Acoustics and musical qualities[a]

Leo Beranek
975 Memorial Drive, Suite 804, Cambridge, Massachusetts 02138

(Received 2 November 1995; accepted for publication 14 February 1996)

This study examines the known acoustical attributes of concert halls and depicts the ways in which they interrelate with the music performed in them. The first section distinguishes between direct sound, early sound, and reverberant sound. The second section concerns how musical factors and acoustical attributes affect vertical and horizontal definition of the music. Figures illustrate the ways that the sounds of fast and slow music are altered by different acoustical conditions. The final section discusses the important acoustical attributes, intimacy, spaciousness, timbre, and tone color.

PACS numbers: 43.10.Ln, 43.55.Gx, 43.55.Hy

I. BASIC ASPECTS OF SOUND IN HALLS

A. Direct sound

Suppose you are seated close to the stage of an outdoor stadium or in a sound-deadened room where there is no echo or reverberation. When a performer near you plays a tone, you hear the tone precisely as the instrument has produced it. The buildup of the tone, its loudness, and the termination that you hear are exactly characteristic of the instrument and the player. The tones of wind instruments cease almost immediately after the excitation is stopped. The tones of stringed instruments persist a little longer because of the vibrations of the strings and the wood. The sounds from drums and freely vibrating piano or harp strings may persist considerably longer unless damped by the player. In any case, the sound that reaches you is only the direct sound. There is no enclosed space, hence, no surfaces to provide reflections that would follow the direct sound.

In a concert hall or opera house, the sound that first reaches the listener before any reflections arrive from the walls and ceiling is called the "direct sound."

B. Early sound, early sound decay, and reverberation time

1. Early sound

The sound transmitted from an orchestra is radiated in all directions and travels through the air at about 1128 ft (344 m) per second and within 1 or 2 s is reflected many times over from the different surfaces of the space. To understand the effect of the acoustical attributes of a hall on the music, we must consider the reflections as divided into two time intervals. First, the "early sound," defined as the direct sound and those reflections that take place within 80/1000 of a second (80 ms) after the arrival of the direct sound. Second, the reverberant sound that is created by the many reflections that occur subsequently.

These reflections come from the side walls, the ceiling overhead, and the walls of the stage enclosure. If they arrive at the listener's position from side walls, that is, from lateral directions, they appear to broaden the source and thus increase the apparent source width (ASW). An increase in ASW lends quality to the music heard in a concert hall. Lateral reflections are easy to achieve at listeners' positions in the center of the main floor of a narrow hall with parallel side walls. But in a fan-shaped hall they are directed to the rear seats only, unless special sound-reflecting panels are attached to the walls (see Fig. 1). Many of the famous old halls, such as those in Boston, Vienna, and Amsterdam, are "shoebox" in shape, which means parallel side walls and horizontal ceiling.

2. Reverberation time (RT)

Each time the traveling sound wave bounces back from a room surface, its strength is weakened by the "sound absorption" of that surface. Thus the sound created by the succession of reflections dies down with time and is said to "decay." The time in seconds that it takes for a loud sound to decay 60 dB, that is, to decay to inaudibility after being cut off, is called the "reverberation time" (RT). RT is usually determined separately at a number of frequencies, such as, 125, 250, 500, 1000, 2000, and 4000 Hz.

3. Early decay time (EDT)

To add to our baggage of new terms, the first 10 dB of the sound decay after a source is cut off is called "early decay time" (EDT), not to be confused with "early sound." EDT becomes very important when intercomparing the acoustical quality of halls.

C. Loudness

Loudness is a function of the energy in a sound divided by the number of people who must share it. The energy is

[a]This is a very slight revision of Chapter 3 of the new book *Concert and Opera Halls: How They Sound* (Acoustical Society of America, Woodbury, NY, 1996). The complete book is available from the Society (see publication list and order form in this issue).

