



Please note that the links in the PEARL logotype above are “live” and can be used to direct your web browser to our site or to open an e-mail message window addressed to ourselves.

To view our item listings on eBay, [click here](#).

To see the feedback we have left for our customers, [click here](#).

This document has been prepared as a public service . Any and all trademarks and logotypes used herein are the property of their owners.

It is our intent to provide this document in accordance with the stipulations with respect to “fair use” as delineated in Copyrights - Chapter 1: Subject Matter and Scope of Copyright; Sec. 107. Limitations on exclusive rights: Fair Use.

Public access to copy of this document is provided on the website of Cornell Law School at <http://www4.law.cornell.edu/uscode/17/107.html> and is here reproduced below:

Sec. 107. - Limitations on exclusive rights: Fair Use

Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phono records or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include:

- 1 - the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- 2 - the nature of the copyrighted work;
- 3 - the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- 4 - the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors

Loudspeaker Measurements and Their Relationship to Listener Preferences: Part 1*

FLOYD E. TOOLE

National Research Council, Ottawa, Ontario K1A 0R6, Canada

Precision measurements on loudspeakers have been possible for some time now, and over the years, various views of their importance have developed as a result of accumulated experience and scientific investigation. A survey of the literature reveals areas of agreement and disagreement among workers. There is also evidence of geographic concentrations of loudspeaker designers and reviewers who appear to share attitudes toward specific measurements. This paper discusses loudspeaker measurements in preparation for Part 2, which presents the results of some current research on the subject.

0 INTRODUCTION

Evaluating loudspeakers by means of measurements is rather like being a detective looking for clues to the existence and origins of misbehavior. The measurements that we use are not, in themselves, virtues or problems; they are merely indicators. Two ears and a brain do not process sounds as do microphones, measuring instruments, and the eyes. The auditory perceptions of musical sounds are not the same as visual analyses of data from clinical sine waves, impulses, and pink noise. And yet, for now at least, there is no choice but to proceed with the traditional methods bearing in mind *always* that evidence that offends the eye may or may not indicate the presence of a problem that is offensive to the ear.

The relationship between listener perceptions and measured quantities is a general problem in psychometrics that has many manifestations in audio. The most basic relationships deal with the detection thresholds of sounds of various kinds, with or without the embellishment of added distortions. From these one progresses through the relationships between the measured quantity of a stimulus and the strength of the associated perception. The scientific literature is replete with learned studies of this kind, dealing, as much as possible, with isolated perceptual dimensions that are logically and clearly related to individually measurable parameters.

The situation becomes more complicated when the experiment is expanded to include multiple stimuli, and vastly more complex with the inclusion of stimulus inputs through more than one sensory modality. If the controls are further relaxed, and the repertoire of experimental signals includes real-world sounds propagated in real-world environments, precise experimental control becomes almost impossible and the perceptions are no longer simple. Yet these are the conditions of a typical listening test within which listeners readily tender opinions of preference, sometimes complete with detailed subjective analyses of what they hear.

It is characteristic of commonplace listening experiences that listener opinions will vary. It is argued by some that it is the variety of listener tastes and preferences that has led to the wide diversity of sounds from loudspeakers, all of which pass, loosely, as "high fidelity." Thus has developed the traditional problem of attempting to match the tastes of the consumer with those of the loudspeaker designer; a messy business indeed.

With enough attention to detail, however, it is possible to demonstrate that much, if not most, of the variation in listener opinion stems not from irreconcilable differences among individuals but from the influence of physical and psychological factors, many of them unrelated to the matter in question. There may indeed be clear differences between listeners with widely different hearing ability, but among listeners with similar, particularly similarly good, hearing the differences of opinion can be very small indeed [1].

* Manuscript received 1985 May 30; revised 1986 January 9.

From the results of the strictly controlled listening tests that have been conducted it is possible to clearly identify loudspeakers that are consistently favored by listeners with normal hearing and those that are not. In fact, the fidelity ratings from these subjective measurements can be used as a means to explore the objective domain, seeking to identify those measured variables that are indicators of listener preference.

The ultimate objective of the present investigation could be stated as follows: to define a set of technical measurements, including the manner of their measurement and the form of their presentation, such that interpretations by experienced unbiased observers correspond to the results of controlled listening tests using experienced unbiased listeners. A loudspeaker that has a good technical performance, in certain specific terms, should sound good in the listening room—and vice versa.

1 AN HISTORICAL PERSPECTIVE

Before presenting the results of the present research work it is important to appreciate the historical context into which they fit. In most scientific and engineering endeavors precise measurements are taken for granted as being of fundamental value. In the design and evaluation of loudspeakers, however, the role of measurements has never been clearly defined.

Part of the difficulty stems from the nature of the device itself. A loudspeaker generates a three-dimensional sound field within a complicated acoustical enclosure, the listening room. Without an understanding of the sound propagation characteristics of the room and of the perceptual processes of the human recipient of the sound, it has not been possible to be definitive about the aspects of loudspeaker performance that are most important.

Inherent in the design or evaluation of loudspeakers is the establishment of some rules by which the work will proceed. By tradition, subjective evaluations have remained the final arbiters of sound quality. However, measurements have been assuming an ever-increasing role in loudspeaker work, and lacking definitive scientific data, workers have used their own experiences and beliefs to develop rules for using measurements. In the early years progress was hampered by primitive instrumentation and poor sound sources. The motivations were, nevertheless, very high indeed, and one need look no further than to a paper by Brittain, in 1936–1937 [2], for a remarkably perceptive review of the hearing process and loudspeaker measurements. A “standard living room” for listening tests is described, as are the techniques, such as spatial or frequency-domain averaging, that help in extracting useful data from measurements in normal rooms. Other measurements were made either in an “acoustically absorbing” room or on a tower in genuine free space.

Brittain sums up his opinions in the following list of “common imperfections of electroacoustic systems . . . , roughly in order of importance.” (Some of the author’s terminology has been updated for the benefit

of modern readers.)

- 1) Frequency response (amplitude response)
- 2) Harmonic distortion
- 3) Spurious noises and intermodulation distortion
- 4) Frequency shift (FM distortion?)
- 5) Dynamic range compression (power compression)
- 6) Transient distortion
- 7) Phase distortion
- 8) Group delay
- 9) Electroacoustic efficiency
- 10) Power-handling capacity
- 11) Constancy of performance.

Looking at this list, one is likely to be overcome with a feeling of *déjà vu*, and rightly so; these conclusions are 50 years old. Yet they hold up well in light of today’s knowledge.

The following is a review of more recent opinions expressed in the literature. Unfortunately the writings are fragmented, with useful contributions to be found in consumer audio publications and trade or corporate publications as often as in professional journals. In all there are few spokespersons for such a large and mature industry. Most of what has transpired in the field of loudspeaker system design has occurred behind closed doors by people who range from enthusiastic amateurs with an artist’s touch to straightforward engineers. Few of these workers have made their ideas public.

In addition, there are the product reviewers. These people, by their journalistic nature, are more candid about their approaches to technical assessment. Some reviewers eschew measurements entirely, or give them only token respect.

Naturally, influences can bear both ways, from designers to reviewers, and vice versa. There are substantial differences of opinion in both camps, and some interesting similarities. For example, there are suggestions of geographical patterns, with the designers and reviewers of specific areas forming what appear to be mutually compatible closed systems. This may or may not have anything to do with listener taste, since adequately controlled and reported listening tests are almost nonexistent, but it certainly does seem to have something to do with the locally fashionable form of amplitude-response measurement.

Almost all workers acknowledge the value of more than one form of amplitude-response measurement. For different reasons, however, everyone seems to have a favorite, and the following rough classification reflects the predominant view held by each author, at least at the time of publication.

The presentation of these views has been organized around the sequence of sounds arriving at a listener’s ears in a normal listening room. First there is the direct sound, normally the sound radiated along the reference axis of the loudspeakers. Next come the sounds that have been reflected once from the adjacent room boundaries—floor, walls, and ceiling—radiated from the loudspeaker at specific and sometimes large angles off axis. Still later there is the reverberant field developed from sounds that have been reflected several

times from the room boundaries and objects within the room. These sounds include, to varying degrees, sounds radiated from the loudspeaker in all directions.

1.1 Amplitude Response as a Function of Listening Perspective

The first item on Brittain's list is amplitude response, referring mainly to on-axis measurements in free space. This curve, as well as being the most flattering to most loudspeakers, describes the direct sound at the ears of properly seated listeners. A number of workers such as Hentsch [3] in Switzerland, and Harwood [4], [5] and Mathers [6] at the BBC and Colloms [7], also in the United Kingdom, clearly believe in the preeminence of the direct sound in establishing perceived sound quality. Their emphasis is therefore on amplitude responses measured in the free field on axis and within a listening window sufficient to include the ears of normally positioned listeners, usually up to 30–45° off axis horizontally, and less vertically.

Moving, philosophically, further into the listening room, we embrace the views of those who believe in including some measure(s) of sounds that arrive at a listener's ears after one or more reflections. Shorter [8], writing of his BBC experiences in 1958, some years before Harwood and Mathers, said that "experience . . . suggests that, if a single quantity representing 'effective' response is to be found at all, it will lie somewhere between the axial and the mean spherical response." He goes on to suggest a spherical integration weighted to give the front response more prominence.

Moir appears to be in general agreement with the suggestion of an ideal loudspeaker having uniform sound distribution over "something less than ± 90 degrees in front of the loudspeaker" [9] and of measurements that take into account both the direct and the reverberant energies by using temporal or spectral integrations [10]. He reaches no final conclusions but notes simply that at the present stage free-field measurements are "in better agreement with the subjectively judged response than any response curve measured in the listening room" [10].

Remaining in the United Kingdom, Walker [11] and Cooke [12], while acknowledging the importance of an uncolored direct sound, point out that indirect sounds should be similarly free of irregularities to avoid perceptible coloration.

Queen in the United States concluded that for sound reinforcement the direct sound was the dominant factor in spectrum perception [13], but after examining stereo imaging closely, he observed that "loudspeaker designs for home music listening rooms must consider directivity not from the standpoint of audience coverage (direct sound) but from the standpoint of uniformity of the intensity of arriving reflections with respect to frequency." From this he went on to specify loudspeaker design objectives of either a nondirectional horizontal radiation, with restricted vertical dispersion, or "a directional loudspeaker providing high uniformity of directional pattern with frequency" [14].

Heyser, also in the United States, without reference to supporting subjective data, selected a combination of free-field measurements and listening-room measurements time-windowed to include early reflections, clearly expressing a compatible philosophy [15].

Komamura et al. [16], in Japan, performed factor analyses on extensive subjective and objective data, and concluded that, of the measured parameters, frequency responses measured in the listening room and in the anechoic room, at 0, 30, and 60°, have a high correlation with listener preference. This, clearly, is a transitional point of view, since it suggests that certain free-field data and listening-room measurements are compatible.

The next significant shift of perspective is to embrace the listening-room sound field in its entirety. Some workers, noting that listeners are often in a predominantly reverberant sound field, alluded to the "substantial evidence that the total acoustic power output of a speaker system over its frequency range is the most important characteristic determining the frequency response heard by the listener" [17]. Brociner and von Recklinghausen were careful to point out, however, that such measures can be fooled, saying that ". . . a nonsmooth power response measured in the reverberation room indicates a similarly irregular frequency response as measured in an anechoic chamber. While it cannot be stated that a speaker system which shows a smooth response in the reverberation room will necessarily sound good or have a smooth pressure response, the reverse is true."

It was at the same time, 1968, that Bose developed his distinctive loudspeaker and, along with it, the view that "when the loudspeaker is properly placed with respect to the rear reflecting wall, the frequency response measured with respect to the total radiated acoustical energy should be flat" [18]. This position was reinforced in a later publication where he concluded that "loudspeakers should be designed to a flat power criterion rather than the conventional flat-frequency-response-on-axis criterion" [19]. Bose, by implication, was referring to his own designs which (then) were either multidirectional or at least widely dispersing, generating a clearly dominant reverberant sound field in the listening room.

Others applied the sound power criterion to conventional loudspeakers, among them several who, perhaps by chance, worked in the same New England (U.S.) area. McShane [20], Allison and Berkovitz [21], Consumers Union [22], Torick [23], and Hirsch [24] all attributed a dominant role to sound power in determining the spectral character of sound in a listening room. Rosenberg [25] and Staffeldt [26] in Scandinavia prevent this from being an exclusively U.S. point of view.

It therefore appears that in the United States and in the United Kingdom there were collections of designers and reviewers sharing perspectives on measurements that would seem to place them at opposite ends of a controversy.

There is probably no better way to illustrate the trans-

Atlantic polarization than to quote from Shorter [8] who says, "The only firm conclusion which can safely be drawn is that with wide-range loudspeakers of conventional directional characteristics a flat axial response may be acceptable but a flat mean-spherical (i.e., power) response is intolerable."

On the face of it, it would seem that both points of view cannot be correct. Consequently there is a tendency for workers, including some of those mentioned above, to perform both kinds of measurements, just to be comprehensive.

Even among the advocates of sound power, the measurements were not seen to be faultless. In addition to the caution stated by Brociner and von Recklinghausen, Allison and Berkovitz commented on a discrepancy in the measurements noting that "the average room does not give the low-frequency support that is commonly assumed." These were attributed to source-boundary interactions, well known in physical acoustics, and identified earlier by Sioles [27], discussed later by Allison [28].

If the purpose of a measurement is to predict loudspeaker performance in a listening room, why not simply do it there? Sioles, in 1963, "recognized that measurements on loudspeakers must be taken in their normal operating environment in order to assess the overall performance unless a large amount of free-field data is accumulated and averaged" [27].

Indeed, the simplicity of the method is appealing, since no special measuring environments are necessary and, using one-third-octave bands of noise, even the instrumentation is simple. Benade, however, cautioned that such filter widths may fail to reveal important irregularities [29].

