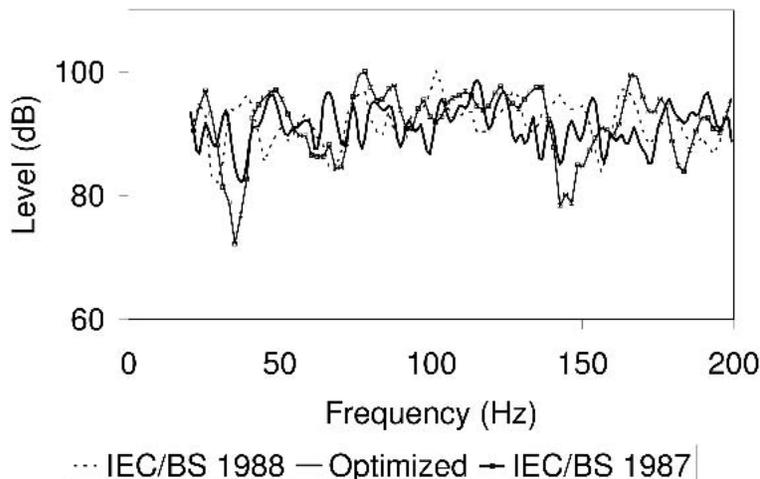


Fig. 7. Comparison of optimized solution with standard rooms.

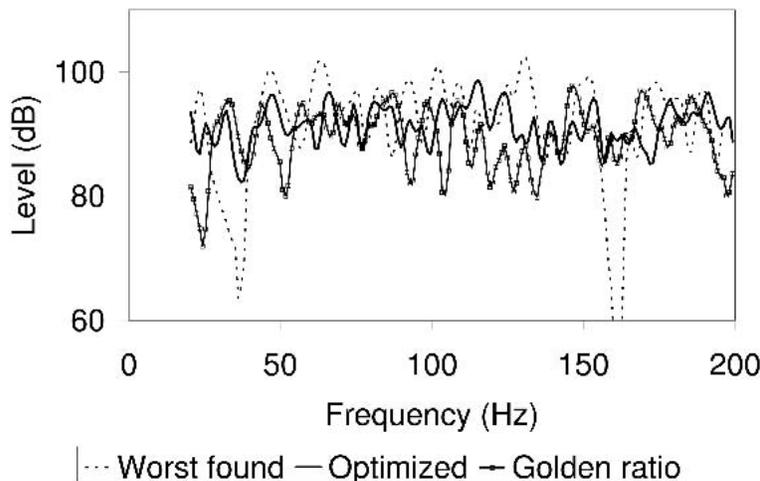


Fig. 8. Comparison of three room modal responses, including "golden ratio."

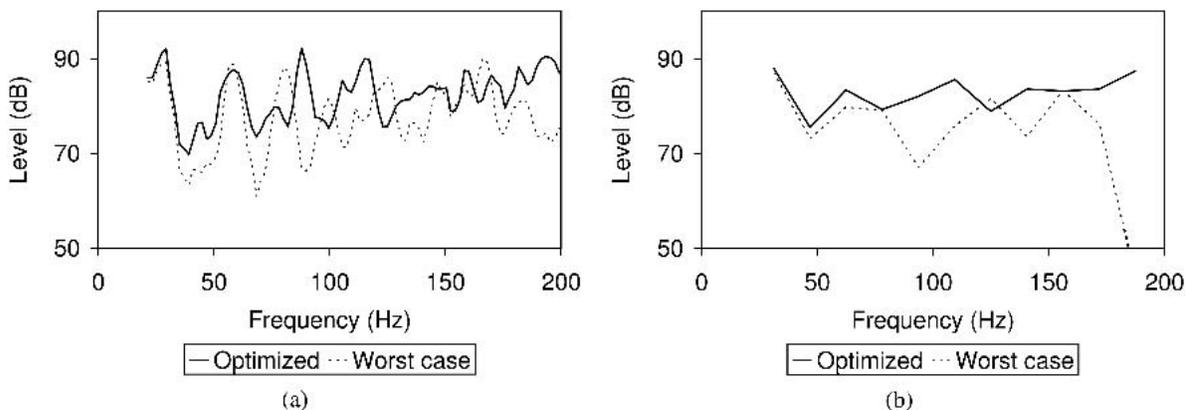


Fig. 9. Results from optimizing for position. (a) Modal response. (b) Transient response.