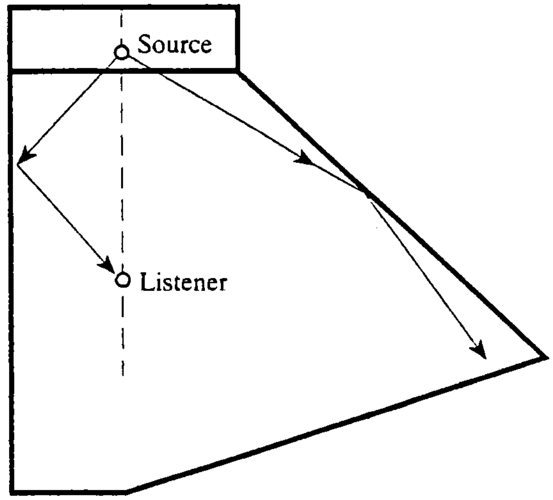

FIG. 1. The left half of the sketch shows how a rectangular hall reflects early sound from a source on the stage laterally to the audience on the main floor. The right side shows how a fan-shaped hall reflects sound to the far corners of a hall, so that early reflections may arrive at main floor listeners only from the ceiling, thus depriving those listeners of early lateral reflections.

also decreased by the absorption caused by the people, furnishings, and materials. Loudness at a particular listener's ears is also divided into two parts, early and reverberant. ''Loudness of the early sound'' is determined by the energy of the sound that comes directly from the source, plus the energy received in the next 80 ms from the early reflections. ''Loudness of the reverberant sound'' is defined by the total sound energy that reaches the listener 80 ms *after* the process of decay has begun.

II. INTERRELATIONS BETWEEN ACOUSTICAL ATTRIBUTES AND MUSIC

A. Fullness of tone and liveness

In a concert hall, when a tone is sounded and is then suddenly turned off, the reverberation is heard for about 1.5 to 2.2 s. Two aspects of reverberation combine to increase the ''fullness of tone'' of an instrument or ensemble at a particular listener's position. The first is the length of the reverberation time (RT), and particularly the length of the early part of the sound decay (EDT). The term ''liveness'' is related to RT alone. The second ingredient of ''fullness of tone'' is the ratio of the loudness of the ''reverberant sound'' to that of the ''early sound.''

Reverberation is not in itself desirable or undesirable; it is one of the components available to the composer (and the performer) for producing a musical effect and as such is actually a part of the music. Some styles of music depend on the tying together of successive tones to produce their overall effect. Many early choral compositions, particularly plainchant, were written to be performed in reverberant cathedrals and require considerable fullness of tone. When these compositions are performed in an acoustically dead environment, they suffer from want of body and lose much of their power. Early composers of organ music often wrote pauses into the music to emphasize the after-ring of the church.

When a hall is designed to guide the sound from the performers directly to the audience, most of the energy produced by the instruments reaches the listeners almost simultaneously with the direct sound, and the tone will lack fullness. If the sound of the music rises freely into the upper reaches of the room, the energy in the early sound will be reduced and the energy in the reverberant sound increased. The interior shaping of the surfaces and the inclusion of special sound-reflecting panels can be combined to control the proportions of these two factors, and, thus, the degree of fullness.

B. Definition or clarity

When a musician speaks of ''definition'' or ''clarity,'' he means the degree to which the individual sounds in a musical performance stand apart from one another. Definition depends critically on musical factors and the skill and intention of the performers, but it is also closely tied to the acoustics of the room. There are two kinds of definition: horizontal, which applies to tones played in succession, and vertical, in which tones are played simultaneously.

''Horizontal definition'' refers to the degree to which sounds that follow one another stand apart. The composer can specify certain musical factors that determine the horizontal definition—such as tempo, repetition of tones in a phrase, and the relative loudness of successive tones. The performer can vary the horizontal definition by the manner chosen to phrase a passage.

The acoustical factors that affect horizontal definition are the length of the reverberation time and the ratio of the loudness of the early sound to that of the reverberant sound—the same two factors that determine fullness of tone, but in inverse relation. That is to say, an increase in horizontal definition goes hand in hand with a decrease in fullness of tone.

''Vertical definition'' refers to the degree to which sounds that occur simultaneously are heard separately. Vertical definition depends on the score, the performer, the acoustics of the room, and the auditory acuteness of the listener. The composer specifies the vertical definition by choosing simultaneous tones and their relation to the tones surrounding them (whether the composition is hymnlike, chordal, contrapuntal, or simply an accompanied melody), and his choice of instruments on which they are played. Performers can alter the vertical definition by varying the dynamics of their simultaneous sounds and through the precision of their ensemble.

Vertical definition is also responsive to such acoustical factors as balance among the sounds of the various instruments as they reach the audience, the degree of blending of the tones of the different instruments in the stage enclosure, the relative response of the hall at low, middle, and high frequencies, and, once again, the ratio of the energy in the early sound to that in the reverberant sound.

