SOUND SYSTEM DESIGN

Marshall Long
Marshall Long/Acoustics
Sherman Oaks, California 91423

At the beginning of certain sound system design projects there is a moment of panic when the whole thing seems totally impossible. Usually the difficulties arise from a combination of a challenging acoustical environment and complicated client demands. At these times it is useful to take a deep breath and review the overall design objectives, which are relatively simple: (1) distribute direct sound evenly to the listening area, (2) provide adequate intelligibility, (3) deliver sufficient level, frequency response, and natural sound quality for the intended use, (4) leave the listener with the sense that the sound is coming from the source, (5) control feedback at the microphone positions, (6) avoid acoustical defects such as long delayed reflections, and (7) respect the architecture of the space.

> *"Vendors should publish the formulas and assumptions behind their software."*

Sound system design combines the arts of architecture, audio, and acoustical engineering and can be surprisingly complicated. The sound system designer is often asked to provide a solution in a less than ideal acoustic environment for a variety of potential uses. Figure 1 illustrates a good example of a challenging project. It is a sketch by the architects, Armando Ruiz and Associates, of their concept of the reconfiguration of a traditional gothic chapel at Mount Saint Mary's College, a small Catholic women's college in Los Angeles. The existing room had a cruciform shape with wooden pews facing the sanctuary. The building has 40-foot-high arched ceilings and is built of stone and concrete with unpadded wooden pews and tile floors.

It is being redesigned to accommodate services in the round in keeping with the Vatican II encyclical that encourages "full active and conscious participation in the liturgy." The existing pews will be removed, and a new seating configuration installed, which can be rearranged using moveable chairs, into any configuration to accommodate various uses. In addition to traditional Masses with presenters located at six different locations distributed throughout the space, the room will be used for lectures, working group presentations, choral singing, and small musical ensembles. Attendance can vary from 6 to 200 at any given time.

The transmission of speech over the existing sound system was difficult to understand, as was unamplified speech beyond about 20 feet. Organ and sacred music sounded wonderful over the existing sound system and the music director did not want to loose those special qualities. The architect wanted to preserve the existing visual ambience. Not surprisingly the budget was small and the time frame short. This project, which was completed in late 2006, serves as a good example of the audio design process.