One significant advantage of this optimization technique is that it is possible to incorporate constraints that may happen in real buildings. To take a room sizing example, if the height of the ceiling is fixed in the building, then it can be fixed in the optimizer, which can then look for the best room width and length within the constraints given by the user.

### 6 MAPS

It is possible to map out the complete error space being searched by the optimizer, and therefore get further insight into the processing of room sizing and optimizing source and receiver positioning. This has been explored for the problem of room sizing rather than layout optimization as this enables a comparison with previous work. Such an approach was carried out by Walker [5], and the findings from his work have been fed into various regulations for designing listening environments. The approach used in this paper is similar, except that the error parameter is more sophisticated as it is evaluated using a spectrum rather than the modal spacing. Furthermore, it has been investigated how robust a particular set of room ratios is to mismatches between theory and reality.

For any particular room volume it is possible to plot the variation of the figure of merit with the room ratios. Fig. 10 shows such a plot. Light areas have a large figure of merit, and correspond to uneven frequency responses; dark areas correspond to the best room sizes. A factor of merit of 7 dB means that 95% of the levels in a spectrum were within  $\pm 7$  dB of the mean. This graph shows great similarities with the contour plots produced by Walker. The main light diagonal line running from the bottom left to the top right corresponds to rooms where two of the dimensions are similar, having a square floor plan, and

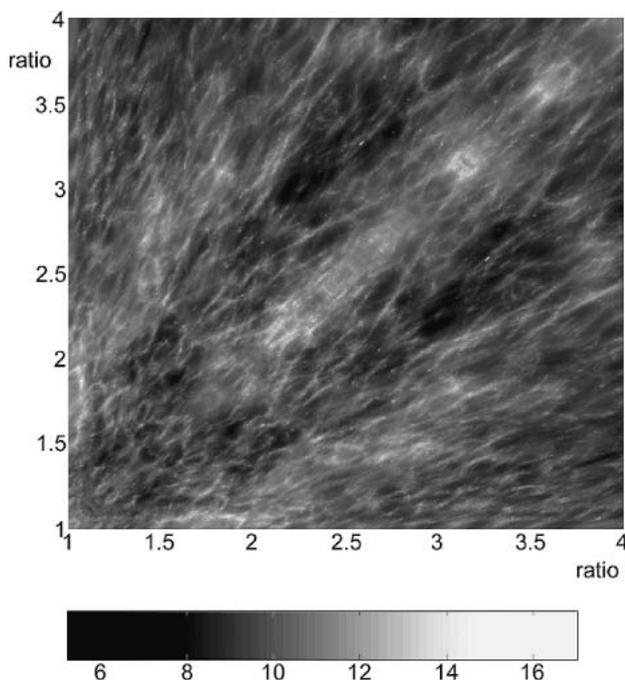


Fig. 10. Variation of figure of merit (in dB) for 100-m<sup>3</sup> room with room ratio.

therefore the distribution of the modes is uneven and the figure of merit poor.