Nevertheless, one-third-octave measurements are widely used for assessments by workers who have found it to be a good indicator of listener preference. Observing that the optimum listening-room curve was not flat, various ideal functions have been proposed, for example, Møller [30]. Russell [31] embellished the method with some spatial averaging and suggested a range of "ideal" functions. Holl [32] also employed spatial averaging over the listening area, but generalized nonflat target curve, he derived a room calibration based on reverberation-time data and rected the actual measured curve so that the ideal response was at least approximately flat. It is equivalent to measuring sound power in a sound

perfectly tests leading him to conclude
 ment is the best indication of what a system will actually sound like." Long [33] reported using a similar method (developed for product reviews by E. Foster, with the cooperation of T. Holl and A. DeKoster).

Schulein [34] tested the technique and concluded that the nonflat room curves in domestic systems were a part of a self-consistent system, including similarly nonflat control-room monitoring improvements in the program material might increase the desirability

The fact that the "ideal" room curve is not flat is a problem for which there appears to be no unambiguous solution. Consequently one sees various approaches to rationalizing the problem; either select a "target curve" and attempt to imitate it, or derive and subtract it from the measurement. Either way important assumptions are involved.

The generalized "target curve" approach used by Møller and Russell assumes that all rooms are similar in their sound diffusion and frequency-dependent absorption. Adjacent-boundary interactions are not included and neither are any considerations of loudspeaker directivity and its effect on the direct-to-reverberant ratio of the sound field. One might expect some uncertainty about the shape of the "ideal" response curve and, in Fig. 1, this is apparent.

The approach used by Holl and Long has rather better prospects since by using standard microphone and loudspeaker placement, the boundary interactions can be maintained as constant factors, or with some calibration data, removed entirely. The room calibrations at middle and high frequencies rely on reverberation-time measurements and thus assume a diffusion in the sound field that may or may not of different directivities would therefore yield data requiring individual interpretation; the method may be useful within the limited scope of fundamentally similar products, but there are likely to be errors in the evaluation of loudspeakers in general. Nevertheless, the "frequency response" of the room used in Holl's example [32, Fig. 2], shown in Fig. 1, is close to Møller's optimum curve; apparently both workers used acoustically similar rooms. The alternative target curves suggested by Russell have very generous tolerances, although the -3-dB-per-decade slope is in the same spirit as the Møller and Holl curves.

1.2 Amplitude Response: A Synopsis

Without recourse to directly comparable measurements to illustrate the point of serious disagreement among knowledgeable and ex-

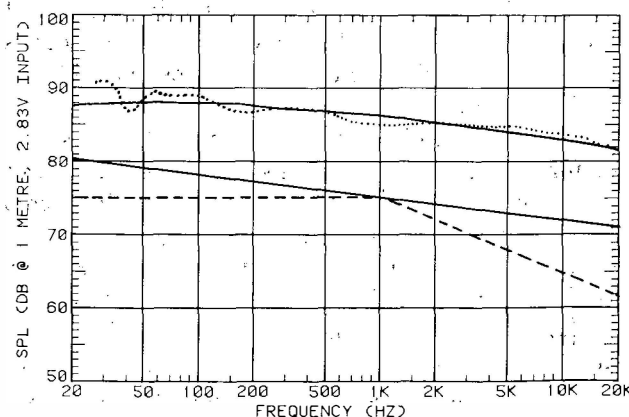


Fig. 1. "Ideal" listening-room amplitude-response curves proposed by Møller (top solid) and Russell (bottom solid and dashed). For comparison a room correction curve used by Holl is shown (dotted).

perienced workers about what constitutes the most useful measure of loudspeaker amplitude response. The magnitude of the disagreement is such that, with the popular forward-facing configuration of loudspeakers, meeting one of the amplitude-response criteria, such as a flat on-axis response, would automatically result in failure of another, the flat sound-power response, because of the directional characteristics of the device. Measurements within the listening room are sensitive to a variety of factors related to both the loudspeaker and the room. If performance in the room is the ultimate criterion, it is unlikely that measurements on the loudspeaker alone would be very reliable, since they convey only part of the information.

To the extent that the general form of the amplitude response is a determinant of listener preference, it would seem that the industry has not yet decided on a design objective.

1.3 Phase Response

In engineering terms it is so logical that phase response should matter that, over the years, several workers have repeated the investigations, only to conclude that in the real world it is, at worst, only a minor problem. With selected test signals and listening with headphones or in an anechoic room, listeners can hear differences attributable to phase shift. Using music, and listening to loudspeakers in normal rooms, even carefully selected listeners appear to have difficulty detecting the presence of quite large phase shifts, much less are they able to establish a preference. While these observations seem to be fairly general for smooth changes in phase as a function of frequency, including substantial group delays, the rapid local phase changes indicative of resonances may be another matter. Always, and especially in loudspeakers with multiple drivers and overlapping crossover regions, there is the uncertainty of whether one is hearing the effects of phase, or of an associated amplitude aberration.

This very point was made by Brittain in 1937, who noted the insensitivity of the ear to phase effects "unless accompanied by some other phenomenon" [2]. Hentsch, in 1951, observed that "phase variations" were unimportant, and large group delays barely detectable, even by experienced listeners [3].

Such results from those early years could be disputed nowadays on the basis of the rather poor microphones and other equipment involved in the tests. They are therefore all the more impressive in view of the fact that much more recent investigations by Preis [35], Bauer [36], Harwood [37], Moir [38], Lipshitz et al. [39], Saponas et al. [40], and Clark [41] among others, have come to basically similar conclusions.

The recent work has been more thoroughly investigative and puts some limits on the thresholds of audibility of various phase and group-delay effects. The limits on overall trends in phase response are very generous, though, and appear not to require special consideration in the design of conventional domestic loud-

speakers. However, Saponas et al. are careful to point out that localized sharp discontinuities in amplitude response tend to be associated with high rates of change of phase shift and the combination causes "frequency distortion."

Others, though, have disagreed with these views, among them Bowers and Roe [42] and Gerzon [43]. In [42], the authors compare two loudspeakers, one adjusted to yield a linear phase response and a similar nonlinear phase version. There was "some departure in measured amplitude response in the nonlinear phase system, but as identical drive was applied to identical units this was not counted a relevant variable." In other words, the electrical input to the two loudspeaker systems was the same, but the acoustical output differed in both amplitude and phase response. One is left to wonder which parameter caused the "clearly audible" differences reported by the listeners. In these experiments it is difficult to manipulate one parameter in isolation, particularly when altering electroacoustic devices, a point well made by Barlow [44] who investigated the same effect and concluded that "any audible effect due to the acoustic centres being in the same plane is very small and is masked by reflection" (from the front panel setbacks).

1.4 Audibility of Amplitude-Response and Time-Domain Anomalies

When Bücklein, in 1962, investigated the audibility of various forms and magnitudes of peaks and dips as a function of frequency, he addressed a fundamental problem in amplitude-response assessment [45]. How flat or smooth does a curve really need to be?

He observed, among other things, that an amplitude-response irregularity appearing as a peak may be highly objectionable when the equivalent dip may be barely audible.

Bücklein also found that the audibility of irregularities increased with their width and their amplitude. The latter is to be expected but the former is interesting, as it implies that low- Q resonances are more easily detected than high- Q resonances. This finding was later confirmed by Stevens [46], looking at enclosure resonances, and by Fryer [47] who explored the audibility of resonances as a function of both Q and frequency. Barlow, discussing Fryer's work, summarized the important points as follows: "The ear evidently detects mainly the energy or area under the peak. In some cases, at high Q , the peak can be well above the general level before being detectable, but in other cases a low- Q resonance, well below the general level, is still detectable. Damping a resonance may not give as great an improvement as hoped for, especially at low frequencies" [44].

In both Stevens's and Fryer's tests, listeners detected low- Q ($Q \leq 10$) resonances at levels 10 to 20 dB below average program level with as much ease as they found high- Q ($Q \geq 20$) resonances with output equal to or above the average level. In terms of measured amplitude response this means that some audible colorations may

be present even though the curve may appear to be quite flat. On the other hand, some narrow spikes or sharp dips that are offensive to the eyes may be virtually ignored by the ears.

Translated into the time domain, there should be a similar sensory contradiction wherein the high- Q resonances with their prolonged decays that are so offensive to the eyes when viewed on an oscilloscope or in "waterfall" diagrams of signal energy in time and frequency [48], [49] may be less audible than rapidly damped lower amplitude overhangs. Hentsch, in fact, concluded just that when he found that viewing transient phenomena on an oscilloscope "is not a certain means of determining the subjective effects of the deformations" [3].

Shorter [50], in 1946, and Corrington [51], in 1955, developed gating systems for looking at the energy in the decaying tails of tone bursts. Both workers were convinced that this view of events in the time domain offered better indications of audible colorations than conventional amplitude-response measurements that included the visual clutter of nonresonant interference effects. Hentsch, in 1951, saw a way around this by using spatial averaging of anechoic swept-tone amplitude responses to more clearly reveal the spectral irregularities caused by resonances. He concluded that "the study of amplitude response alone is sufficient to determine the quality of reproduction in the transient regime" [3]. In 1961 Larson and Adducci [52] arrived at the same conclusion and observed further that "little correlation exists between the transient performance of a loudspeaker and musical listening tests."

In 1971 Sapoñas, Matson, and Ashley (40) examined a variety of transient and steady-state test signals and concluded that most of the commonly used transient signals were inappropriate for evaluating loudspeakers, and for even a simple impulse, it would be "nearly impossible to look at the output from a microphone and say if this is a good or bad result." They concluded that "the sine wave is still very much the king of test signals."

Driscoll, writing in 1974, basically agreed with this result saying that while he thought that impulse tests can be "a useful pointer to at least some qualities of a loudspeaker's performance," the test by itself cannot be used to rank products "in any order of performance quality" [53]. Unwin expressed almost the directly opposite point of view, placing great importance in the visual form of the impulse response and arguing that "testing a loudspeaker with a steady sine wave tells us almost nothing about the way it will handle transients which are the most important constituents of musical performance" [54]. Unfortunately no proofs were offered.

In distinct contrast to all of these endeavors to identify and quantify the sources of resonance and transient colorations there is the conclusion of Bose who "learned that it is possible to produce music without audible coloration from distortion, resonances or transient response irregularities by the use of a multiplicity of full

range loudspeakers" [19]. These tests were done with loudspeakers that generated a predominantly reverberant sound field so that, in addition to the small differences among the various drivers, there was the confusion of the mostly reflected sound field to disguise the audible effects of these technical imperfections. Briggs, writing earlier, in 1958, noted that directing the axis of a loudspeaker toward a wall or ceiling "camouflages peaks in response to a remarkable extent because the room gets to work on the sound waves before they reach the ear of the listener" [55]. These findings might be related to that of Harwood [5] who obtained the "remarkable result" that listeners could not hear a series of low- Q resonances at spacings of less than one octave over the entire frequency range.

Although there are some clear differences of opinion, there are nevertheless some recognizable patterns in this collection of findings and they can be described as follows: audible colorations are often caused by resonances in a loudspeaker system. These resonances will be revealed, in varying degrees, in both steady-state amplitude responses and time-domain responses to interrupted sounds.

In order for amplitude responses to reveal unambiguous clues to the presence of resonances the measurement must have high resolution in the frequency domain and incorporate some form of spatial averaging to remove the visual clutter caused by acoustical interference. This done, evidence of potentially audible coloration will take the form of upward thrusting peaks deviating from the underlying general shape of the curve. Wide, low-amplitude bumps can be as annoying as much higher, narrow peaks. Dips in the amplitude response appear to be much less important than peaks. The technique is limited, however, in that it will not necessarily reveal the low- Q resonances 10 to 20 dB below the level of the primary signal that have been shown to be sources of audible coloration.

Explorations of time-domain performance can identify most resonance problems, but the interpretation appears to run contrary to common preconceptions. The highly visible long "tails" caused by high- Q resonances are not always the most severe problems. The broader band, better damped tails of low- Q resonances need to be ferretted out of the time-domain data, as these are often associated with the dominant audible coloration.

It is interesting to speculate why this phenomenon occurs. First, from a purely statistical point of view, a broad resonance will be excited more often by sounds in music than a narrow resonance. Second, since musical sounds at middle and high frequencies are either transient in nature or are at least strongly amplitude modulated, a low- Q resonance will reach full amplitude more often and more quickly than a high- Q resonance. And, finally, since these are forced resonances, once the exciting signal is removed, the resonant system reverts to its natural frequency. For a high- Q resonance the resonant overhang will have the same, or close to the same, frequency as the signal that stimulated it,

and its presence may be difficult to detect, given that musical instruments are themselves resonant devices. A low- Q resonance, on the other hand, can ring at a frequency significantly different from the one that initiated the response, thus imparting a monotonal coloration to a range of exciting frequencies.

The frequency at which the amplitude aberration occurs also matters. All studies have indicated a reduction in audibility at both frequency extremes. Bücklein's results show two regions of high sensitivity: between about 200 Hz and 600 Hz and 2 kHz and 6 kHz. Fryer's data indicate a slightly increased sensitivity between 2 kHz and 5 kHz.

Between about 600 Hz and 2 kHz the ears seem to be marginally less sensitive to these problems. It may be more than coincidence that this is the frequency range within which there appears to be a kind of physiological crossover of auditory function. For example, below about 1 kHz the inner ear is capable of neurally encoding accurate information about the acoustic waveform; above this frequency there is an increasing reliance on the amplitude envelope, rather than carrier information. In sound localization this is apparent in the importance of interaural time differences at lower frequencies while, at high frequencies, the sensitivity to timing information seems to apply principally to the amplitude-modulated envelope rather than to the fine structure of the underlying signal [56]. It is fortunate that this frequency range, within which there may be some uncertainty about the dominant perceptual cues, includes the majority of woofer- or midrange-to-tweeter crossovers, where amplitude and phase response problems are routinely found.