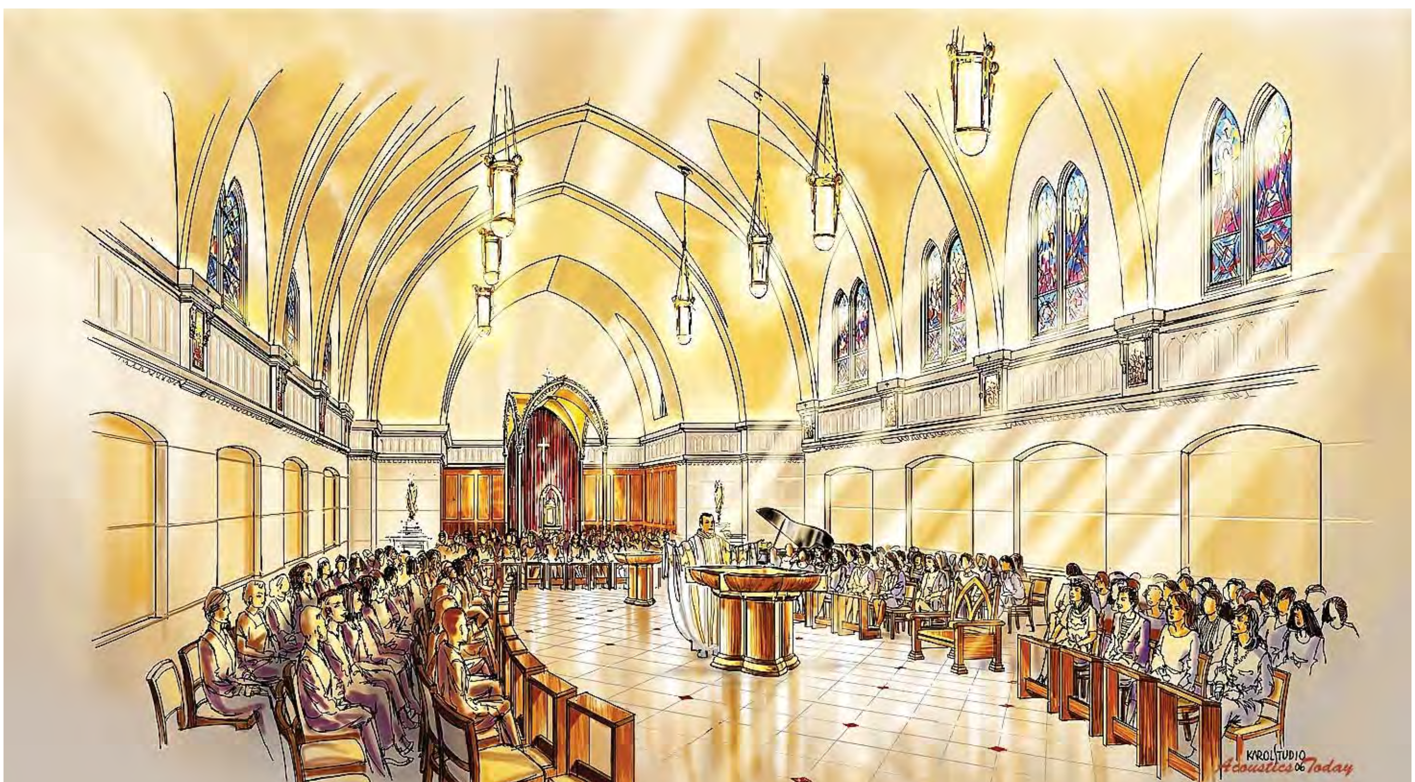

Fig. 1. Mount Saint Mary's College Chapel.

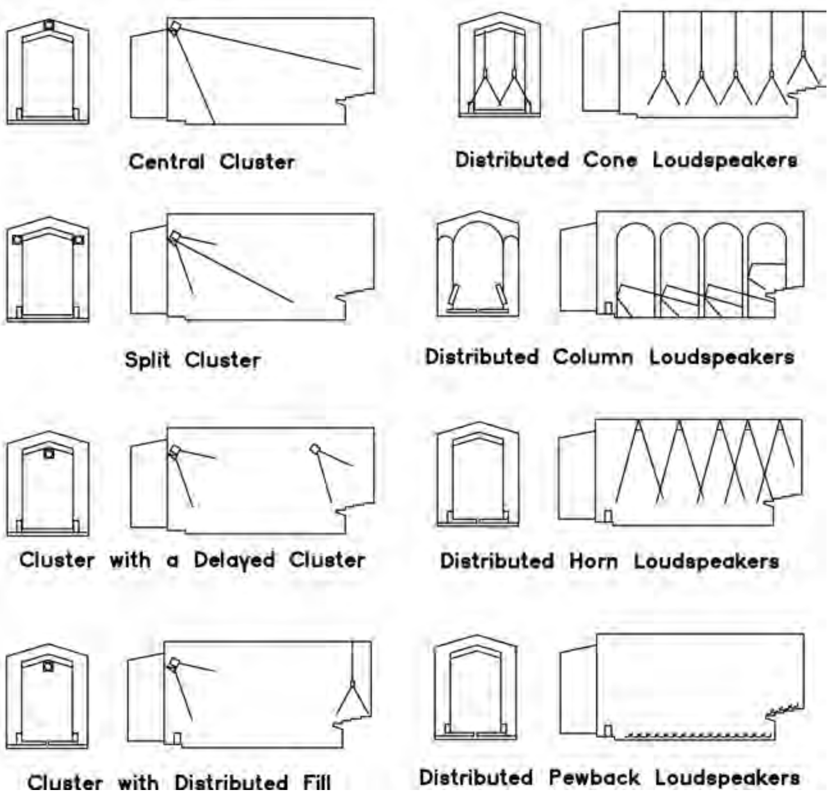

Fig. 2. General types of sound systems (Klepper, 1999).

The directivity of a cone loudspeaker at a given frequency depends on its diameter. The coverage (-6 dB) angle is about 90 degrees when the wavelength is equal to the cone diameter and narrower above that frequency. Home stereo systems use small diameter cones to obtain wide dispersion. Sound reinforcement systems use large area horns, line arrays, or distributed loudspeakers to obtain high direct field levels and good feedback control. Cone loudspeakers can now generate relatively high sound pressure levels with low distortion and excellent fidelity.

Systems used only for speech do not require large transducers. A good example is the telephone, where adequate intelligibility can be achieved in a modest space. Music systems, with their extended bass requirements, use larger loudspeakers, although in home systems bass drivers are rarely bigger than 10–12-inches in diameter. Distributed systems for music need 12-inch diameter loudspeakers, often coaxial, to cover the frequency range. Subwoofers, usually 15- to 18-inch cones, can be used for bass reinforcement where organ or other low frequency instruments are supported.

Many manufacturers offer high quality cabinet loudspeakers, with a combination of horns and cone drivers in a two- or three-way enclosure. Another approach uses a number of small two- or three-way cabinets, each having two or more cones bracketing a high frequency wave guide. These are linked together like tank treads and hung vertically in a line or J-shaped ribbon.

In the chapel it was decided to use 12-inch diameter coaxial loudspeakers in 2 cubic foot cylindrical enclosures. This combination offers excellent frequency response and provides enough bass for live music and choral singing. Where distributed systems are used, the bass is supported by the combined area of many loudspeakers so subwoofers may not be required. Where necessary, 18-inch drivers in individual boxes can be added.