In interpreting these data it is important not only to look at the value of the figure of merit at one point, but also to look at whether nearby points are similarly good. For a set of room dimensions to be useful, it is necessary for it to be robust to changes in room size due to construction tolerances in terms of the size of the room and the properties of the construction material. The theoretical model will not exactly match reality, and it is important that the set of room dimensions chosen be not overly sensitive to these construction tolerances. Otherwise it is likely that the real room may fail to perform as well as expected. Consequently the best room dimensions are surrounded by other good dimensions, and would be shown as broad dark areas in Fig. 10. Using this principle, Fig. 10 can be further analyzed to clarify this point. The figure of merit for a room dimension ( $L_x, L_y, L_z$ ) is calculated as the largest value for similar sized rooms whose dimensions are bounded by

$$\begin{aligned} L_x - \Delta L_x &\leq L_x \leq L_x + \Delta L_x \\ L_y - \Delta L_y &\leq L_y \leq L_y + \Delta L_y \\ L_z - \Delta L_z &\leq L_z \leq L_z + \Delta L_z \end{aligned} \tag{22}$$

where  $\sqrt{\Delta L_x^2 + \Delta L_y^2 + \Delta L_z^2} < 5$  cm.

Typically 20 rooms around a particular set of room dimensions are considered in the process. The largest figure of merit value is taken rather than an average because this assumes a worst-case scenario and reduces the risk of a poor listening room being designed.

Fig. 11 shows the plot after this “averaging” process. Now dark areas of the graph indicate room ratios that are not only good, but are robust to construction tolerances. The plot is, however, difficult to read and interpret, and consequently an additional process is undertaken. While

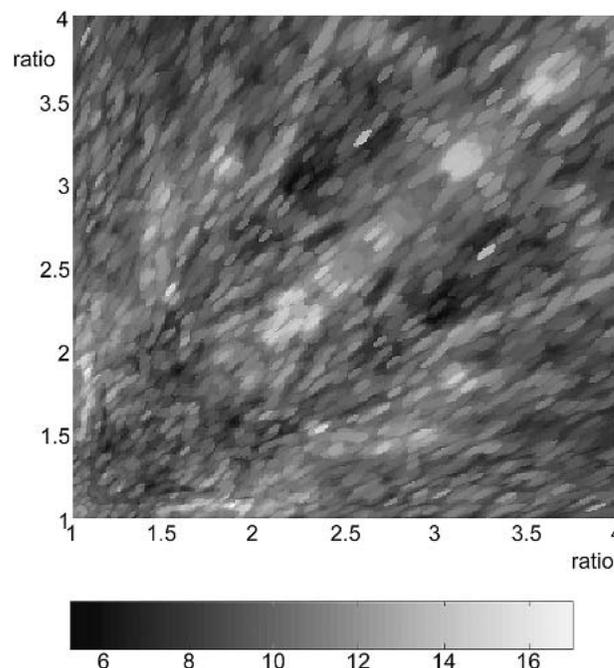


Fig. 11. Variation of figure of merit (in dB) for 100-m<sup>3</sup> room with room ratio after “averaging” to allow for parameter sensitivity.

the plot shows how the quality of a room varies with its dimensions, what the designer is primarily concerned with are which ratios give the best room response. For this purpose the figures of merit are categorized into three classes: best ratios, reasonable ratios, and others. Fig. 12 shows the categorized plot.

The problem is that it is difficult to know the sensitivity of the listener to the merit factor values and so it is difficult to make the categorization. To put it more simply, is a room with a 9 dB figure of merit as good as one of 7 dB? What is the smallest perceivable difference—the difference limen? This is a problem common to all the methods that look at the flatness of the frequency response since rigorously derived subjective data describing the factors of merit are not available. By inspecting some of the spectra

for different factors of merit it is suggested that 2 dB would be a reasonable first guess at the difference limen. Spectra of merit factors differing by 2 dB exhibit clearly visible differences in response, which might be expected to be audible. Further work is proposed in this area to define a subjective sensitivity accurately.

Figs. 13–15 show the categorized plots for three room volumes. Using these plots it is possible to design rooms with good low-frequency responses. It is also possible to investigate a few key issues, as discussed in the next sections.