1.5 Nonlinear Distortions

Shorter, in 1958, offered the advice that "with some forms of distortion it would probably be easier, as well as more profitable, to remove the cause by appropriate design than to discover rules for assessing the effect" [8]. In electronic components this principle has been applied with considerable success, but in loudspeakers, it is commonplace to measure distortions in whole percentages at modest sound levels. Either listeners are remarkably tolerant of these imperfections or there is still a major problem to be solved.

From accounts of different investigations, however, one can reach a number of conclusions. Bose concluded that with most high-fidelity loudspeakers "audible nonlinear distortion in music or speech is definitely one of the more minor of their shortcomings, notwithstanding the fact that distortion measurements . . . can be quite large" [18]. Kantrowitz arrived at a fairly generous 3% as the level of CCIF (intermodulation) distortion above which it became objectionable in high-frequency drivers [57]. Here, though, is a case where one might suspect that the source material may have had an influence on the results since the 1962-vintage disk recordings, and playback devices had similar problems of their own. The more recent work of Fryer [47] indicated similarly large values for the detection

of intermodulation distortion. Noting that typical loudspeakers at typical listening levels produce less than 1% intermodulation distortion, he concluded that "it is not a particularly serious issue for designers."

However, there are several forms of nonlinear distortion and a number of important variables influencing their audibility. Cabot, in a recent review of previous work, discusses the subject in some detail [58]. While acknowledging some of the high detection thresholds that have been reported, he comments on the need for more experiments with better controls and using the superior program material available today. Selecting what he regards as possibly "the most reliable work to date," Cabot arrives at a performance objective of 0.05% total harmonic distortion, with the provision that high-order components not be dominant. If this is the case, then there is considerable room for improvement in loudspeakers, and Shorter's advice might prove to be difficult and rather expensive to implement.

Even with this small sampling of the literature on nonlinear distortion it is evident that there is no unanimity of opinion. Clearly, there is a need for more work on this subject.

2 DISCUSSION

Given the age and size of the industry it may appear to be somewhat surprising that there is so little agreement on the desirable measured performance objectives of loudspeakers. That this is true for something as basic as the amplitude response is particularly disconcerting.

Errors in the physical measurements are not likely to be at fault here, the differences at issue are simply too large. There are some suggestions that the manner of measurement and data presentation may be important, but, there would appear to be secondary factors. There is even the suggestion that geography might be a consideration.

Throughout the published accounts of loudspeaker design and evaluation methods there are references to listening tests and subjective preferences lending support to one or another technical evaluation scheme. Rarely are the subjective measurement methods described in any detail. Knowing now the attention to detail that is necessary to obtain reliable subjective data, it is reasonable to speculate that another important variable underlying these uncertainties is listener opinion.

Recent work on subjective measurements [1] has produced a large body of reliable subjective data on a number of loudspeakers. In Part 2 of this paper these data provide the basis for an examination of certain aspects of measured loudspeaker performance.

3 REFERENCES

- [1] F. E. Toole, "Subjective Measurements of Loudspeaker Sound Quality and Listener Performance," *J. Audio Eng. Soc.*, vol. 33, pp. 2-32 (1985 Jan./Feb.)
- [2] F. H. Brittain, "The Appraisal of Loudspeakers," *GEC J.*, pt. 1, vol. 7, pp. 266-276 (1936)

Nov.); pt. 2, vol. 8, pp. 121–130 (1937 May).

[3] J.-C. Hentsch, "La Fidélité des Haut-Parleurs dans la Reproduction des Phénomènes Transitoires," *Tech. Mitt.-PTT* (Bern), no. 6 (1951 June).

[4] H. D. Harwood, "The Hows, Wheres and Whys of Testing High Quality Loudspeakers," *Audio*, vol. 55, pt. 1, pp. 16–20 (1971 Aug.); pt. 2, pp. 18–21 (1971 Sept.).

[5] H. D. Harwood, "Some Factors in Loudspeaker Quality," *Wireless World*, vol. 82, pp. 45–54 (1976 May).

[6] C. D. Mathers, "Design of the High-Level Studio Monitoring Loudspeaker Type LS5/8," British Broadcasting Corp., Rep. BBC-RD-1979/22 (1979 Nov.).

[7] M. Colloms, *High Performance Loudspeakers*, 3rd ed. (J. Wiley, New York, 1985).

[8] D. E. L. Shorter, "A Survey of Performance Criteria and Design Considerations for High-Quality Monitoring Loudspeakers," *Proc. IEE*, vol. 105, pt. B, pp. 607–621 (1958 Nov.); reprinted in *J. Audio Eng. Soc.*, vol. 7, pp. 13–28 (1959 Jan.).

[9] J. Moir, "Speaker Directivity and Sound Quality," *Wireless World*, vol. 85, pp. 61–63, 96 (1979 Oct.).

[10] J. Moir, "Loudspeakers in Rooms, Which Response," *Hi-Fi News Rec. Rev.*, vol. 27, pp. 22–25 (1982 June).

[11] P. J. Walker, "The Loudspeaker in the Home," *J. Brit. Inst. Radio Eng.*, vol. 13, pp. 377–380 (1953 July).

[12] R. E. Cooke, "Misleading Measurements," *Hi-Fi News Rec. Rev.*, vol. 21, (1976 Oct.); reprinted in *KEFTOPICS*, vol. 1, no. 3, KEF Electronics Ltd., Maidstone, England.

[13] D. Queen, "Relative Importance of the Direct and Reverberant Fields to Spectrum Perception," *J. Audio Eng. Soc. (Project Notes/Engineering Briefs)*, vol. 21, pp. 119–121 (1973 March).

[14] D. Queen, "The Effect of Loudspeaker Radiation Patterns on Stereo Imaging and Clarity," *J. Audio Eng. Soc.*, vol. 27, pp. 368–379 (1979 May).

[15] R. C. Heyser, "Breakthrough in Speaker Testing," *Audio*, vol. 57, pp. 20–30 (1973 Nov.).

[16] M. Komamura, K. Tsuruta, and M. Yoshida, "Correlation between Subjective and Objective Data for Loudspeakers," *J. Acoust. Soc. Japan*, vol. 33, pp. 103–115 (1977 Mar.).

[17] V. Brociner and D. R. von Recklinghausen, "Speaker System Design Using a Reverberation Chamber," presented at the 34th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 16, p. 342 (1968 July), preprint 579.

[18] A. G. Bose, "On the Design, Measurement and Evaluation of Loudspeakers," presented at the 35th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 17, p. 86 (1969 Jan.), preprint 622.

[19] A. G. Bose, "Sound Recording and Reproduction, Part 1: Devices, Measurements, and Perception," *Technol. Rev. (MIT)*, pp. 19–25 (1973 June); "Part 2: Spatial and Temporal Dimensions," *ibid.*, pp. 25–33 (1973 July/Aug.).

[20] C. L. McShane, "The Meaning of Quantitative Loudspeaker Measurements," presented at the 36th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 17, p. 344 (1969 June),

preprint 651.

[21] R. F. Allison and R. Berkovitz, "The Sound Field in Home Listening Rooms," *J. Audio Eng. Soc.*, vol. 20, pp. 459–469 (1972 July/Aug.).

[22] "Loudspeaker Accuracy: CU's Tests," *Consumer Rep.*, vol. 38, pp. 456–457 (1973 July).

[23] E. Torick, "In the Loudspeaker Testing Lab," *High Fidelity*, vol. 27, pp. 69–73 (1977 Oct.).

[24] J. D. Hirsch, "Testing Speakers," *Stereo Rev.*, vol. 47, pp. 24–25 (1982 Aug.).

[25] U. Rosenberg, "Loudspeaker Measurement and Consumer Information," Statens Provningsanstalt, 11486 Stockholm 5, Sweden, Rep. 244 (1973).

[26] H. Staffeldt, "Correlation between Subjective and Objective Data for Quality Loudspeakers," *J. Audio Eng. Soc.*, vol. 22, pp. 402–415 (1974 July/Aug.).

[27] G. W. Sioles, "Loudspeaker Measurements in Live Rooms," *J. Audio Eng. Soc.*, vol. 11, pp. 203–206 (1963 July).

[28] R. F. Allison, "The Influence of Room Boundaries on Loudspeaker Power Output," *J. Audio Eng. Soc.*, vol. 22, pp. 314–320 (1974 June).

[29] A. H. Benade, *Fundamentals of Musical Acoustics* (Oxford University Press, New York, 1976), chap. 12.

[30] H. Møller, "Relevant Loudspeaker Tests," Brüel and Kjaer Application Note 15–067; also "Relevant HiFi Tests at Home," presented at the 47th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 22, p. 272 (1974 May).

[31] R. H. Russell, "Speaker Evaluation: Ear or Machine?," *The Audio Amateur*, pt. 1, pp. 10–15, issue 1; pt. 2, pp. 20–25, issue 2 (1975).

[32] T. Holl, in conversation with J. Atkinson, "Speakers, Rooms and Response," *Hi-Fi News and Rec. Rev.*, vol. 28, pp. 40–43 (1983 May).

[33] R. Long, "Loudspeaker Testing and the Listening World," *High Fidelity*, vol. 31, pp. 18–19 (1981 June).

[34] R. B. Schulein, "In Situ Measurement and Equalization of Sound Reproduction Systems," *J. Audio Eng. Soc.*, vol. 23, pp. 178–186 (1975 Apr.).

[35] D. Preis, "Linear Distortion, Measurement Methods and Audible Effects: A Survey of Existing Knowledge," presented at the 2nd International Conference of the Audio Engineering Society, Anaheim, California, 1984 May 11–14, preprint C1005.

[36] B. B. Bauer, "Audibility of Phase Distortion," *Wireless World*, vol. 80, pp. 27–28 (1974 Mar.).

[37] H. D. Harwood, "Audibility of Phase Effects in Loudspeakers," *Wireless World*, vol. 82, pp. 30–33 (1976 Jan.).

[38] J. Moir, "Phase and Sound Quality," *Wireless World*, vol. 82, pp. 80–84 (1976 Mar.).

[39] S. P. Lipshitz, M. Pocock, and J. Vanderkooy, "On the Audibility of Midrange Phase Distortion in Audio Systems," *J. Audio Eng. Soc.*, vol. 30, pp. 580–595 (1982 Sept.).

[40] T. A. Saponas, R. C. Matson and J. R. Ashley, "Plain and Fancy Test Signals for Music Reproduction Systems," *J. Audio Eng. Soc.*, vol. 19, pp. 294–305 (1971 Apr.).

[41] D. Clark, "Measuring Audible Effects of Time Delays in Listening Rooms," presented at the 74th Convention of the Audio Engineering Society, *J. Audio*

Eng. Soc. (Abstracts), vol. 31, p. 972 (1983 Dec.), preprint 2012.

[42] J. Bowers and S. Roe, "Phase and Loudspeakers," *Hi-Fi News Rec. Rev.*, vol. 21, pp. 56–61 (1976 Apr.).

[43] M. A. Gerzon, *Wireless World (Letter to the Editor)*, and reply by Harwood, vol. 82, pp. 60–61 (1976 Mar.).

[44] D. A. Barlow, "Loudspeaker Coloration," *Wireless World*, vol. 84, pp. 34–36 (1978 Mar.).

[45] R. Bücklein, "The Audibility of Frequency Response Irregularities," *J. Audio Eng. Soc.*, vol. 29, pp. 126–131 (1981 Mar.), a translation of a 1962 original paper in German.

[46] W. R. Stevens, "Loudspeakers—Cabinet Effects," *Hi-Fi News Rec. Rev.*, vol. 21, pp. 87–93, 97 (1976 Sept.).

[47] P. Fryer, "Loudspeaker Distortions—Can We Hear Them?," *Hi-Fi News Rec. Rev.*, vol. 22, pp. 51–56 (1977 July).

[48] J. M. Berman and L. R. Fincham, "The Application of Digital Techniques to the Measurement of Loudspeakers," *J. Audio Eng. Soc.*, vol. 25, pp. 370–384 (1977 June).

[49] C. P. Janse and A. J. M. Kaizer, "Time-Frequency Distributions of Loudspeakers: The Application

of the Wigner Distribution," *J. Audio Eng. Soc.*, vol. 31, pp. 198–223 (1983 Apr.).

[50] D. E. L. Shorter, "Loudspeaker Transient Response—Its Measurement and Graphical Representation," *BBC Quart.*, vol. 1, pp. 1–9 (1946 Oct.).

[51] M. S. Corrington, "Correlation of Transient Measurements on Loudspeakers with Listening Tests," *J. Audio Eng. Soc.*, vol. 3, pp. 35–39 (1955 Jan.).

[52] R. J. Larson and A. J. Adducci, "Transient Distortion in Loudspeakers," *IRE Trans. Audio*, pp. 79–85 (1961).

[53] R. Driscoll, "New Transient Test for Loudspeakers," *The Gramophone*, pp. 789–790 (1974 Oct.).

[54] P. Unwin, "Loudspeaker Coloration," *Hi-Fi News Rec. Rev.*, vol. 19, pp. 71–75 (1974 Mar.).

[55] G. A. Briggs, *Loudspeakers*, 5th Ed. (Wharfedale Wireless Works, Idle, Bradford, Yorkshire, 1958).

[56] J. Blauert, *Spatial Hearing* (1974); transl. by J. S. Allen (M.I.T. Press, Cambridge, MA., 1983).

[57] P. Kantrowitz, "Distortion of High-Frequency Loudspeakers," *J. Audio Eng. Soc.*, vol. 10, pp. 310–317 (1962 Oct.).