Definition, both horizontal and vertical, results from a complex of factors, both musical and acoustical—a certain piece of music, played in a certain way, in a certain environment. The degree of definition that the composer intended is necessary in order for the music to be communicated faith-

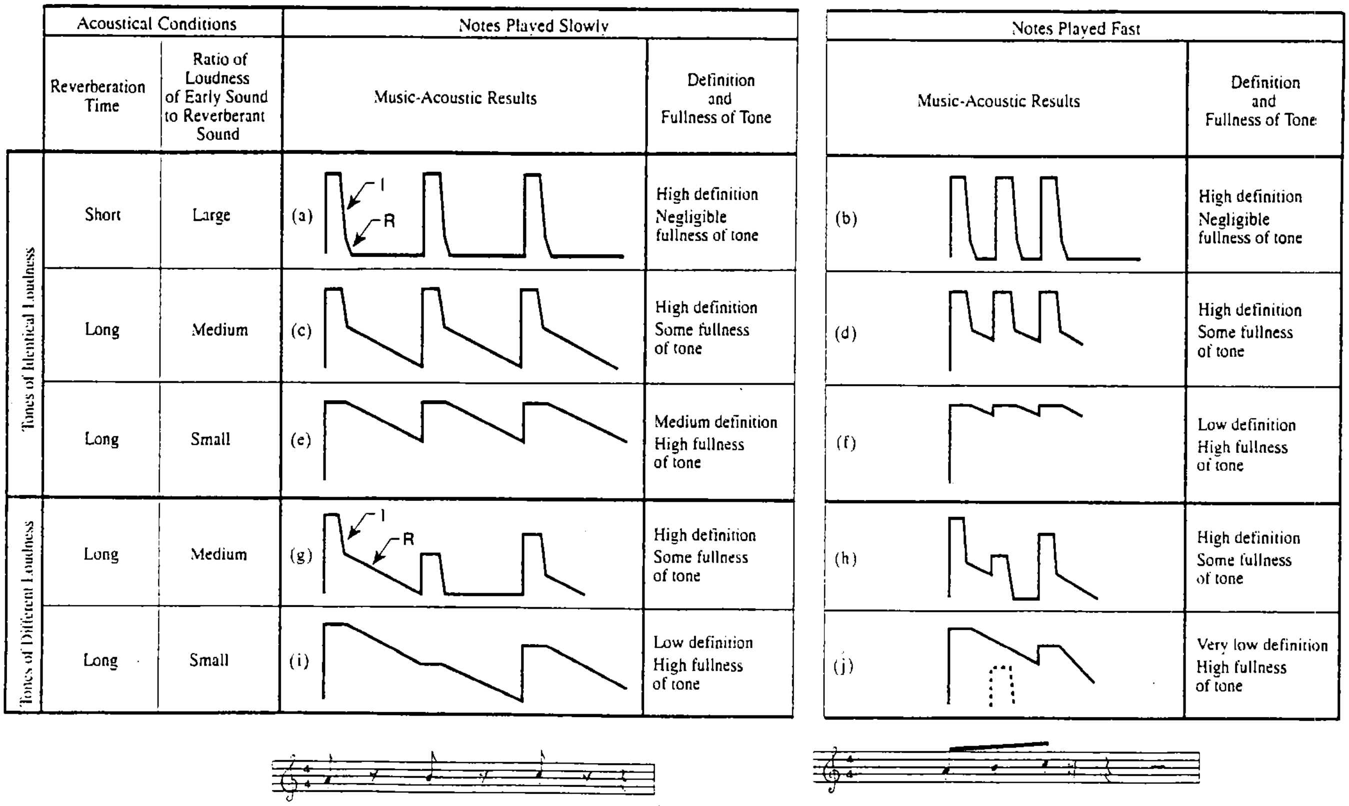

FIG. 2. Chart illustrating the interrelations among speed of music, reverberation time, and ratio of loudness of early sound to that of reverberant sound, and the music itself (tempos are identical). Here, I is the decay of the instrumental sound and R is the decay of the room sound.

fully to the audience. Gregorian chant—with its slow melodic lines that build and recede gradually, is performed with little horizontal definition; it is usually sung in a room with a very long reverberation time and a high ratio of reverberant to early sound energy. Bach's *Toccata in D Minor* for organ needs a reverberation time of at least 3 s in order for the full sonorities to be realized. At the other end of the spectrum, a piano concerto by Mozart—with its rapid solo passages and the delicate interplay of piano and different orchestral voices—needs considerable horizontal and vertical definition. It should be performed in a room with a relatively short reverberation time and a large ratio of early to reverberant sound energy. Mozart, after listening to a performance of *Die Zauberflote* from various locations in the hall and backstage, wrote in October 1791:

> By the way, you have no idea how charming the music sounds when you hear it from a box close to the orchestra—it sounds much better than from the gallery.