### 6.1 Room Ratios

How relevant are room ratios to room design? If a simple room ratio could be used regardless of the room volume, then it would be expected that the shaded maps

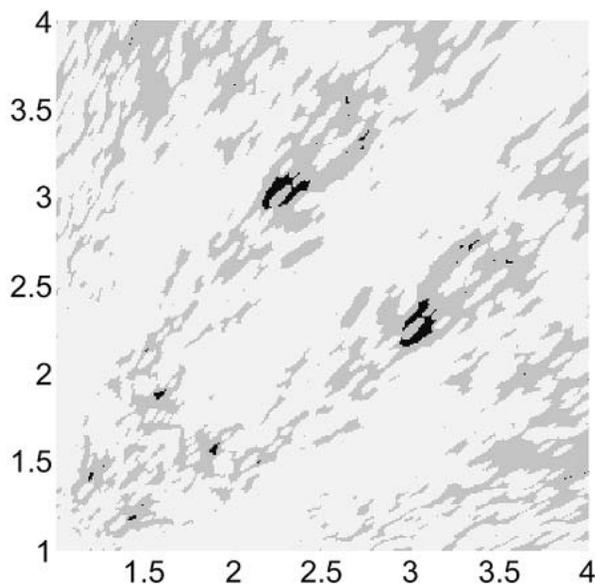


Fig. 12. Variation of room quality with room ratio after categorization for 100-m<sup>3</sup> room.

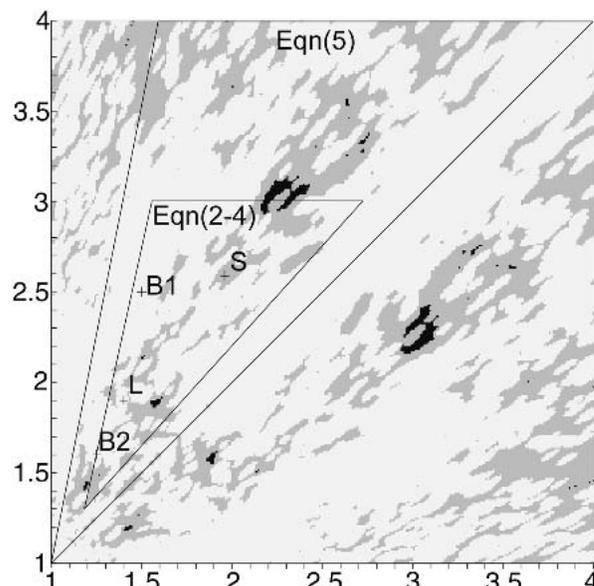


Fig. 14. Variation of room quality for 100-m<sup>3</sup> room. S—standard room size 7 × 5.3 × 2.7 m. Legend same as in Fig. 13.

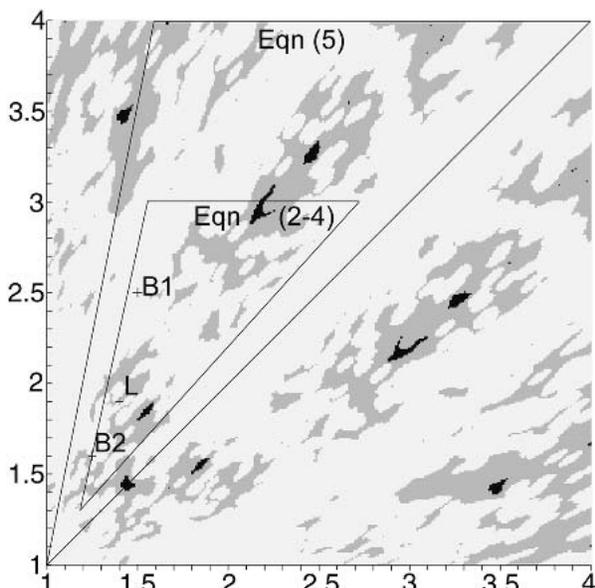


Fig. 13. Variation of room quality for 50-m<sup>3</sup> room. Triangular regions are mapped out by equations indicated. B1, B2—location of two ratios attributed to Bolt; L—location of best ratio of Louden.