Design approaches

Sound system design is handled using three basic tools: (1) loudspeaker selection, placement, and orientation, (2) electronic component selection and utilization, and (3) room acoustics. Thus, it is important for the designer to be able to control of each of these tools. There are a number of approaches to the loudspeaker layout that can be used. Several are shown in Fig. 2 and can be summarized as: (1) use one or more groups or clusters of loudspeakers to cover the room from a position above the focus of interest, (2) use multiple distributed (usually overhead) loudspeakers positioned relatively close to the listeners, and (3) use a combination of 1 and 2 with electronic delays.

In this church, a cluster solution was inappropriate since it would be unsightly and inflexible when the source origination point was moved. Similarly loudspeakers mounted on the side walls were a poor choice, since it would not be possible to control the perceived direction of the sound source. Thus an overhead distributed system remained the best approach. It would provide even coverage, with minimal visual intrusion, while offering enough flexibility to maintain source localization.

Sound quality

The quality of a sound system is something of an aesthetic choice. The simplest aspects—level and frequency range—have much to do with the size and type of loudspeakers. Larger loudspeakers extend lower in frequency, can handle greater power, and provide more directional control than smaller loudspeakers. Cone loudspeakers sound more realistic than horns. Horns are more efficient and have better pattern control than cones.

Modeling

Most designers use a computer model to assist with the loudspeaker layout. This has led to the plaintive cry, heard after installation, "But the computer said it would be OK," which is why it is important to look behind the curtain and ask the Wizard of 0's (and 1's) what is in his secret code. Vendors should publish the formulas and assumptions behind their software. Otherwise designers have no clue about what they are calculating. A computer model locates and orients the loudspeakers, a three dimensional coordinate system. An individual loudspeaker is characterized mathematically by its sound power, as a function of frequency, and its directivity, as a function of both frequency and direction. The direct field levels are calculated at various receiver locations, based on direction and distance from the loudspeakers. The power and directivity data, furnished by the loudspeaker manufacturer, are calculated from sound pressure level measurements, taken on the surface of an imaginary sphere surrounding the device. Based on these measurements, manufacturers publish the directional properties in octave or third octave bands, expressed as the on-axis sensitivity (the sound pressure level at 1 meter for a 1 watt electrical input), and the change in level in directions other than on-axis. Angular increments of 5 or 10 degrees are most commonly used. Programs interpolate level values between the measurement points. Directivity data, given in standard unencrypted ASCII files, are the most useful since it allows the raw data to be examined directly.

When loudspeakers are arranged in lines or clusters, they should be treated as individual sources by the model. When the sound fields from two loudspeakers overlap, there is an interaction due to the differences in phase, which creates local increases and decreases in the received signal. In these cases the signals from each loudspeaker must be combined by taking both their level and relative phase into consideration. Direct field levels should be calculated in this manner when the sound pressure levels of overlapping sound fields are nearly the same.

It is also important to use the sound power levels of individual loudspeakers to calculate the reverberant field contribution, since the phase relationships are not maintained at large distances or after many reflections. Programs that bundle groups of loudspeakers together and treat them as one source, do not accurately calculate their reverberant field contribution since the reverberant field sound power level is underestimated.

Coverage

Calculating the direct sound level at each receiver is one of the main objectives of a computer model. Direct field data can be presented numerically as sound pressure levels at each receiver location, or can be displayed graphically. The numeric approach has the advantage of allowing adjustment of the gain of individual loudspeakers to achieve the most even coverage. A calculation of the standard deviation of the direct field level within the intended coverage area is a useful measure of evenness (usually plus or minus 2 dB). Programs should, at a minimum, calculate levels in the 500 Hz, 1 kHz, and 2 kHz bands. The 500 Hz band is particularly important for feedback control since loudspeakers tend to be less directional in this band.

The final design in the Mount Saint Mary's Chapel is shown in Fig. 3, a drawing of the floor plan with the loudspeakers superimposed. Normally distributed loudspeakers are positioned at an elevation approximately equal to their spacing but preferably not farther than 20-25 feet above the listener. In this design they were pointed straight down, but that is not always necessary. Even where the ceilings are slanted, distributed loudspeakers can be used, although the spacing depends on the pitch of the ceiling. In this church computer calculations yielded standard deviations of 0.8, 1.1, and 1.4 dB in the 500 Hz, 1 kHz, and 2 kHz bands, which is satisfactory.

Intelligibility—Liveness

Even coverage is not the only goal. Don Davis once observed that even coverage can be obtained by pointing all the loudspeakers at the ceiling. The next concern in the design is intelligibility, clearly the sound should be intelligible, and this requires pointing the loudspeakers at the listeners. One of the traditional measures of intelligibility is the number of consonant-vowel-consonant syllables misunderstood. In early studies listening tests were carried out using a group of spoken words in a neutral carrier sentence. The intelligibility was expressed in terms of badness, that is, loss of intelligibility, instead of goodness. Intelligibility metrics were studied at Bell Labs in the 1920's and 30's using a single source and receiver in listening tests.

Later the intelligibility was expressed in terms of badness, that is, loss of intelligibility in a measure called Liveness.