Mozart knew what later generations have confirmed, that his style of music sounds best where the ratio of the early to reverberant energy is large (e.g., at a box near the orchestra where the direct sound is stronger).

C. Speed of music relative to fullness of tone and definition (clarity)

Fullness of tone and definition or clarity are both related to the speed with which music is played. The chart of Fig. 2 shows the effect on music played slowly and music played rapidly of two different reverberation times and various ratios of the loudnesses of early to reverberant sound.

In (a) and (b) the individual tones of both the slow and fast music stand out distinctly, and the full extent of their attack and decay is discernible, because the reverberation time is very short and the ratio of early to reverberant sound energy is very large. In other words, there is negligible fullness of tone. Note that the rapid decay of the instrumental sound I is followed by the slower decay of the room sound R, because the reverberant sound arrives at a listener's ears after the early sound.

In (c) we see the effect on slow music of a long reverberation time and a medium ratio of early to reverberant sound energy. The full extent of the attack and part of the natural decay of the sound are heard before the reverberation takes over. A certain amount of the natural decay of the instrumental tone is buried in the reverberation.

In (d), because of the closer spacing of the tones in the fast music, both part of the attack and part of the natural decay of the tone are buried in the reverberation. The reverberation adds fullness to the tone, although the definition is still high.

In (e) the length of the reverberation time is the same as in (c), but the ratio of the loudness of early sound to reverberant sound is small. Here, part of the attack and all of the natural instrumental decay of sound are obscured by the reverberation. The tones are bonded, one to another, so that no one sound stands out clearly. When the speed of the music is increased, in (f), each tone is still observable, but most of the natural attack and the decay of the instrument are buried in the reverberation. The music is almost continuous. Definition is sacrificed to fullness of tone and there is no possibility of a staccato sound.

Let us see how successive tones of different loudnesses fare in these rooms (lower graphs). When the music is played slowly and the ratio of the loudness of early to reverberant sound is fairly large, as in (g), each tone is heard clearly. With the faster speed (h) definition is sacrificed, but the tones are still distinct. In (i), where the ratio of the loudness of the early to reverberant sound is small, the weakest tone (in the middle) is barely audible in the reverberation ''tail'' of the preceding tone. Finally, in (j), when the music is fast, the weakest tone is lost completely in the tail of reverberation of the preceding tone; the definition is very low, and the fullness of the tone is great.

The interplay of speed of music, definition, and fullness of tone illustrated in Fig. 2 is characteristic of music performed in concert halls. With a mental picture of the acoustical environment in which the piece is likely to be played, the composer may be able to create a composition with a particular degree of definition and fullness of tone in mind. The performer must then select a tempo and phrasing that seem to fulfill the composer's intent. And if the performance is presented in an acoustically appropriate hall, the audience is rewarded with music that follows the composer's conception accurately.

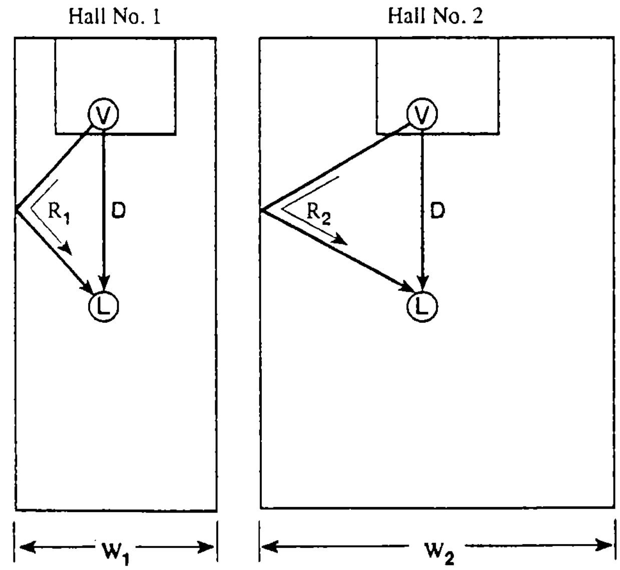

FIG. 3. Drawings showing the effect of hall width W on the difference between the length D and the length R_1 or R_2. The direct sound travels path D from the violinist V to the listener L. The reflected sound travels path R. The distance (R_2-D) in Hall No. 2 is longer than (R_1-D) in Hall No. 1. The ceilings of the two halls are assumed to be so high that reflections from the ceilings occur later than the reflections (as shown) from the side walls.