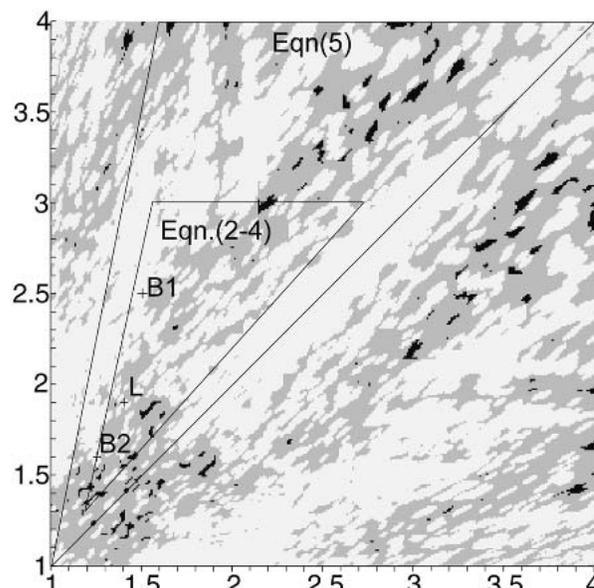


Fig. 15. Variation of room quality for 200-m<sup>3</sup> room. Legend same as in Fig. 13.

produced would be the same whatever the room volume. Figs. 13–15 show the plots for rooms of 50, 100, and 200 m<sup>3</sup>. The smallest room has fewer good ratios compared to the larger rooms, although the general pattern has some similarities. If the strictest criterion for a room is taken, using the darkest areas in the figures, then there are very few room ratios (about 20 clustered around 1:2.19:3) that can be applied to all three room volumes. These 20 ratios are useful because they are robust to room volume. However, it is overly restrictive to work with these ratios alone, because there are plenty of other solutions with different aspect ratios that might be useful in a particular room design. Using an optimizing procedure, as outlined previously, frees the designer from the need to work with only a small number of ratios that apply across all volumes.

## 6.2 Comparison with Best Ratios from Literature

The shaded plots also allow the results to be compared with previous work. Bolt suggested a ratio of 2:3:5 (equivalent to 1:1.5:2.5), which is labeled B1 in Figs. 13–15. Interestingly this is not a ratio with a high factor of merit for any of the room volumes tested. The best ratio found by Loudon (1:1.4:1.9, labeled L) is also shown, along with the ratio of 1:1.25:1.6 suggested by Bolt (labeled B2). Again these do not appear to represent a good choice. It is suggested that this is because the ratios are not robust to constructional variations. The best room (7 × 5.3 × 2.7 m) suggested in standards, which has a volume of 100 m<sup>3</sup>, fared well in comparison to the optimized room as discussed previously. This is labeled S in Fig. 14 and is shown to be a reasonable and robust solution, although better ratios do exist.

Figs. 13–15 also show the regions suggested in various standards. (They also have the stipulation that the ratios should not be within ±5% of integer values, but these exclusions are not marked.) The standards identify wedges shaped sets of ratios which avoid the worst possible rooms. In light of the analysis shown here, it might be appropriate to alter the criterion to better identify regions where the probability of building a good room is increased. In particular, the upper bound in Eq. (2), relating to rooms with large aspect ratios, allows rooms with relatively poor quality. It is suggested that revising the criterion would solve this problem:

$$\frac{1.1L_y}{L_z} \leq \frac{L_x}{L_z} \leq \frac{1.32L_y}{L_z} + 0.44. \quad (23)$$

Alternatively, desirable room ratios can be read directly from Figs. 13–15, which enable robust, good-quality rooms to be achieved.