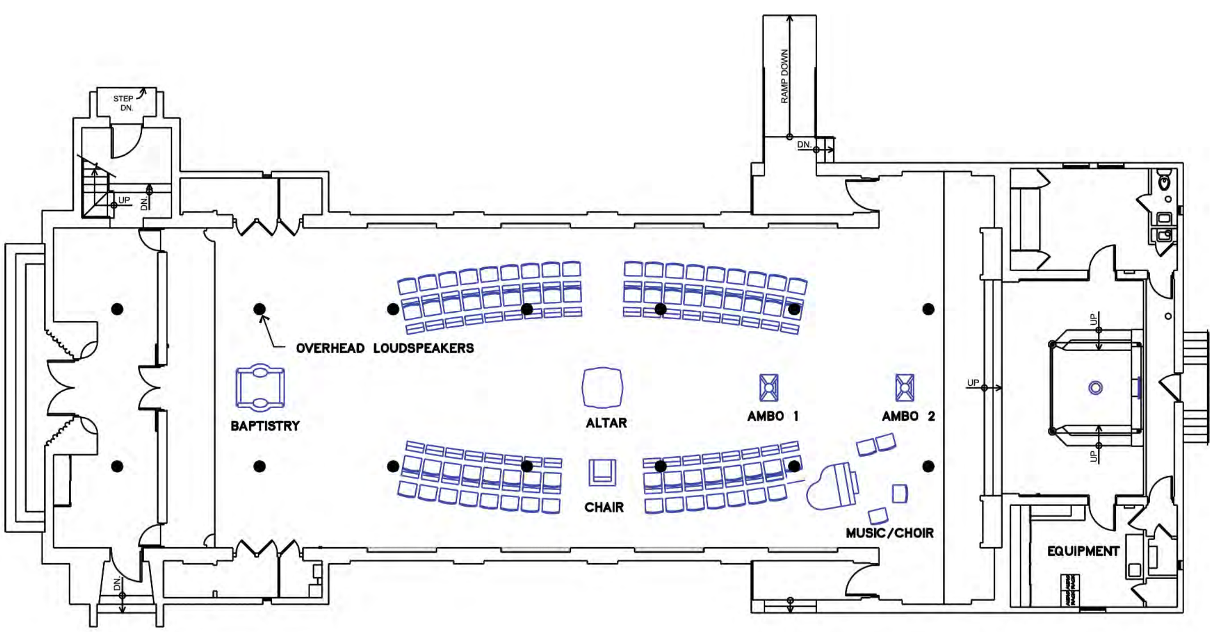

Fig. 3. Floor plan with loudspeaker layout.

Maxfield and Albersheim, (1947) wrote about the Liveness in terms of the ratio of the reverberant-to-direct energy densities times the reverberation time for a source having a directivity of one.

$$L = \frac{T_{60}\,D_r}{13.82\,D_d} = 22.6\,\frac{T_{60}^2\,r^2}{V} \tag{1}$$

where

D_r = reverberant field energy density (W sec/cu m)

D_d = direct field energy density (W sec/cu m)

T_{60} = reverberation time, or the time for sound to decay by 60 dB in a room (s)

V = volume of the room (cu m)

r = distance between the source and receiver (m)

The familiar reverberation time in metric units is defined in the usual way as

$$T_{60} = 0.161\frac{V}{A} \tag{2}$$

where

A = total area of room absorption (metric sabins)

$= S_1\alpha_1 + S_2\alpha_2 + S_3\alpha_3 + \cdots + S_n\alpha_n$

S_i = surface area of material i (sq meters)

α_i = fraction of the sound absorbed by material i

Intelligibility—Articulation loss

In 1971 Peutz published a formula similar to that previously found, for the percentage articulation loss of consonants or %ALcons. In metric units an equation can be written in terms of the properties of a room for a single source, having a directivity of 1.

$$\%ALcons = 200\,\frac{T_{60}^2\,r^2}{V} \tag{3}$$

It is intuitive that when the receiving room is highly reverberant, speech is more difficult to understand. The standard practice is to use the 2000 Hz octave band to calculate the articulation loss of consonants. Values less than 5–10% are considered good.

Intelligibility—Signal to noise ratio

While there are many ways of measuring intelligibility, they are all based on some sort of signal-to-noise ratio. Different metrics use different definitions of what constitutes the signal and what constitutes the noise. The consensus is that the direct signal (or most of it) is good for intelligibility and the reverberant noise (or most of it) is bad for intelligibility, and a ratio of the two is a measure of how good or bad the intelligibility is.

Notice that noise, other than the reverberant field, is left out of this discussion to make things easier, even though it, too, is bad.

The articulation loss formula is relatively easy when there is only one source. Early pioneers converted levels into distances to make the comparison easier. When there are multiple loudspeakers of different types, with different gains, distances become awkward to use. Under these conditions, the simple form of the articulation loss equation can no longer be used.