[58] R. Cabot, "Perception of Nonlinear Distortion," presented at the 2nd International Conference of the Audio Engineering Society, Anaheim, California, 1984 May 11–14, preprint C1004.

THE AUTHOR



Floyd E. Toole was born in Moncton, New Brunswick, in 1938. He received the B.Sc. degree in electrical engineering in 1960 from the University of New Brunswick and the Ph.D. and D.I.C. in electrical engineering in 1965 from the Imperial College of Science and Technology, University of London, U.K. Since then he has been with the acoustics section, division of physics, National Research Council, Ottawa.

Dr. Toole's early research was concerned with sound localization and the mechanisms of binaural hearing. After an interval of activity in the measurement and control of noise, including organizational and standards-writing work with the Canadian Standards Association, he returned to audio. In recent years he has been involved with loudspeakers, rooms, and listening tests. A routine program of measurements and listening tests

is regularly used by loudspeaker manufacturers, acoustical consultants, and audio publications for purposes that range from product design to product reviewing. A parallel, research-oriented effort is aimed at improving the precision and utility of measurements and listening tests. Part of this energy has been put into working groups of the International Electrotechnical Commission where he is active in standards writing for loudspeaker measurements, listening tests, headphones, and amplifiers. In the field of professional audio, Dr. Toole has designed recording studios, control-room monitor loudspeakers and sound reinforcement systems for large concert halls and theaters.

Dr. Toole is a member of the Audio Engineering Society, the Acoustical Society of America, and the Canadian Acoustical Association.

Loudspeaker Measurements and Their Relationship to Listener Preferences: Part 2*

FLOYD E. TOOLE

National Research Council, Ottawa, Ont. K1A 0R6, Canada

Using the highly reliable subjective ratings from an earlier study, loudspeaker measurements have been examined for systematic relationships to listener preferences. The result has been a logical and orderly organization of measurements that can be used to anticipate listener opinion. With the restriction to listeners with near-normal hearing and loudspeakers of the conventional forward-facing configuration, the data offer convincing proof that a reliable ranking of loudspeaker sound quality can be achieved with specific combinations of high-resolution free-field amplitude-response data. Using such data obtained at several orientations it is possible to estimate loudspeaker performance in the listening room. Listening-room and sound-power measurements alone appear to be susceptible to error in that while truly poor loudspeakers can generally be identified, excellence may not be recognized. High-quality stereo reproduction is compatible with those loudspeakers yielding high sound quality; however, there appears to be an inherent trade-off between the illusions of specific image localization and the sense of spatial involvement.

4 INTRODUCTION—PART 2

The review of published opinion in Part 1 of this paper¹ indicates that there are aspects of loudspeaker measurements about which there is significant disagreement among various workers. Prominent among these is the measurement of amplitude response, where there is no universally accepted method of measurement. Most designers and reviewers acknowledge that there is some merit in measurements of free-field on-axis and off-axis performance, and in the consolidated measures of sound-power or listening-room response. Nevertheless, individuals generally express a preference for one of these, based on their experience or investigations.

For loudspeakers of conventional design, using forward-facing radiators, the same criterion of excellence cannot be applied to all of the measurement options. A conventional loudspeaker system with a flat amplitude response on its reference axis cannot have a flat sound-

power response, and vice versa. Neither a flat axial-response nor a flat sound-power response ensures a measured response in a typical listening room that is flat, or even of any particular shape that could be standardized.

Thus without even considering the performance of loudspeakers with regard to phase response, time-domain response, or nonlinear distortions there are disagreements in basic measurement methodology. The present study is an attempt to resolve some of the apparent contradictions.

Restating the long-term objective of this project, it is to define a set of technical measurements and the form of their presentation, such that interpretations by experienced unbiased observers correspond to the results of controlled listening tests using experienced unbiased listeners. A loudspeaker that has a good technical performance, in certain specific terms, should sound good in the listening room, and vice versa.

5 MEASUREMENTS

Fundamental to this investigation is the clear identification of the subjective and objective variables to be used in the attempted correlations. The subjective

* Manuscript received 1985 May 30; revised 1986 January 13.

¹ F. E. Toole, *J. Audio Eng. Soc.*, vol. 34, pp. 227–235 (1986 Apr.).

data came from experiments conducted under strict experimental controls, all of which have been described at length in a recent paper by this author [1]. The results of these experiments were consistent and repeatable but, as was pointed out in this earlier work, it is possible to be consistently in error (due, for example, to a prejudicial selection of program material or to an acoustical aberration in the listening room). Therefore there is a second purpose to the present investigation: to test for objectively identifiable bias in the listener ratings.

5.1 Subjective Measurements

The listeners who participated in the subjective measurements from which the present data are taken ranged from professional sound-recording engineers to audiophiles. Many were musicians, but all of them had a background of serious critical listening. In all, there were 42 listeners in these tests, but the data used here pertain only to the 28 who exhibited low judgment variability. These listeners all had hearing threshold levels within 10 dB of the ISO audiometric zero at frequencies below 1 kHz, and within 20 dB, up to 6 kHz. As noted in the previous paper, these were the listeners whose "fidelity ratings" showed the greatest consistency within individuals and the closest agreement across the group of individuals.

Not all listeners auditioned all loudspeakers and not all loudspeakers were included in each experiment. There were, in fact, six separate experiments, with certain "anchor" products being common to several of them. All of the loudspeakers were conventionally enclosed two- or three-way systems with forward-facing drivers.

In the course of each 30-min exposure to randomized presentations of a test group of four loudspeakers, listeners completed questionnaires, quantifying their perceptions of various audible attributes and arriving finally at an overall fidelity rating. This rating is on a scale of 0 to 10, where 10 identifies a reproduction that is perfectly faithful to the ideal, no improvement being possible. The number 0, in contrast, denotes a reproduction with no similarity to the ideal—a worse reproduction cannot be imagined. This scale has been used extensively by the author and others, and is embodied in IEC Publication 268-13 [59]. Each experiment ran from one to several days, until each listener had auditioned each loudspeaker three to five times in the randomized presentations.

All of the fidelity ratings used here came from monophonic listening tests. Over half of the loudspeakers were evaluated in stereophonic comparisons as well, but these data are not included. As pointed out previously [1], highly rated loudspeakers receive closely similar ratings in both stereo and mono tests, but loudspeakers with lower ratings tend to receive elevated ratings in stereo assessments. This scaling distortion, combined with the increased judgment variability in stereo tests, encouraged the use of the monophonic test results.

5.2 Objective Measurements

As discussed in Part 1, the identification of a single defensibly superior method of measuring the frequency response of loudspeakers has been a matter of debate for years. Anechoic on-axis measurements have had a logical appeal on the basis that they describe the "direct" sound at a listener's ears. From a marketing point of view there is the further advantage that they tend to show the product in the most flattering light. Detractors have argued that listening rooms are not anechoic spaces. In fact some have argued that in normal rooms listeners experience a sound field that is predominantly reverberant. Consequently total sound-power output is promoted by these experts as the single most reliable indicator of sound quality. Between these extreme views are those that embrace both, either explicitly—saying that a good loudspeaker should measure well in both respects—or implicitly—advocating measurements within the listening room.

The discussion of results will deal with these alternatives and their merits in more detail. For the moment let us consider a measurement system that embraces most if not all of these alternatives and that has sufficient flexibility to process and present the data in new forms, if required.

Over nearly two decades of intermittent investigation of loudspeaker performance, the author developed a preference for high-resolution anechoic swept-tone frequency responses. These were not restricted to on-axis views of performance, but included measurements at various angles off axis [60]–[62]. The analog equipment yielded graphic plots with ease and accuracy, but one was forced to use visual analysis and integration to deduce much of value from the collection of curves.

It is a modern variation of the time-honored method that provided the basis for the following evaluations. A computer-controlled programmable oscillator was stepped through a predetermined set of frequencies on each scan. After each step the output was held until the sound field stabilized before the measurement was made. In the anechoic chamber the delay was near zero, but in the listening room a delay of about 400 ms was necessary to accommodate the reverberation characteristics of the room.

With the measurements stored in digital form, it was a simple matter to "filter" the data after the fact. Computing the energy average over several adjacent measurements is equivalent to measuring that bandwidth with a perfect filter. Similarly, because the same set of test frequencies was used in each measurement scan, it was possible to combine the results for several loudspeaker and/or microphone positions and to compute averaged measurements, to do spatial integrations over various solid angles, to compute sound-power output, directivity index, and so on, all with high resolution in the frequency domain.

In the present evaluation the pure tone was stepped through 200 frequencies uniformly spaced on a logarithmic scale running from 20 Hz to 20 kHz. Free-

field measurements were made at 2 m in an anechoic chamber.² With the loudspeaker on a turntable, measurements were made on the reference axis (if specified by the manufacturer) and at 15° increments to ±90° horizontally and vertically. Measurements continued at 30° increments in the rear hemisphere, for a total of 34 curves.

Measurements in the anechoic chamber developed progressively larger errors at frequencies below 200 Hz due to imperfect sound absorption at the room boundaries. To allow useful data to be accumulated at lower frequencies, the error was stabilized by fixing the locations of the woofer and the microphone within the standing-wave pattern in the chamber. The error was measured by comparing results obtained inside the chamber with those obtained using a duplicate configuration on a 10-m tower outdoors. Incidentally, a significant problem was created by thermal sensitivity of various parts of the loudspeaker systems. Between the interior of the anechoic chamber and the tower, differences in air temperature and the more localized effects of solar heating and wind cooling could cause frequency-response fluctuations of up to 2 dB.

Once the error function was identified, the amplitude-response correction became a routine part of the data-processing operation. Apart from occasional narrow-band anomalies caused by unusually large enclosures or remotely located reflex ports, it is believed that the measurements are accurate within ±1 dB down to 30 Hz.

Other measurements included the phase response, which was calculated by Fourier transformation from the time-domain impulse response of the loudspeaker, total harmonic distortion at two input power levels, and sensitivity. These measurements were made only on the reference axis.

5.3 Evolutionary Frequency-Response Measurements

In this investigation, all of the most common forms of amplitude-response measurements were used, and some new variations were tried. Accordingly the following responses were measured and processed for the test loudspeakers (all measurements at 2 m):

- 1) On-axis free-field measurements
- 2) Off-axis free-field measurements up to ±180° horizontal and vertical at 15° increments in the front hemisphere and 30° increments in the rear hemisphere
- 3) Mean amplitude response in front hemisphere
- 4) Mean amplitude response in total sphere
- 5) Mean amplitude response in ±15° "listening window"
- 6) Mean amplitude response in ±30–45° annulus
- 7) Mean amplitude response in ±60–75° annulus
- 8) Total sound power

² This places the microphone in, or close to, the far field of most conventional loudspeakers. The sensitivity at the standard distance of 1 m was calculated from measurements at 2 m.

9) Directivity index

10) Phase response.

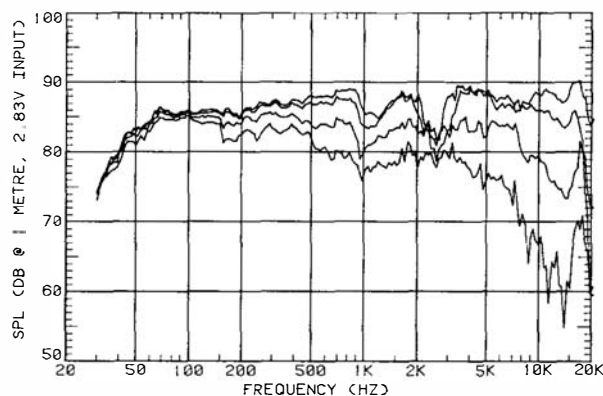
The combinations of measurements appearing throughout this paper as spatial averages or mean amplitude responses were achieved by a process of energy averaging as follows:

$$\text{SPL}_{\text{total}} = 10 \log_{10} \left\{ \sum_{n=1}^N \frac{10^{\text{SPL}_n/10}}{N} \right\}$$

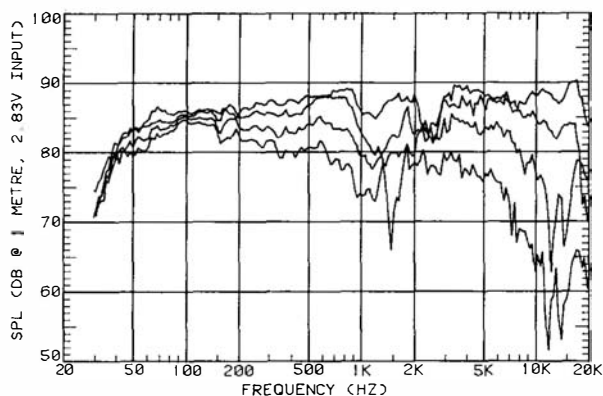
Unfortunately listening-room measurements were not possible for all of the loudspeakers, but some selected units were tested in this manner. This matter will be addressed separately in a later section.

These measures will be explained with reference to measurements performed on loudspeaker D, a conventional two-way design consisting of an 8-in (200-mm) woofer and a 1-in (25-mm) tweeter. This loudspeaker was given a mean fidelity rating of 7.5, an interesting fact to remember when reviewing these data in the context of later discussions [1, Fig. 17].

Conventional on- and off-axis amplitude-response measurements are shown in Fig. 2. From the selection of curves displayed it is evident that this kind of information, while accurate, is rather difficult to interpret. The problem is that some of the irregularities in the curves are caused by resonances within the loudspeaker,



(a)



(b)

Fig. 2. Free-field measurements on loudspeaker D at 2 m on axis and at (a) 30, 60, and 90° off axis horizontally (left) and (b) 30, 60, and 90° off axis vertically (up).

while others are associated with interference between multiple drivers, reflections, or diffraction. Resonances tend to cause similar irregularities in most measurements, while irregularities associated with interference tend to change in both form and frequency depending on the orientation of the measuring microphone. Visual "integration" by experienced eyes can sort out much of the inconsistent clutter caused by relatively innocuous interference effects, revealing the important evidence of persistent discontinuities, due to resonances, and directional trends. The novice examiner, however, is likely to be unnecessarily alarmed by what is seen.