## 7 CONCLUSIONS

A method has been presented to enable determining the size of small critical listening spaces as well as appropriate positions for the loudspeakers and listeners. Criteria for room size and layout have been adopted so as to minimize the coloration effects of low-frequency modes, as pre-

dicted in steady-state modal and transient responses. This method is an improvement over previous methods for determining room sizes in that the theoretical basis relies on a more accurate model of the room response, as opposed to a simple examination of the modal frequency spacing for a rigid box. The system is flexible in that it can search for the best dimensions within constraints set by the designer. Furthermore, the procedure has flexibility in that as better prediction models of rooms become available, they can be used within the general optimization design procedure. Results demonstrate that the new search method produces room sizes that match or improve on the room ratios published in the literature. When applied to room layout, the optimization procedures are successful in finding positions with flatter modal and transient responses. Variations in room quality for different room sizes have been examined and plots produced which enable the designer to select appropriate room sizes. These plots have included allowance for constructional variations, which had not been considered in previous work. Future work will include adding this concept of construction variation to the optimization algorithm, in addition to investigations into the subjective characterization of the chosen cost parameter.

## 8 ACKNOWLEDGMENT

The authors would like to thank Y. W. Lam for his advice on prediction models, and Andrew West for carrying out the room measurements.

## 9 REFERENCES

- [1] R. H. Bolt, "Note on The Normal Frequency Statistics in Rectangular Rooms," *J. Acoust. Soc. Am.*, vol. 18, pp. 130–133 (1946).
- [2] C. L. S. Gilford, "The Acoustic Design of Talk Studios and Listening Rooms," *J. Audio. Eng. Soc.*, vol. 27, pp. 17–31 (1979 Jan./Feb.).
- [3] M. M. Loudon, "Dimension Ratios of Rectangular Rooms with Good Distribution of Eigentones," *Acustica*, vol. 24, pp. 101–104 (1971).
- [4] O. J. Bonello, "A New Criterion for the Distribution of Normal Room Modes," *J. Audio. Eng. Soc.*, vol. 29, pp. 597–606 (1981 Sept.); Erratum, *ibid.*, p. 905 (1981 Dec.).
- [5] R. Walker, "Optimum Dimension Ratios for Small Rooms," presented at the 100th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 44, p. 639 (1996 July/Aug.), preprint 4191.
- [6] R. Walker, "A Controlled-Reflection Listening Room for Multichannel Sound," *Proc. Inst. Acoust. (UK)*, vol. 20, no. 5, pp. 25–36 (1998).
- [7] EBU R22-1998, "Listening Conditions for the Assessment of Sound Programme Material," Tech. Recommendation, European Broadcasting Union (1998).
- [8] IEC 60268-13, BS6840 13, "Sound System Equipment—Part 13: Listening Tests on Loudspeakers," International Electrotechnical Commission, Geneva, Switzerland (1988).
- [9] IEC 268-13:1985, BS6840 13, "Sound System Equipment—Part 13: Listening Tests on Loudspeakers," International Electrotechnical Commission, Geneva, Swit-

zerland (1987).

[10] J. Borwick, Ed., *Loudspeaker and Headphone Handbook*, 2nd ed. (Focus Press, Oxford, UK, 1994).

[11] F. A. Everest, *The Master Handbook of Acoustics*, 4th ed. (McGraw-Hill, New York, 2001), pp. 404–405.

[12] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, 1968), pp. 555–572.

[13] S. E. Olive, P. L. Schuck, J. G. Ryan, S. L. Sally, and M. E. Bonneville, “The Detection Thresholds of Resonances at Low Frequencies,” *J. Audio. Eng. Soc.*, vol. 45, pp. 116–128 (1997 Mar.).

[14] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, 4th ed. (Academic Press, New York, 1997).

[15] Y. W. Lam, “An Overview of Modelling Tech-

niques for Small and Large Performance Spaces,” *Proc. Inst. Acoust. (UK)*, vol. 22, pp. 297–304 (2000).