Fortunately %ALcons can also be expressed in terms of a signal to noise energy ratio, or rather the noise to signal energy ratio times the reverberation time (Bistafa and Bradley, 2000). This allows us to apply it to complex sound systems.

$$\%ALcons = 8.9\left[\frac{T_{60}\,(\text{Reverberant Noise Energy})}{13.82\,(\text{Direct Signal Energy})}\right] \tag{4}$$

There are more complex versions of this relationship that include extraneous noise and extend the formula to long distances, but for our purposes this will suffice.

Another approach is to use raw signal to noise ratios in several octave bands. For our purposes the signal is the combined direct field level from all loudspeakers and the noise is the combined reverberant field level from all loudspeakers. With these assumptions

$$\text{Direct-to-Reverberant Level} = 10\log\left[\frac{\sum \text{Direct Signal Energy}}{\sum \text{Reverberant Noise Energy}}\right] \tag{5}$$

In this metric the signal and the noise are expressed as energies, both of which are steady state values. There is no consideration of when the signals arrive or how they have reflected. We rely on the relative levels to sort this out. Loudspeakers contribute to the useful energy only to the extent that the levels they produce affect the direct field level at a particular receiver.

Table 1 (Long, 2006) shows a chart of steady state signal-to-noise ratios which can be used to judge how intelligible a system will be. The comparisons are made in three octave bands centered at 500, 1k, and 2k Hz.

In almost all locations in a room the calculated reverberant field levels are higher than the direct field levels, but even in these areas, good intelligibility can be achieved. By comparing signal to noise ratios in several octave bands we have a useful tool for sound system design.

As part of the computer modeling the reverberation times and the signal to noise ratios were calculated at Mount Saint Mary's Chapel. In the existing church in the empty condition, the mid frequency reverberation times were around 4 seconds, much too high for optimal results. Even with the distributed loudspeakers, the signal-to-noise ratios were about -10 dB, which yielded only fair intelligibility. As a result it was decided to add some absorption to the room. Normally it is best to do this in the high ceiling areas, where surfaces do not provide useful reflections for envelopment or intelligibility. In this room those surfaces were not available, so panels were designed to fit into the niches of the side walls, as shown in Fig. 4. These reduced the calculated reverberation times to about 1.7 seconds and the signal to noise ratio to -6 to -7 dB, both satisfactory values. This calculation confirmed the viability of the loudspeaker layout.

Intelligibility—Arrival times

Many have argued that not all the direct field sound energy is good and not all the reverberant field energy is bad. This has led to intelligibility metrics that take the signal arrival times into account. In these metrics all sound arriving before a given time, after the first sound, is good, and all sound arriving after that time is bad. Not infrequently the first arrival is not the loudest sound especially when there are multiple sources: a talker on stage, a point source in the face of the stage, a central cluster, and perhaps a delayed distributed system plus reflections. Consequently some judgment must be exercised in selecting the cutoff time.

As we can see, there are at least two ways of parsing the signal-to-noise: (1) by level, with the lower level contributions having less influence, and (2) by time, with later arrivals counted as noise. The second approach, using metrics that integrate (add up) the energy arriving at a receiver over time, requires a great deal more calculation than a static model. It is not clear that the added information is worth the extra calculation. Thus I prefer the first method.

To summarize we have a choice of ways of controlling intelligibility. We can raise the direct signal level at a receiver by using high directivity loudspeakers or we can use many distributed loudspeakers placed close to the audience. We can also decrease the reverberant field level by lowering the level produced by individual (distributed) loudspeakers or by using directional loudspeakers that contribute lower levels of sound power to the reverberant field. Finally we can add absorption to the receiving space to decrease the reverberant field as well as the reverberation time.

Perception of direction

The perception of direction is another important design consideration. It is important to maintain the illusion that the amplified sound is coming from the original source. This is accomplished by using the human reaction to the first arrival sound, which determines the perceived direction, even when the later arriving sound is louder. The phenomenon is dependent on the relative level and time delay. Normal

Table 1 Direct-to-Reverberant Sound Levels for Speech Intelligibility in Rooms

Direct-to-Reverberant Level (dB)	Intelligibility
> -3	Excellent
-3 to -6	Very Good
-6 to -9	Good
-9 to -12	Fair
-12 to -15	Poor
< -15	Very Poor

design practice is to control the time of the second arrival so that it occurs about 5 milliseconds after the first. This preserves the directional cue obtained from the first arrival.

In a cluster design we do this by locating the loudspeakers above the source, taking advantage of the fact that our perception of direction is less sensitive in the vertical plane than in the horizontal plane. We can also add localizing loudspeakers near the sound source, and delay the signals fed to the other loudspeakers. The illusion can be preserved even though the delayed loudspeakers may be contributing more to the overall level at a given receiver.