In Fig. 3 the on-axis response is repeated for comparison with the mean of the 25 amplitude-response measurements in the front hemisphere. The smoothing achieved by this spatial averaging is substantial. The evidence of acoustical interference is considerably reduced, but the indicators of nonuniform sound output remain. Gone also is all but the most elementary directional information. This artifice yields what amounts to an axially weighted measurement of the sound propagated in the general direction of the listener, including most of the early reflections and most of the reverberant energy.

Also shown in Fig. 3 is the mean of all 34 amplitude-response measurements in the total sphere surrounding the loudspeaker. It should be noted that this is not the mean spherical response referred to by Gee and Shorter [63], who used a surface-area weighting of sound-pressure-level measurements to produce what is now called total sound power (item 8).

Fig. 4 shows processed data that incorporate spatial averaging while retaining some directional information. The $\pm 15^\circ$ "listening window" measure is the mean of the on-axis and the four 15° off-axis measurements. It represents the direct sound for listeners located within a reasonable distance from the conventional stereo seat. The $30\text{--}45^\circ$ measure averages the eight measurements within that angular range and is thus indicative of the direct sound received by listeners who might be poorly located. For the principal listeners these sounds might represent floor or ceiling reflections, and for all listeners, they represent part of the reverberant sound field. The $60\text{--}75^\circ$ measure represents sounds that, in many room arrangements, arrive at the listener's ears after reflection from sidewalls and also as reverberation. The early reflections arrive within the first 20 ms, and assuming no attenuation at the boundaries, with amplitudes 2–10 dB lower than the direct sound.

The purpose of spatial averaging is to eliminate irregularities in the curves caused by interference effects that change with microphone location, leaving intact the irregularities due to other causes. The idea is not new. For example, Hentsch [3] and Shorter [8], in the 1950s, advocated the technique which has the characteristic of retaining high resolution in the frequency domain by trading off directional resolution. The common method of smoothing a messy curve is to use one-third-octave filtered noise as a test signal, which retains spatial precision, while yielding data with low resolution

in the frequency domain. This is an unsatisfactory trade, it would seem.

The estimate of sound power (item 8) shown in Fig. 5 is derived from the 34 amplitude-response measurements, each weighted according to its contribution to the total power radiated. The directivity index (item 9) is as defined in IEC Publication 268-5: "Under free-field conditions, the ratio, expressed in decibels, of the intensity measured at a chosen point on the reference axis to the intensity that a point source radiating the same acoustic power as the loudspeaker under test would

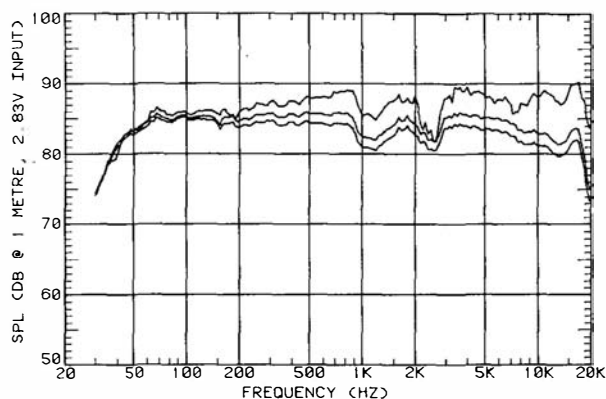


Fig. 3. Free-field measurements on loudspeaker D at 2 m. Top—on axis; middle—mean response in front hemisphere; bottom—mean response in total sphere.

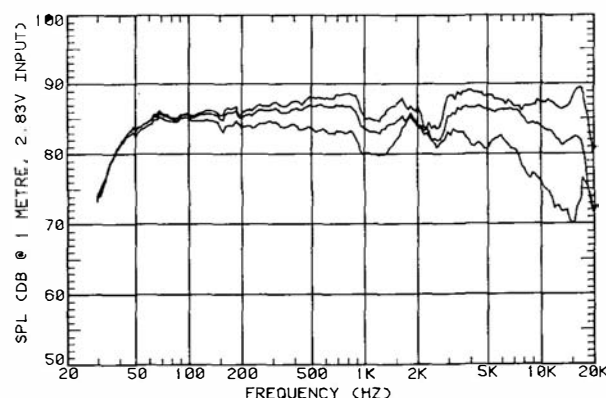


Fig. 4. Free-field measurements on loudspeaker D at 2 m. Mean responses within selected angles (horizontal and vertical combined). Top— $0\text{--}15^\circ$; middle— $30\text{--}45^\circ$; bottom— $60\text{--}75^\circ$.

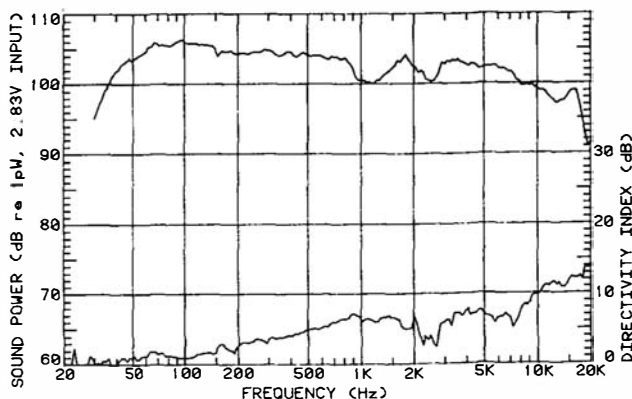


Fig. 5. Free-field measurements on loudspeaker D at 2 m. Top—calculation of sound-power output; bottom—calculation of directivity index.

produce at the same measuring position, the measurements being made at a specified frequency . . .” [64].

Phase response, shown in Fig. 6, was measured by fast Fourier transform analysis of the on-axis impulse response. Data below about 300 Hz were not reliable because of the proximity of reflecting surfaces in the measuring environment. It is possible to see the clear correlations between the localized irregularities in this response and those in the amplitude response that survived the spatial averaging process. However, it should be noted that, like the amplitude response, the phase response is dependent upon microphone location. The data shown apply only to the reference axis of the loudspeaker.

6 VISUAL CORRELATIONS OF SUBJECTIVE AND OBJECTIVE DATA

Ideally all possible measurements would have been performed on all loudspeakers that were evaluated in the subjective measurements. Unfortunately this was not possible, as the hardware and software for physical measurements were developed in parallel with those for the subjective measurements. The data shown here apply to some products that were used in the experiments of [1] and some that have been conducted since, in the course of routine product evaluations. Throughout, however, the experimental procedures and controls have been identical. The distinction that these products share is that the technical data were all acquired using the computer-controlled measurement system. Similar data on other products acquired by analog means, and consequently without the benefit of processing, were shown in an earlier presentation [65].

Ideally also, the technical measures would be reduced

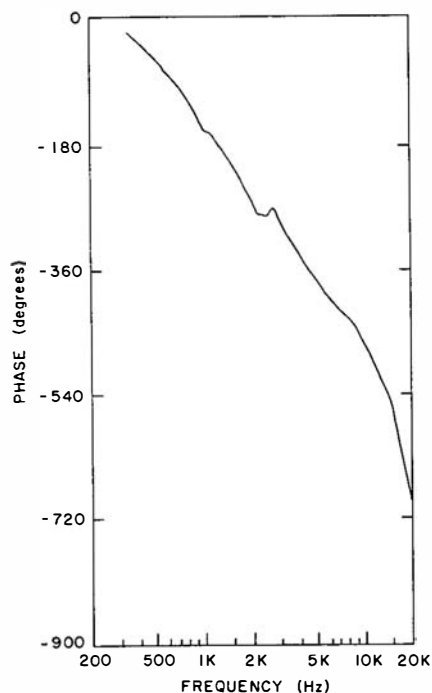


Fig. 6. Free-field measurements on loudspeaker D at 2 m. On-axis phase response.

to single-number ratings so that the correlations with the fidelity rating and other subjective measures would be simple numerical exercises. However, the dimensions of these measures are numerous, and it seemed unreasonable to contaminate the data by further processing based on speculation or simplistic assumptions. Consequently the following presentations rely on visual recognition of those features that the groupings of data have in common, and those that appear not to follow the trend of the grouping parameter.

The measured data were first organized by placing all products that received mean fidelity ratings (all listeners) within ranges of 0.5 scale unit in the same group. This resolution was selected on the basis that it represents a difference in fidelity ratings that most listeners with near-normal hearing threshold levels were able to distinguish with high statistical significance. In common parlance, loudspeakers with fidelity ratings that differed by about 0.5 of a scale unit (that is, about 5% of the total rating scale) were recognized as being sufficiently different that a preference could be stated, with substantial agreement among listeners with near-normal hearing threshold levels. Loudspeakers with fidelity ratings that differed by less than this amount, although exhibiting audible differences, tended to be similarly acceptable in an overall assessment of sound reproduction accuracy.

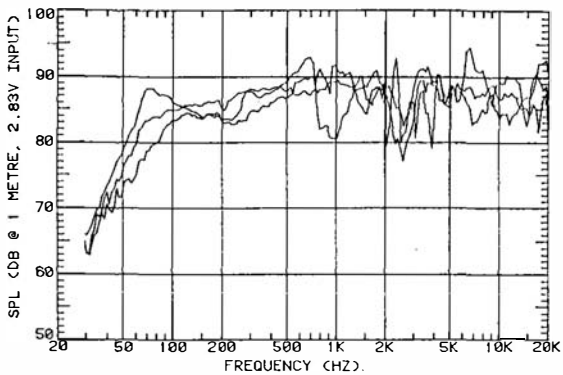
Since loudspeaker sensitivity is not a factor in sound reproduction accuracy, the data were adjusted to compensate for this factor using the mean on-axis sound level between 300 and 3000 Hz as the indicator, in a manner similar to that used in the equal-level adjustments in the listening tests.

6.1 On- and Off-Axis Amplitude Response

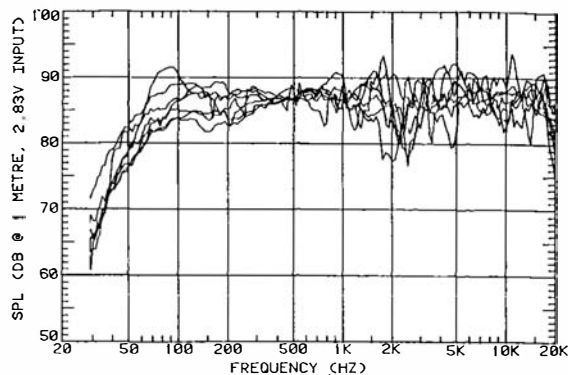
Fig. 7(a) shows measurements on the three loudspeakers in the test that received mean fidelity ratings in the range of 6.0–6.4. These were the lowest ratings in the group, and an examination of the data suggests why.

The on-axis responses are quite irregular, with numerous sharp discontinuities. Bass output ranges from fairly weak to rather resonant, and none of these products has much output below about 60 Hz. The spatially averaged measurements between 30 and 45° off axis have disguised most of the effects of acoustical interference, so that the remaining irregularities are indications of more serious problems. Also evident is the attenuation of high-frequency output due to the directionality of the tweeters. The 60–75° averaged curves also reveal many of the same features that are present in the other data, including further high-frequency roll-off.

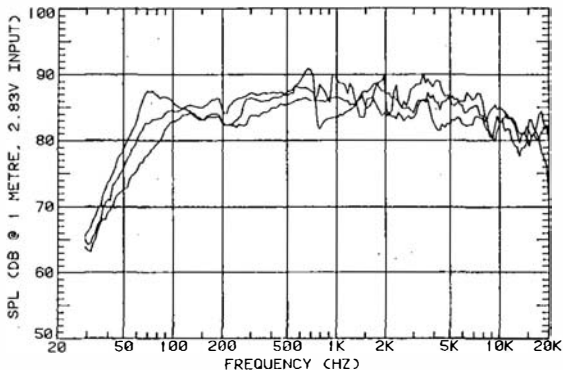
The persistence of certain features suggests that the problems are not minor, that they will be present in the direct, early reflected, and reverberant sound fields in the listening room. The sharpness of the discontinuities, even after spatial averaging, implies underlying resonances with the attendant time-domain and phase



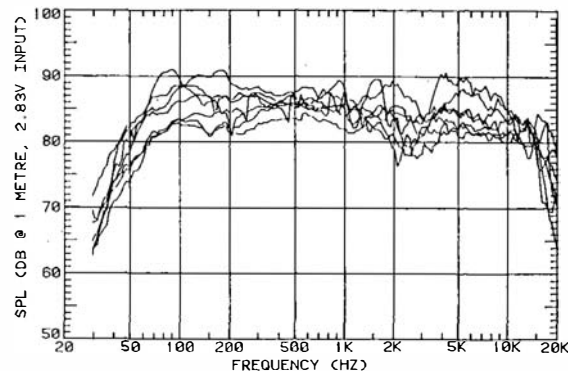
ON-AXIS FREQUENCY RESPONSE
THREE LOUSPEAKERS WITH FIDELITY RATINGS OF 6.0-6.4



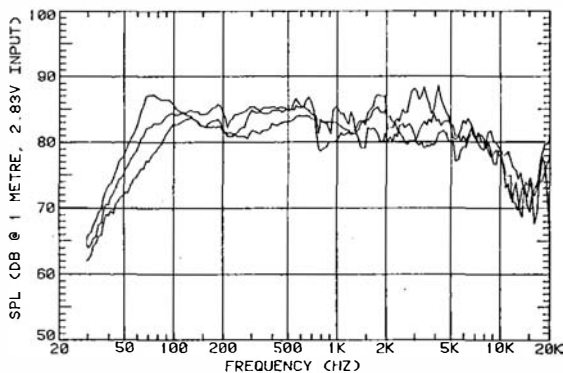
ON-AXIS FREQUENCY RESPONSE
SEVEN LOUSPEAKERS WITH FIDELITY RATINGS OF 6.5-6.9



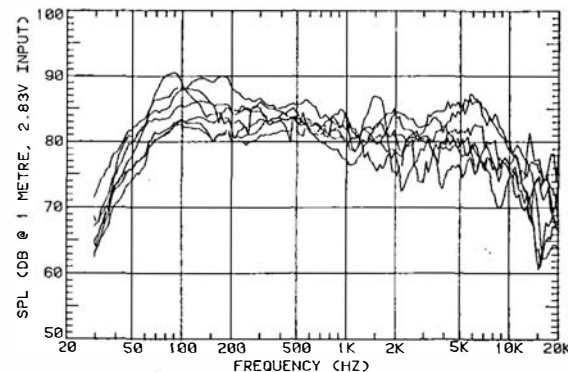
AVERAGE FREQUENCY RESPONSE 30 TO 45 DEG. OFF AXIS HOR. AND VERT.
THREE LOUSPEAKERS WITH FIDELITY RATINGS OF 6.0-6.4.