[16] R. Walker, “Low-Frequency Room Responses. Part 1: Background and Qualitative Considerations,” BBC Research Dep., Rep. RD1992/8 (1992).

[17] J. B. Allen and D. A. Berkley, “Image Method for Efficiently Simulating Small Room Acoustics,” *J. Acoust. Soc. Am.*, vol. 65, pp. 943–950 (1979).

[18] C. L. Peckeris, “Ray Theory versus Normal Mode Theory in Wave Propagation Problems,” in *Proc. Symp. Applied Mathematics*, vol. II (1950), pp. 71–75.

[19] W. H. Press et al., *Numerical Recipes, the Art of Scientific Computing* (Cambridge University Press, Cambridge, MA, 1989), pp. 289–292.

## THE AUTHORS



T. J. Cox



P. D'Antonio



M. R. Avis

Trevor J. Cox was born in Bristol, UK, in 1967. He obtained a B.S. degree in physics from Birmingham University and a Ph.D. degree from the University of Salford.

Since the early-mid 1990s he has worked at both South Bank and Salford Universities, and as a consultant for RPG Diffusor Systems, Inc. Currently he is professor of Acoustic Engineering at Salford University, where he teaches room acoustics and signal processing at the undergraduate and postgraduate levels. His research centers on room acoustics, and his best known work concerns the measurement, prediction, design, and characterization of diffusing surfaces. His innovation of using engineering optimization methods to enable the design of arbitrarily shaped diffusers has enabled designs that can satisfy both aesthetic and acoustic requirements.

Dr. Cox is a coauthor of *Acoustic Absorbers and Diffusers: Theory, Design and Application*. He served as vice chair of the AES Subcommittee on Acoustics Working Group SC-04-02, is a member of the ISO/TC 43/SC 2/WG25 Working Group on scattering coefficients, and is associate editor for Room Acoustics for *Acustica/Acta Acustica*. He is a member of the Institute of Acoustics and the Audio Engineering Society.

Peter D'Antonio was born in Brooklyn, NY, in 1941. He received a B.S. degree from St. John's University, New York, in 1963 and a Ph.D. degree from the Polytechnic Institute of Brooklyn in 1967.

In 1974 he developed a widely used design for modern recording studios at Underground Sound Recording Studio, Largo, MD, utilizing a temporal reflection-free zone and reflection phase grating diffusers. He is the founder and president of RPG Diffusor Systems, Inc., established in

1983, and has significantly expanded the acoustical palette of design ingredients by creating and implementing a wide range of novel number-theoretic, fractal, and optimized surfaces, for which he holds many trademarks and patents.

Dr. D'Antonio has lectured extensively, published numerous scientific articles in technical journals and magazines, and is a coauthor of *Acoustic Absorbers and Diffusers: Theory, Design and Application* (Spon Press, 2004). He served as chair of the AES Subcommittee on Acoustics Working Group SC-04-02, is a member of the ISO/TC 43/SC 2/WG25 Working Group on scattering coefficients, and has served as adjunct professor of acoustics at the Cleveland Institute of Music since 1991. He is a fellow of the Acoustical Society of America and the Audio Engineering Society and a professional affiliate of the American Institute of Architects.

Mark Avis was born in Essex, UK, in 1970. He was sponsored by the BBC to study at Salford University, UK, and graduated with a B.Eng. degree (Hons.) in electroacoustics in 1993. After a time in acoustic consultancy, he returned to Salford to pursue postgraduate research in the area of active control with specific application to the modification of modal behavior of rooms at low frequency. His Ph.D. degree on this subject was awarded in 2001. He has maintained interests in related areas, such as low-frequency subjective perception and active acoustic diffusion. More generally, his research and teaching interests center on electroacoustics and transducer design, and he leads the master's program in Audio Acoustics at Salford. He is also active in vibration and structure-borne sound and when not working indulges these mechanical interests with decrepit east European motorcycles. He is a member of the Institute of Acoustics.