In small rooms such as the chapel, the talker can be its own localizing source. In larger rooms point sources must be added in the face of the stage in auditoria or in the face of steps in worship spaces. They are relatively easy to disguise in these locations. Localizing loudspeakers do not have to reproduce the full bandwidth of sound. Step mounted transducers are limited to 6-inch diameter cones in a 7-inch-high step face. Delays can be fixed by calculating the propagation time and adding 5 milliseconds. After installation they can be adjusted by measurement or by ear. It is not difficult to set delay times by ear simply by listening for changes in perceived source direction. In fact, differences of one or two milliseconds are quite perceptible.

Another important consideration in the design of delays is to provide a smooth transition as the listener moves away from the source. The loudspeakers covering a given area must not only provide even level coverage but also consistent source imaging. Often distributed overhead loudspeakers are

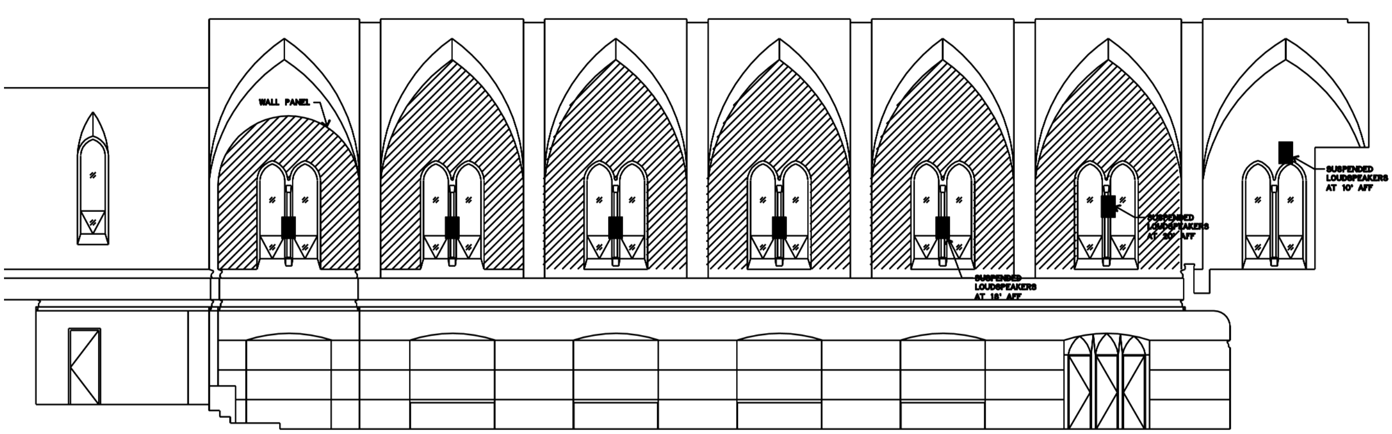

Fig. 4. Elevation showing loudspeakers and wall panels.

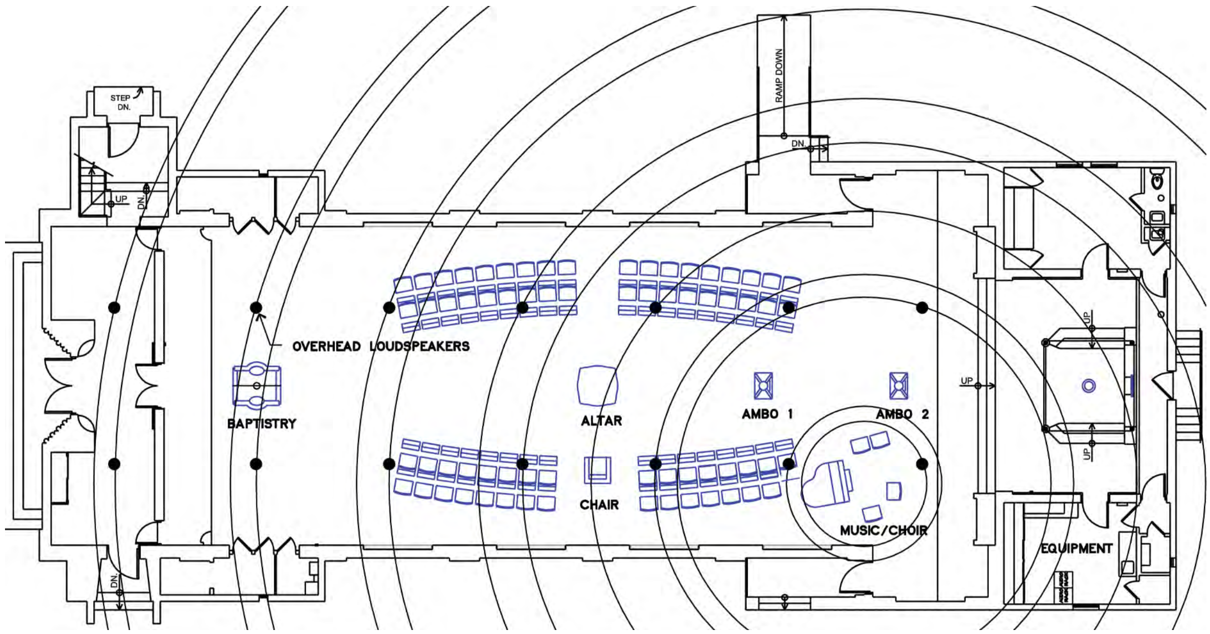

Fig. 5. Floor plan with music delay zones.

the best solution. If the loudspeaker heights are uniform and the loudspeaker type is the same, delays are increased as the listener position moves farther from the source. In this way, loudspeakers covering an area in front of a listener can provide additional directional cues to the area behind. This is an excellent approach in the case of large reverberant churches.