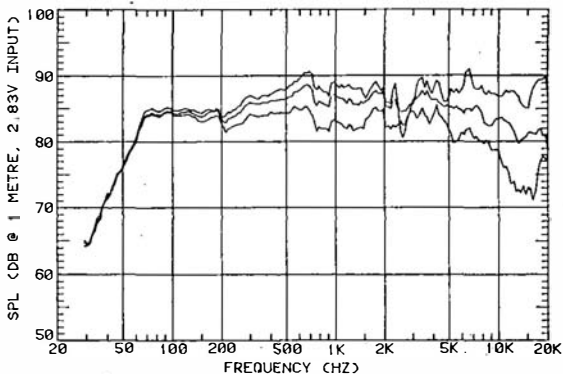
AVERAGE FREQUENCY RESPONSE 30 TO 45 DEG. OFF AXIS HOR. AND VERT.
SEVEN LOUSPEAKERS WITH FIDELITY RATINGS OF 6.5-6.9.



AVERAGE FREQUENCY RESPONSE 60 TO 75 DEG. OFF AXIS HOR. AND VERT.
THREE LOUSPEAKERS WITH FIDELITY RATINGS OF 6.0-6.4.

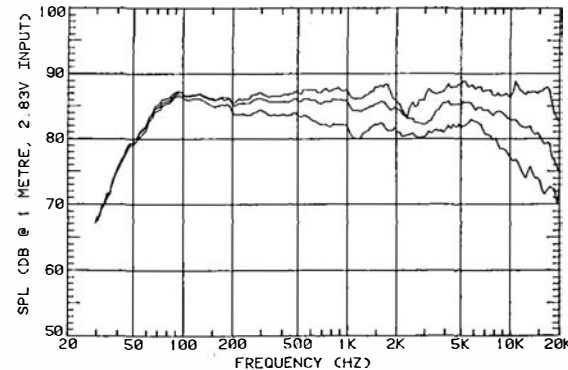


AVERAGE FREQUENCY RESPONSE 60 TO 75 DEG. OFF AXIS HOR. AND VERT.
SEVEN LOUSPEAKERS WITH FIDELITY RATINGS OF 6.5-6.9.



TOP-TO-BOTTOM, AVG. ON AXIS, AVG. 30-45 DEG., AVG. 60-75 DEG.
THREE LOUSPEAKERS WITH FIDELITY RATINGS OF 6.0-6.4.

(a)

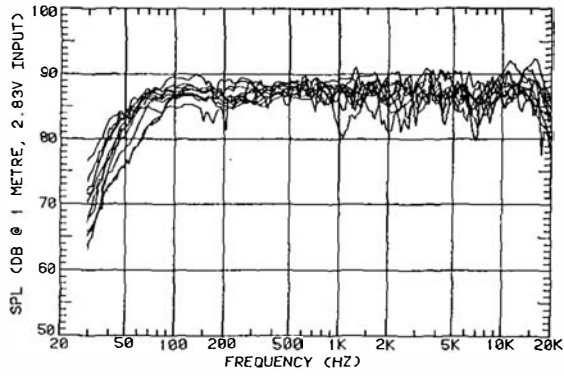


TOP-TO-BOTTOM, AVG. ON AXIS, AVG. 30-45 DEG., AVG. 60-75 DEG.
SEVEN LOUSPEAKERS WITH FIDELITY RATINGS OF 6.5-6.9.

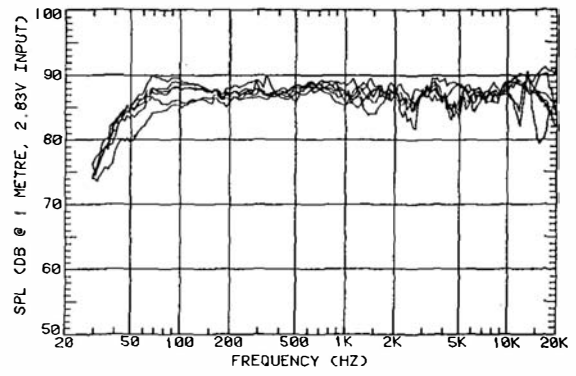
(b)

Fig. 7. Amplitude response measurements of loudspeakers with fidelity ratings (a) 6.0-6.4.

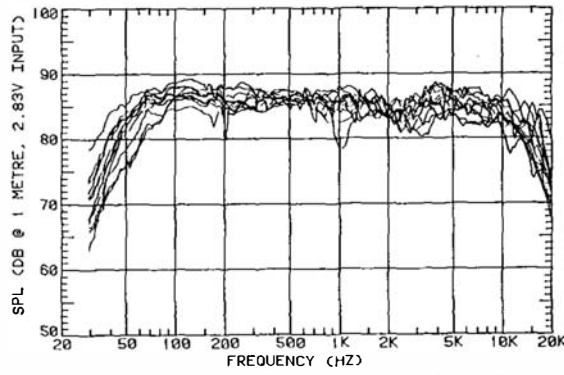
Fig. 7. Amplitude response measurements of loudspeakers with fidelity ratings (b) 6.5-6.9.



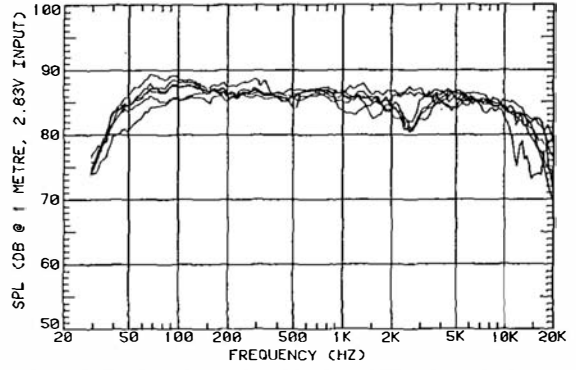
ON-AXIS FREQUENCY RESPONSE
ELEVEN LOUSPEAKERS WITH FIDELITY RATINGS OF 7.0-7.4



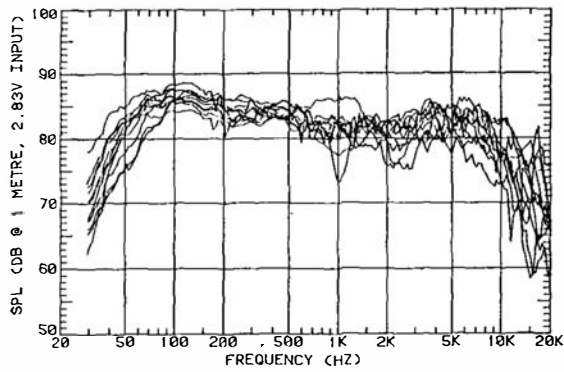
ON-AXIS FREQUENCY RESPONSE
SIX LOUSPEAKERS WITH FIDELITY RATINGS OF 7.5-7.9



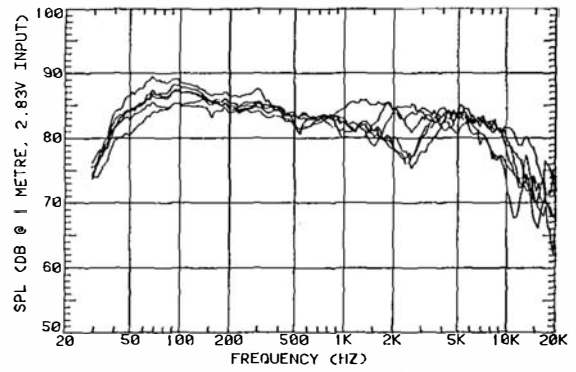
AVERAGE FREQUENCY RESPONSE 30 TO 45 DEG. OFF AXIS HOR. AND VERT.
ELEVEN LOUSPEAKERS WITH FIDELITY RATINGS OF 7.0-7.4.



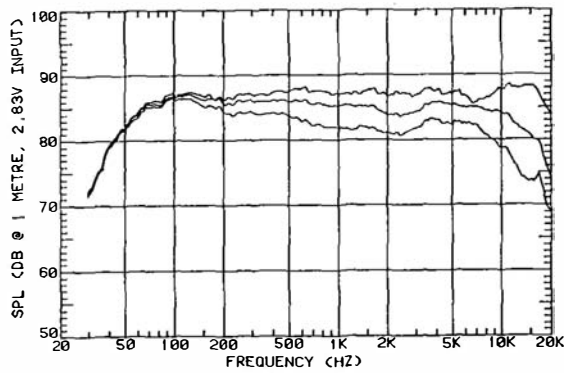
AVERAGE FREQUENCY RESPONSE 30 TO 45 DEG. OFF AXIS HOR. AND VERT.
SIX LOUSPEAKERS WITH FIDELITY RATINGS OF 7.5-7.9.



AVERAGE FREQUENCY RESPONSE 60 TO 75 DEG. OFF AXIS HOR. AND VERT.
ELEVEN LOUSPEAKERS WITH FIDELITY RATINGS OF 7.0-7.4.

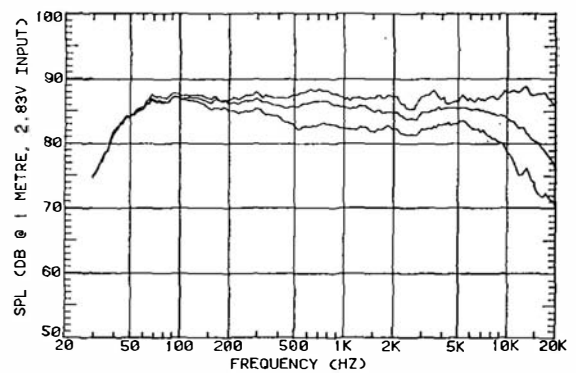


AVERAGE FREQUENCY RESPONSE 60 TO 75 DEG. OFF AXIS HOR. AND VERT.
SIX LOUSPEAKERS WITH FIDELITY RATINGS OF 7.5-7.9.



TOP-TO-BOTTOM, AVG. ON AXIS, AVG. 30-45 DEG., AVG. 60-75 DEG.
ELEVEN LOUSPEAKERS WITH FIDELITY RATINGS OF 7.0-7.4.

(c)



TOP-TO-BOTTOM, AVG. ON AXIS, AVG. 30-45 DEG., AVG. 60-75 DEG.
SIX LOUSPEAKERS WITH FIDELITY RATINGS OF 7.5-7.9.

(d)

Fig. 7. Amplitude response measurements of loudspeakers with fidelity ratings (c) 7.0-7.4.

Fig. 7. Amplitude response measurements of loudspeakers with fidelity ratings (d) 7.5-7.9.

anomalies.

The bottom curves are the averages of each of the previous three groups of curves, suggesting possibly the design objectives for a product of this caliber: restricted bass output, moderately directional high-frequency output, and a consistent unevenness through the midrange. The actual shape of the unevenness is probably not so important as its magnitude, as all three products appeared to have rather clearly identifiable "signatures." In short, the design objectives for products with low fidelity ratings appear not to be very demanding.

Moving up the fidelity rating scale to those products that were rated between 6.5 and 6.9, we see in Fig. 7(b) many of the same characteristics identified in the previous group, but with small differences. The bass output is generally more extended; the midrange, about 300 Hz to 2 kHz, shows a consistently flatter tendency that, with increased variability, extends to the highest frequencies in the on-axis curves. Off-axis responses continue to show directional high-frequency roll-off, and the irregularities in the off-axis measurements continue to echo many of the distinctive features seen in the axial responses. Overall, however, these products exhibit less severe output fluctuations than those seen in the previous group.

The averaged curves at the bottom of the presentation show clearly that the "design objectives" are somewhat more demanding, although the tolerances are generous enough to permit some artistic input.

With the eleven loudspeakers given fidelity ratings between 7.0 and 7.4 we have reached the level of pleasant-to-good listening, a fact that one might have suspected by reviewing the data [Fig. 7(c)]. In all respects previously noted these products are superior. There is no longer any doubt that a flat, uniform amplitude response was the objective and that the permissible deviations were fairly small. With only individual localized exceptions, all eleven products *combined* show on-axis amplitude responses that are within 3 dB of a horizontal line at the 87-dB level from about 80 Hz to 16 kHz. Over much of that frequency range most products did even better. In the off-axis responses these products also distinguish themselves. The curves drift away from the horizontal, due to the directivity of the drivers, but there is still considerable control of the quality of output, even at 60–75° off axis.