III. IMPORTANT INDICATORS OF ACOUSTICAL QUALITY

A. Intimacy

Acoustical ''intimacy'' suggests to the listener the size of the space in which it is performed. Different styles of music sound best in a hall with the appropriate degree of acoustical intimacy. It is not necessary that the room have a particular size, but only that it *sound* as though the size is appropriate. Acoustical intimacy is largely determined by the initial-time-delay gap—the difference in time of arrival of the direct sound and the first reflected sound—but also in part by the loudness of the overall sound since the listener assumes that a performance sounds louder in a small room than in a large one.

The initial-time-delay gap is illustrated in Fig. 3 which pictures the direct sound D traveling from a violin V on the stage to a listener L. If D equals 80 ft (24.4 m) the sound will take about 7/100 of a second, or 70 ms, to reach the listener. The first reflection occurs from a side wall and, as drawn here, will travel a distance of 97 ft (29.6 m), taking 85 ms to reach the listener. Thus the initial-time-delay gap for that listener would be 50 ms. In the best halls for today's symphonic repertoire, the initial-time-delay gap for a listener seated in the center of the main floor usually lies between 15 and 30 ms.

In every period composers have written chamber music to be played by small groups of instruments in small rooms. Between 1700 and 1900 most orchestral music was written for larger groups who performed in the larger but quite narrow concert halls of Europe, such as the Grosser Musikvereinssaal of Vienna. Opera is usually performed in relatively small horseshoe-shaped opera houses like La Scala in Milan or the Staatsoper in Vienna. Beginning in the sixteenth century composers have written choral and liturgical music for performance in both large cathedrals and smaller, less reverberant churches or chapels. These varied settings confirm that the composer conceives of each musical work with a particular degree of intimacy in mind. If the work is performed in a hall whose intimacy is not scaled to it, the listener is quickly aware of the inappropriateness of the acoustic environment. The organist E. Power Biggs once said, ''The listener immediately senses something wrong when he hears one of the organ works, such as those by Bach, played in a small college auditorium.''

Since chamber music has for almost three centuries been written for performance in rooms of similar acoustical characteristics—intimacy and high definition, imparting relatively low fullness of tone—little need be said about the acoustical design of such rooms. The following discussion concerns the acoustics of concert halls and opera houses seating more than 1500 persons, with occasional references to the acoustics of churches and cathedrals.

B. Spaciousness

A recent addition to the indicators of acoustical quality is the two-pronged attribute of spaciousness, composed of apparent source width (ASW) of the early sound and listener envelopment (LEV) by the reverberant sound. ASW is a major indicator of acoustical quality in concert halls.