Another consideration in sound system design is that the source origination point, whether speech or music, may vary depending on the source position. For example a talker located at an altar will provide a different source origination point than a choir. Using individually amplified distributed loudspeakers, different delays can be employed that depend on the source origination point. In this way even coverage can be maintained while the perceived source direction is changed.

At Mount Saint Mary's Chapel the room was small enough that the talker could be the localizing source. For each talker location, the delays were set for each distributed speaker. Figure 5 shows a typical configuration. The design is much like dropping a pebble in a pond and watching the waves move outward from the origination point. By locating microphone plug boxes in the floor near the originating point, each box would uniquely define the associated delay pattern.

Feedback control

In the design of every sound system we must address the phenomenon of feedback. Figure 6 shows the geometry of a simple sound system having a talker, a microphone, and a loudspeaker. Feedback occurs when the direct sound level from a loudspeaker exceeds the sound level from the talker at the output of the microphone. Under these conditions a feedback loop is created which produces a howl in the system. To avoid this condition, the talker level must exceed the loudspeaker level by a certain margin of safety called the feedback margin of stability

$$L_{T\text{-}M} - L_{H\text{-}M} - D_M(\theta) > L_{FMS} \text{ for no feedback} \tag{6}$$

where

$L_{T\text{-}M}$ = direct field level produced by the talker (T) at the microphone (M), (dB)

$L_{H\text{-}M}$ = direct-field level produced by the loudspeaker system at microphone, (dB)

$D_M(\theta)$ = directivity index of microphone in the direction of loudspeaker relative to talker (usually negative), (dB)

L_{FMS} = feedback margin of stability (about 10 dB)

From this we see the ways to control feedback:

1) Move the talker closer to the microphone so that the system gain can be reduced and thus the level at the microphone from the loudspeaker;
2) Select a directional microphone that preferentially emphasizes sounds coming from the talker, relative to the loudspeaker;
3) Design loudspeakers that deliver more sound to the listeners than to the microphones, either by using directional loudspeakers or a distributed system;
4) Use equalization, frequency shifting, compression, and other electronic techniques.

At Mount Saint Mary's Chapel all of these techniques were utilized. The system was equalized for music with a house curve which rolled off 4 dB per octave above 3k Hz. Additional equalization was included in the speech path, tuned to the major feedback frequencies. Cardioid microphones were used at the fixed talker positions. Earset microphones were used for the wireless transmitters. A frequency shift of 10 cents (1200 cents/octave) was also electronically introduced as a safety measure.

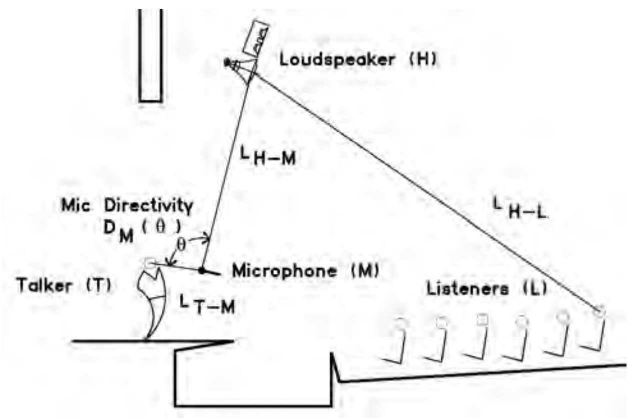

Fig. 6. Basic parameters for feedback control.

Sound system setup

The setup and adjustment of a sound system after installation is rarely discussed but is critically important. When the electronics have been installed, it is usually necessary to spend at least several hours, and with a larger system, a couple of days, to set up the system. In general, the following steps are recommended:

1) Excite each group of like-type loudspeakers independently with pink noise and adjust the third octave equalizer assigned to that group to be flat (1 dB) from 100 Hz to 3 kHz. When there are subwoofers the low frequency limit should extend down to 80 Hz. Roll off the high frequencies about 4 dB per octave above 3 kHz. For concert venues, season to taste;
2) Measure the spatial coverage throughout the receiving space separately using pink noise in the 500 Hz, 1 kHz, and 2 kHz octave bands. It should match the predicted coverage levels, usually no more than 2 dB and preferably less. Adjust the level for each loudspeaker to achieve this balance. Note that a distributed system requires a separate amplifier channel for each loudspeaker. While theory predicts that the reverberant field will be uniform throughout the space, this is seldom the case;
3) Set the delay times for each delay zone and each source position so that the sound appears to come from the origination point at every seat. This must be tested using a talker at each microphone position. A convenient way of testing this effect is to have a talker walk toward a high quality fixed microphone while talking. As the person enters the field of the microphone the source image should remain fixed on the talker. The test must be repeated for each zone and again with all speaker zones on;
4) Play a familiar piece of music through the system. The music should contain a good mix of high-, mid-, and low-frequency energy. I have used Jennifer Warnes' recording of "Bird on a Wire," and "Stay," by Allison Krauss, but every engineer has his or her own favorites. Make the final adjustments to the music equalizers by ear;
5) Using a talker, increase the system gain at each microphone position until the initiation of feedback. Measure the frequency at which the ringing occurs. Increase the equalizer notch depth at that frequency using a narrow band or parametric filter for that microphone type and position. When sufficient gain before feedback has been achieved, review the equalizer settings to make sure they are not too extreme and do not affect the naturalness of the sound;
6) Add about 10 cents of frequency shift to the speech feed as added feedback insurance;
7) Save all system settings on a computer disk and make backup copies.

Conclusions

There are a number of approaches to sound system design. The recent trend has been to utilize large clusters, often line arrays to cover the audience area. Where very high sound pressure levels are not required, good source localization, excellent intelligibility, low feedback, and a discrete appearance can be maintained using a distributed overhead system.

Figure 7 shows a photograph of the completed installation at the Mount Saint Mary's College Chapel. The loudspeakers were positioned so that they complemented the existing light fixtures. The approach also allowed us to limit

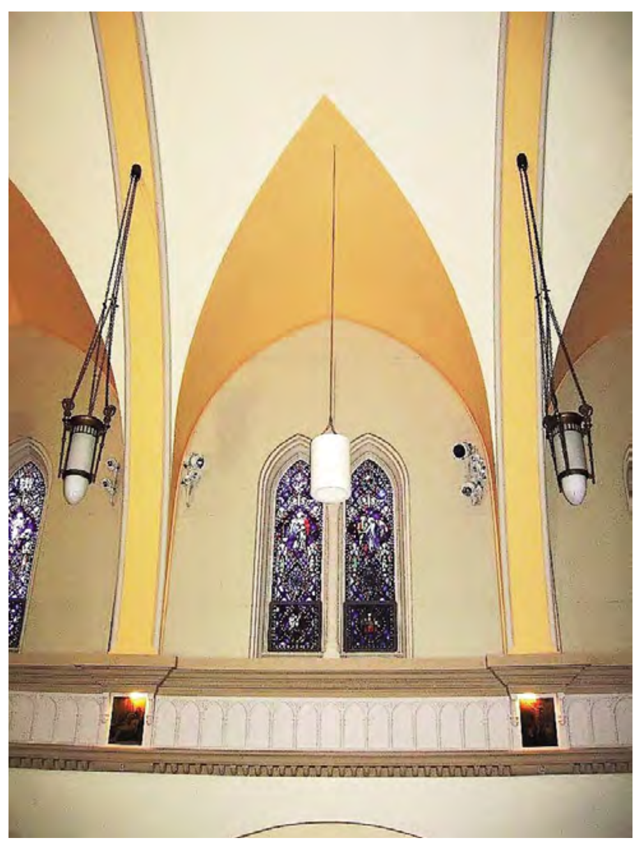

Fig. 7. Completed installation.

the amount of absorption added to the room so the reverberation time remained long enough to provide a suitable environment for choral singing. Speech intelligibility and the music environment were both quite satisfactory.**AT**

References

Bistafa, Sylvio R., and Bradley, John S. (**2000**). "Revisiting algorithms for predicting the articulation loss of consonants AL_{cons}," J. Audio Eng. Soc. **48**(6), Fig. 2, p. 537, and Fig. 4, p. 539.

Klepper, David L. (**1999**). "A different angle." Sound & Video Contractor **17**(1)

Long, Marshall (**2006**). *Architectural Acoustics* (Academic Press,), Table 18.3, p. 619.

Maxfield, J. P., and Albersheim, W. J. (**1947**). "An acoustic constant of enclosed spaces correlatable with their apparent liveness," J. Acoust. Soc. Am. **19**(1), 71-79.

Peutz, V. M. A. (**1971**). "Articulation loss of consonants as a criterion for speech transmission in a room," J. Audio Eng. Soc. **19**(11).

Marshall Long received a BSE from Princeton University in 1965, attended the University of Grenoble in France and the University of Madrid in Spain in 1966, and received MS and Ph.D. degrees in engineering from University of California-Los Angeles (UCLA) in 1971. While still a graduate student, he founded his own consulting firm now in its 37th year. Marshall Long/Acoustics specializes in architectural acoustics, audiovisual design, noise and vibration control, and other technical areas related to acoustics. He enjoys sailing, judo, soccer, reading, and writing, and is living with his family in Sherman Oaks, CA.