In contrast, below 100 Hz the spread is considerable, especially considering that at low frequencies there is a substantial change in apparent loudness for a small change in sound-pressure level. (In the familiar equal-loudness contours the curves crowd together at low frequencies [66].) Overall, however, the averaged curves show a small extension of bass output compared to the previous sets of data.

The smoothness and flatness of the curves and the rather small tolerances that have been noted suggest that designers of loudspeakers that are to achieve this level of subjective praise will have to pay rather close

attention to virtually all aspects of amplitude-response measurement. Interestingly, though, the tolerances at very low and very high frequencies are still rather large, an observation which, when reversed, points up the fundamental importance of performance in the middle six to seven octaves.

The six loudspeakers awarded mean fidelity ratings between 7.5 and 7.9 are in the highest category of subjective praise [Fig. 7(d)]. The excellence of their sound reproduction is noteworthy, but so also is their similarity. With many musical sounds these loudspeakers were easily confused. In this presentation the measurements also might be confused. Apart from the off-axis depression in the 2–4-kHz range that is characteristic of many two-way designs, the curves exhibit more areas of similarity than difference. In this group, with one exception, even the bass responses are closely similar. Some manufacturers would be content to see this kind of consistency in the long-term production of any one model, yet these six loudspeakers came from six different designers in six different companies in three different countries.

Although the groups of curves are useful to identify general trends in performance, a more detailed analysis requires data on individual loudspeakers. Fig. 8 shows measurements on the 20 loudspeakers that received fidelity ratings in the range of 6.8–7.9. This range represents about two standard deviations in the fidelity ratings of listeners with near-normal hearing threshold levels.

Fig. 8 displays, from left to right, the single on-axis measurement, the mean amplitude response in the "listening window" of $\pm 15^\circ$ horizontally and vertically (a five-curve spatial average), and eight-curve spatial averages representing the mean amplitude responses in the ranges 30–45° and 60–75° off axis. Also shown are the –10-dB low-frequency cutoffs measured relative to the mean on-axis amplitude response in the 300–3000-Hz range, and the retail price in Canadian dollars at the time of the tests.

Inspecting the curves in each of the columns, it is possible to see a progressive increase in the smoothness as a function of the fidelity ratings. The loudspeakers with lower ratings tend to exhibit more fine structure in the curves which, in the spatially averaged data, indicates the presence of resonances within the loudspeaker system. When these detailed imperfections are combined with more general spectral variations, the ratings tend to be still lower.

The small amount of spatial averaging in the 0–15° curves appears to be helpful in removing some of the interference clutter that confuses the unprocessed on-axis (0°) curves. Sharp discontinuities that are visible in the eight-curve spatial averages (30–45° and 60–75°) carry a significance that far exceeds a similar discontinuity in a single measurement. Almost all of the lower rated loudspeakers exhibit such problems.

The careful observer may detect some apparent anomalies in the data, with some loudspeakers looking as though they ought to have been higher or lower in

the subjective ratings. Two points are relevant in these instances. First, these data pertain only to amplitude response; performance in another domain can be responsible for the anomalous appearance. Loudspeaker 12, for example, looks as though it might have warranted a higher subjective rating. Indeed, listeners generally praised the lack of coloration and the wide bandwidth of this unit. However, they also noted excessive distortion at high sound levels, particularly in the low frequencies. This was confirmed in measurements of nonlinear distortion. The basis for the problem is that the designer chose to restrict the dynamic range in order to achieve an extended low-frequency range—a reasonable trade-off in the context of small, inexpensive loudspeakers.

The second point relates to the narrow range of fidelity ratings embraced by these products and to the fact that several of them have identical or very similar ratings. As carefully pointed out in the previous paper [1], even with the close experimental controls of these listening tests, ratings that differ by much less than about 0.5 fidelity rating unit may not be significant. In other words, repeating the listening tests may result in a rearrangement in the ranking of closely rated loudspeakers.

If these factors are not at issue, the differences in rating may be traceable to features in the amplitude-response data. For example, loudspeaker 16 appears to have relatively few imperfections compared to some of the neighboring curves. However, a close examination reveals that the off-axis response at frequencies above about 2 kHz rolls off more rapidly than is typical for loudspeakers with higher ratings. It is perhaps significant that in this case, the smoothness of the off-axis response is not an adequate compensation for the lack of output. Listeners commented on the generally mellow or muted sound of this loudspeaker, a characteristic that would not have been anticipated from an examination of the axial response.

Loudspeaker 11 exhibited amplitude responses that in most respects were commendable. However, a single small peak at about 1.3 kHz can be seen in the on-axis response and in the spatially averaged 0–15° response. The fact that the feature survived spatial averaging suggests that it is a possible source of coloration; that it does not appear in measurements further off axis indicates the directionality of the acoustical output of the misbehaving element. This feature may have been responsible for listener reports of slight midrange coloration.

In the latter example spatial averaging enabled the observer to distinguish between an irregularity in the on-axis curve that could have been associated with acoustical interference and one that was, in fact, caused by a resonance in the woofer. The directionality of the output from the large woofer diaphragm at that frequency is another clue that can be seen in the measurements. Interpretations of other measurements can also benefit from spatial averaging. Some of the irregularities seen in the on-axis responses of loudspeakers

3, 8, and 12, for example, disappear with spatial averaging within the 15° listening window. However, other features persist in all of the combination measurements, emphasizing the severity of the problems.

Loudspeaker 15 exhibits a number of detailed irregularities that are clearly recognizable in all of the measurements. The clear evidence of these imperfections relies upon the data having high resolution in the frequency domain. If the data were one-third-octave filtered, as is commonly the case when data are combined for spatial averaging or sound-power integration, much of the evidence would be lost. Fig. 9 shows the effect of one-third-octave swept filtering on the on-axis amplitude response of loudspeaker 15; the smoothing effect is considerable. Even more information would be lost in measurements using filters with fixed center frequencies, resulting in the familiar “staircase” curves, with a horizontal plateau extending over each one-third-octave band.

It was noted in discussions of Fig. 7, and it is evident in the presentation of Fig. 8, that there is a relationship between low-frequency response and fidelity rating. Loudspeakers with high fidelity ratings tended to have the most extended bass response. Looking at this factor in isolation revealed a moderate correlation between the low-frequency cutoff and the fidelity rating. Fig. 10 shows the scatter diagrams, linear regressions, and correlation coefficients for the top 23 loudspeakers in the ratings. Loudspeakers with lower ratings performed so erratically through the midrange that it was difficult to arrive at a reference level against which to compare performance at the low-frequency extreme. For the loudspeakers shown, the reference level was the mean on-axis amplitude response in the 300–3000-Hz range, selected on the basis that this is the essential voice-frequency range and that it includes much of the frequency range of maximum auditory sensitivity. The low-cutoff frequencies were determined at levels 5, 7.5, and 10 dB below the reference level. The respective correlation coefficients were -0.35 , -0.46 , and -0.50 ; the data for two of these are shown in Fig. 10.

Evidently, listeners regard the extension of low-frequency response to be one of the important factors determining overall loudspeaker performance. As a single-number rating of this parameter, the -10 -dB level, compared to the 300–3000-Hz midband level, would appear to have a slight advantage over measures based on smaller attenuations.

6.2 Sound-Power Output and Directivity Index

A visual integration of the data presented in Figs. 7 and 8 suggests much of what is now about to be seen. Calculations of sound power and directivity index at the 200 test frequencies show that the loudspeakers in the upper performance categories exhibit sound-power and directivity-index curves with fewer small-scale irregularities (Fig. 11).

Because of the similarity in physical dimensions and design configurations of these loudspeakers, the underlying shape of the collections of curves is rather

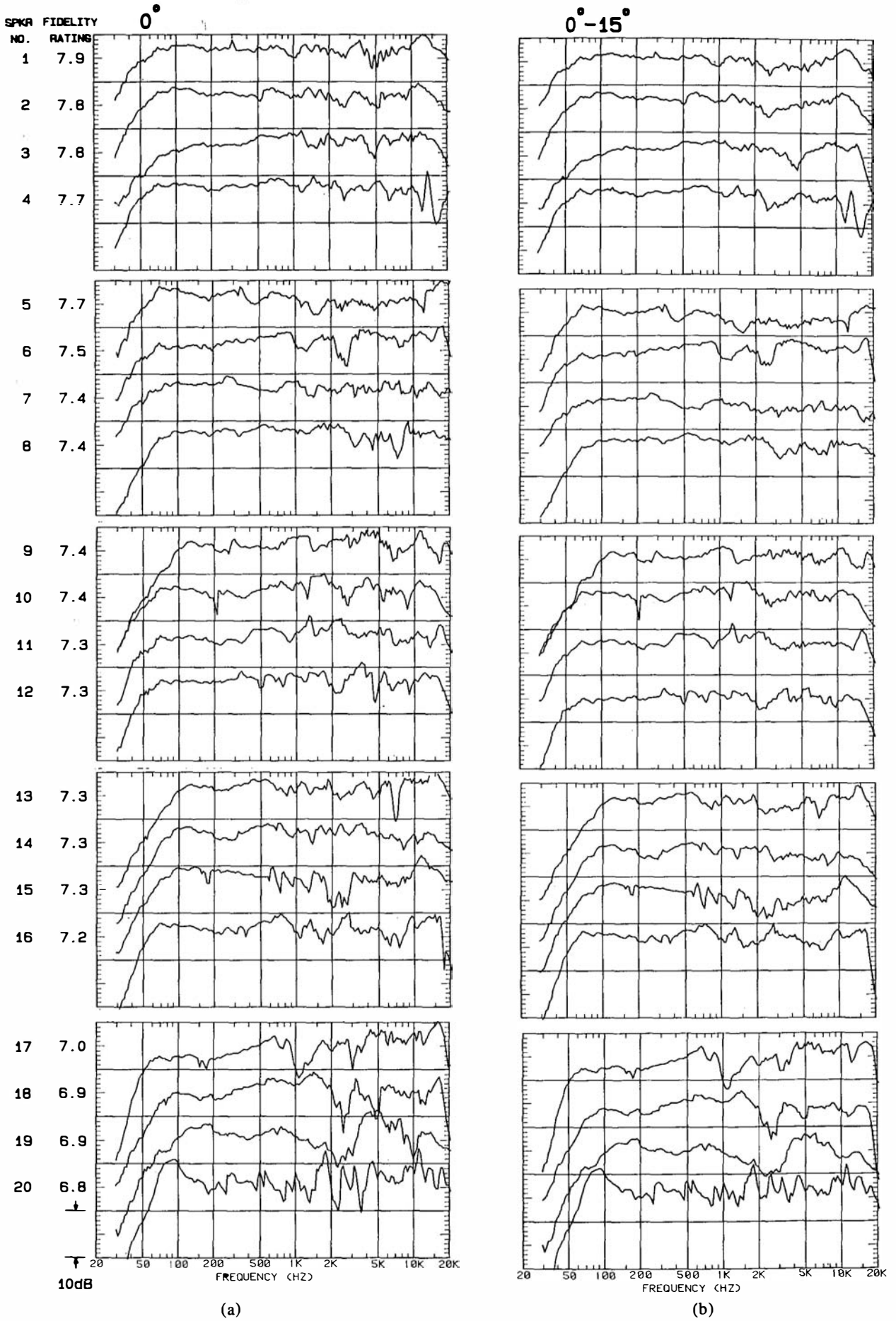


Fig. 8. Amplitude response measurements of 20 loudspeakers with fidelity ratings between 6.8 and 7.9.

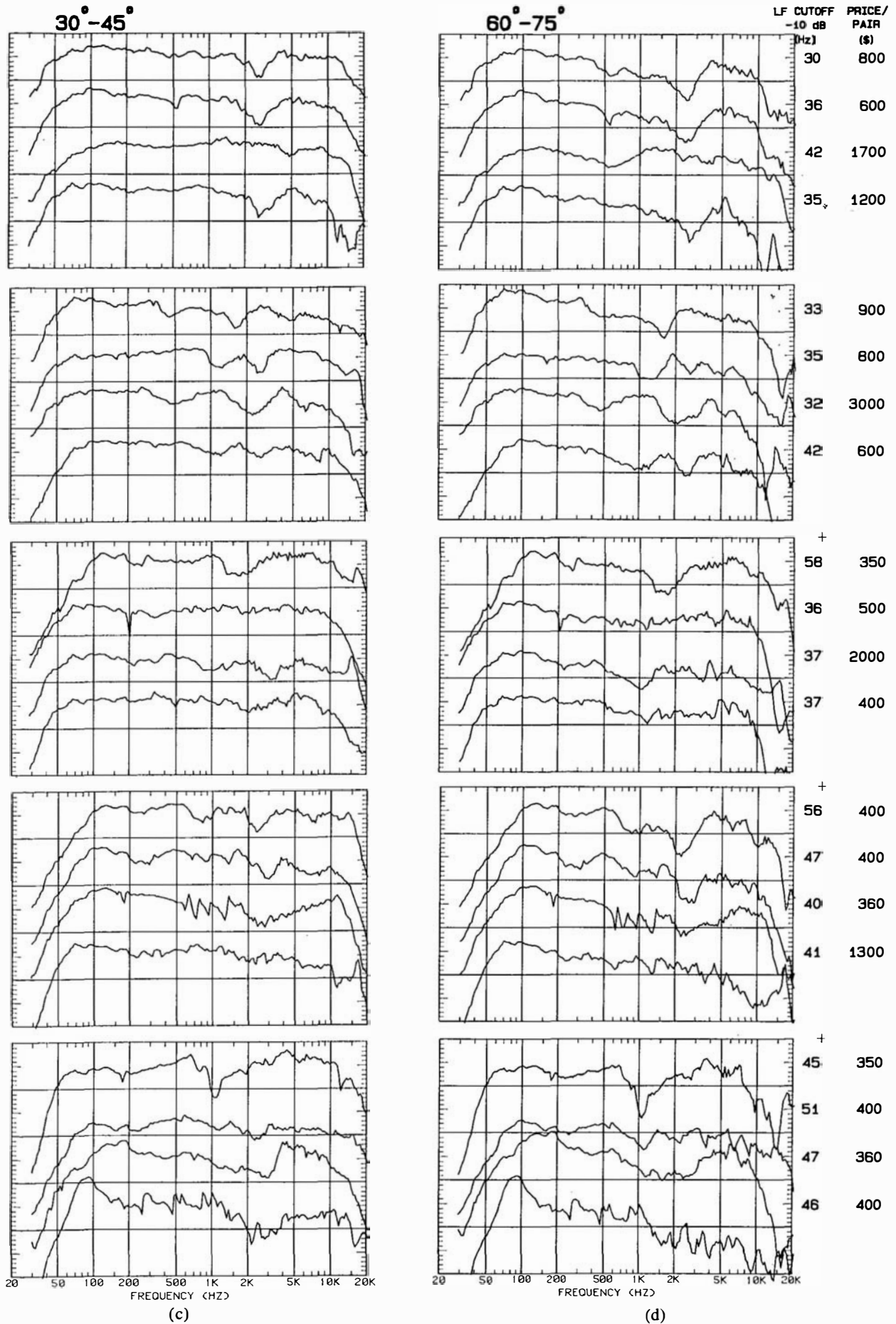


Fig. 8. Amplitude response measurements of 20 loudspeakers with fidelity ratings between 6.8 and 7.9.