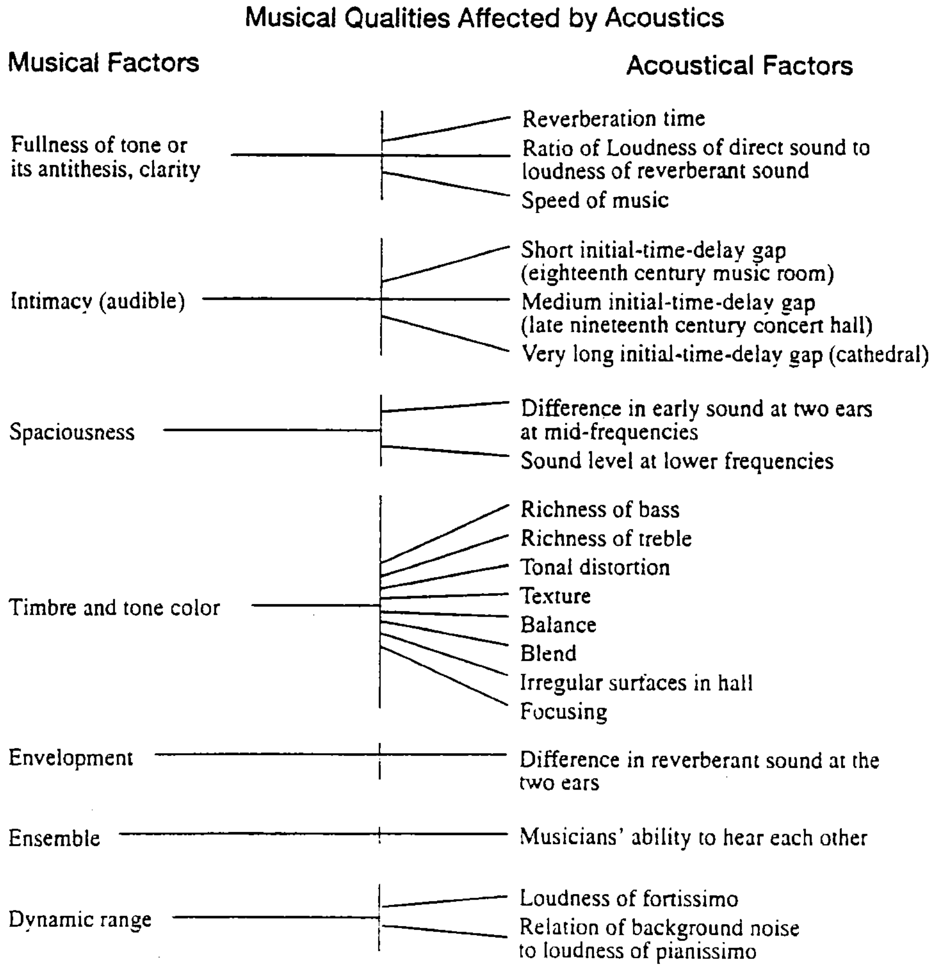

FIG. 4. Chart showing the interrelations between the audible factors of music and the acoustical factors of the halls in which the music is performed.

C. Apparent source width (ASW): the larger the better

At this writing, the best measure of ASW at mid-frequencies appears to be a quantity called "interaural cross-correlation coefficient" ($IACC_E$), where the subscript E indicates that only the early sound is considered in the calculation. This calculation measures the degree of dissimilarity of the musical sounds that reach the two ears. The less similar they are, the lower the value of $IACC_E$ and the greater the value of ASW. ASW is also increased by the level of the music at low frequencies, usually rated in decibels and designated $G_{\rm low}$.

Measurements have been performed in some of the world's best concert halls to set a standard for the desirable values of $IACC_E$ and $G_{\rm low}$, and these are discussed in later chapters entitled, "Spaciousness" and "Loudness."

D. Listener envelopment (LEV)

The reverberant sound that reaches the listener after 80 ms is most pleasant if the listener hears it coming from all directions. In Symphony Hall, Boston, a listener on the main floor hears the reverberation as filling the entire space above, ahead, and behind. By contrast, a listener seated in a steeply raked balcony will hear the reverberation as coming primarily from the front. At this writing, listener envelopment can best be estimated by visual inspection of the hall—including whether the sound has the freedom to surround the listener, and the degree of irregularities and ornamentation on side wall and ceiling surfaces and on balcony fronts.

E. Timbre, tone color, and bass warmth

"Timbre" is the quality of sound that distinguishes one instrument from another or one voice from another. "Tone color" describes balance between the strengths of low, middle, and high frequencies, and balance between sections of the orchestra. It is affected by the acoustic environment in which the music is produced. If a hall either amplifies or absorbs the treble sound, either brittleness or a muffled quality may mar the music. If the surfaces of the walls or ceiling or seats absorb the low frequencies, the full orchestra may sound deficient in basses or cellos. If the stage enclosure or the main ceiling directs certain sounds toward some parts of the hall only and not toward others, the tone color will be affected differentially.

One facet of concert hall acoustics that has a clearly observable effect on orchestral timbre occurs when the length of the RT at low frequencies is either too long or too short compared to the RT at middle frequencies. This ratio is directly related to the musical attribute called ''warmth.'' A hall lacks warmth when the reverberation times are lower at low frequencies than at mid-frequencies, i.e., a low bass ratio BR. *But that measurement is meaningful only when the hall is fully occupied, because people absorb the high frequencies more than the low frequencies.*

IV. SUMMARY OF THE MUSICAL QUALITIES AFFECTED BY ACOUSTICS

Figure 4 summarizes the interrelations between the musical qualities heard in a concert hall and the acoustical factors that affect those qualities. This chart, together with these definitions cover the known interrelations between acoustics and the quality of music performed in a concert hall.