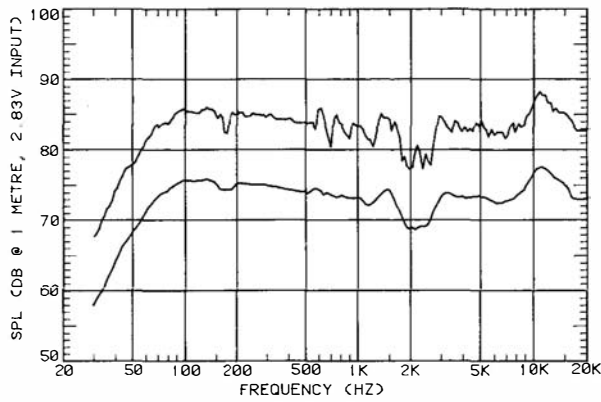


Fig. 9. Comparison of on-axis amplitude responses of loudspeaker 15 with (bottom) and without (top) one-third-octave swept filtering.

similar. There is, however, a trend toward better bass and less erratic curves in loudspeakers with higher fidelity ratings.

Sound-output irregularities strong enough to survive the spatial averaging of 34 separate amplitude-response measurements are probably indicative of audible colorations. Some of these are clearly rather narrow-band problems, and other wider bandwidth effects show steep, almost vertical transitions between levels. These clues indicate time- and phase-domain aberrations, and it is these clues that are disguised in measurements performed using one-third-octave or octave-band analysis of pink noise. Such measurements, in themselves, therefore have reduced discrimination and need the support of other techniques that examine performance explicitly in the time and phase domains.

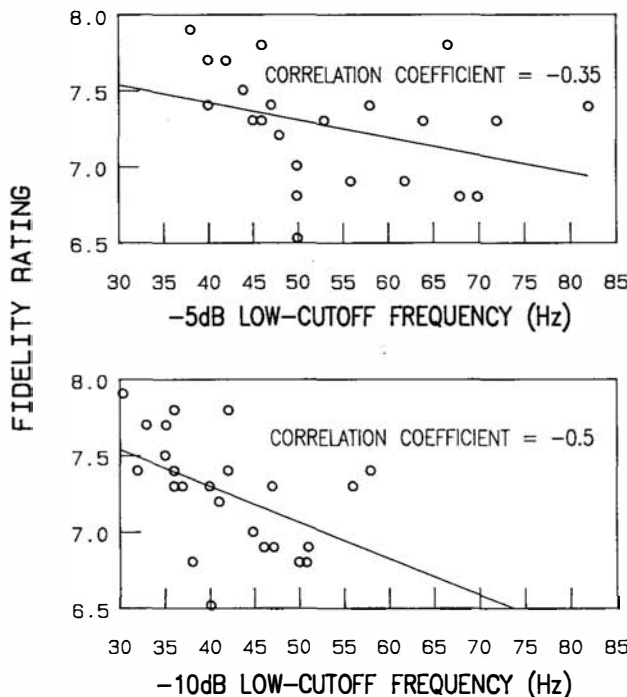
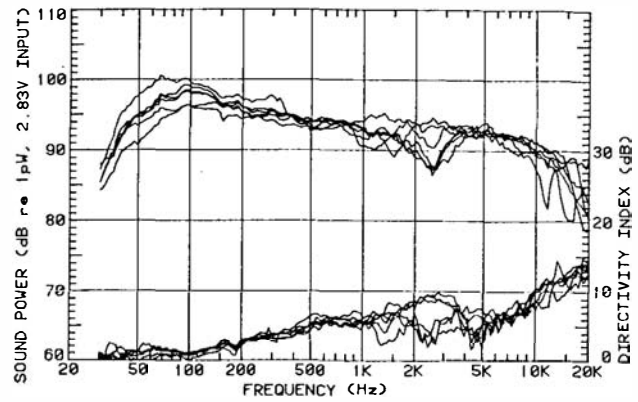
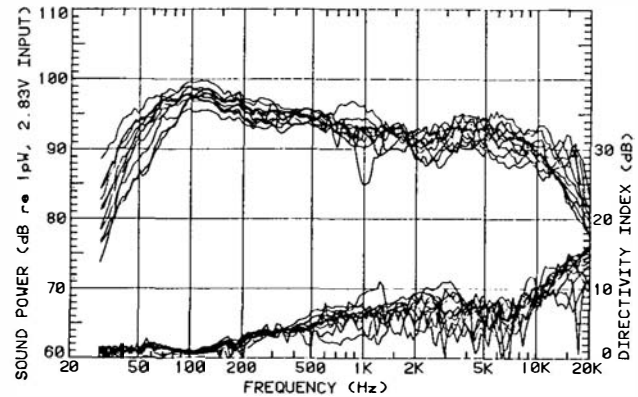


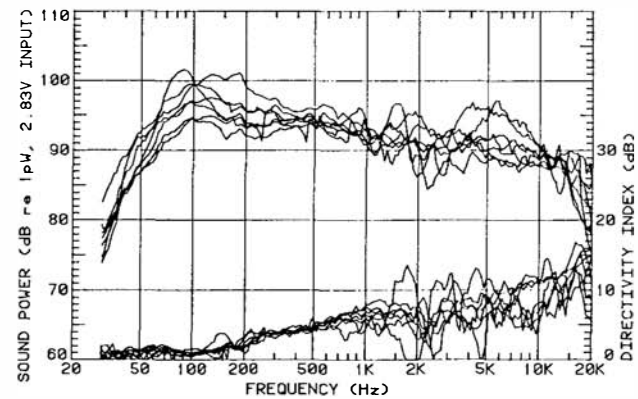
Fig. 10. Relationship between fidelity rating and low-frequency cutoff measured at -5 and -10 dB relative to mean on-axis amplitude response between 300 and 3000 Hz.



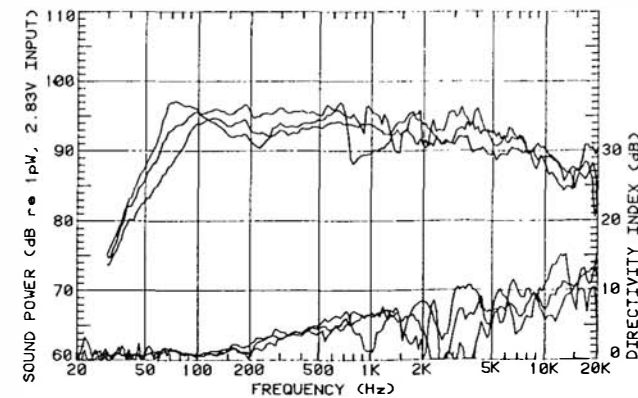
FIDELITY RATING 7.5-7.9



FIDELITY RATING 7.0-7.4



FIDELITY RATING 6.5-6.9



FIDELITY RATING 6.0-6.4

Fig. 11. Sound-power response and directivity index for loudspeakers in four preference groups.

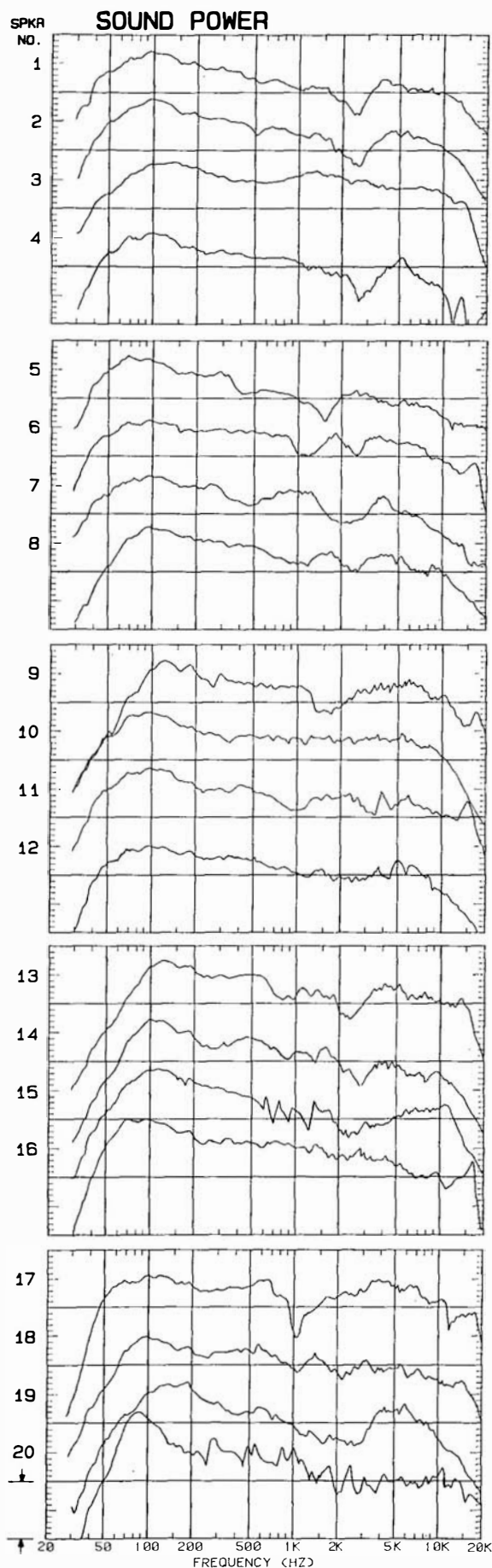


Fig. 12. Sound-power response for 20 loudspeakers with fidelity ratings between 6.8 and 7.9.

A display of the individual sound-power responses for the top 20 loudspeakers (Fig. 12) is supportive of these observations. In fact, if the fine structure were removed from these curves by one-third-octave or wider bandwidth filtering, it would be difficult to discern any trend whatsoever in the data. Some of the lowest rated loudspeakers, such as loudspeakers 17, 19, and 20, do however appear to have some prominent spectral excesses or deficiencies that would be seen even in filtered data.

6.3 Phase Response

Fig. 13 shows phase-response measurements for most of the loudspeakers in the performance groups. Technical problems resulted in the loss of these data on four of the products.

Here it seems that it is not the overall form of the phase response but rather the presence of discontinuities that appears to correlate with the performance ranking. In general, the best sounding loudspeakers had phase responses that were smooth and only gently undulating in contour. Nonminimum-phase behavior and group delay were not explicitly examined in this set of products.

6.4 Discussion

The purpose of this investigation was to see if there were obviously recognizable characteristics in the objective data that could be construed as confirming the fidelity ratings resulting from subjective measurements using listeners with basically normal hearing. The conclusion is that there are.

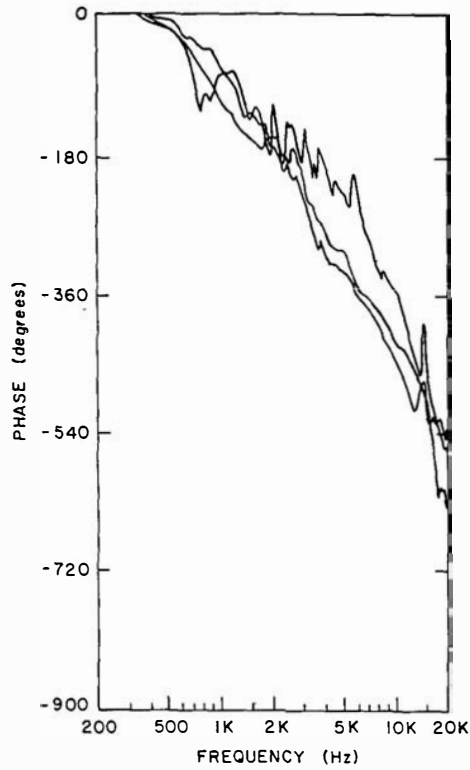
In fact there are strong suggestions that, with a few rather basic instructions, observers could rank-order the measured data in a manner that would closely parallel the subjective rankings of the products. Listeners, it seems, like the sound of loudspeakers with a flat, smooth wideband on-axis amplitude response that is maintained at substantial angles off axis. If this is achieved, the loudspeakers will exhibit smooth (but not flat) sound-power responses and directivity indices. The phase responses will also tend to be smooth, but not of any particular overall shape.

In a precursor to the present paper [65] a similar presentation of swept-tone anechoic data on 23 loudspeakers led to very much the same conclusion, without the niceties of computer processing.

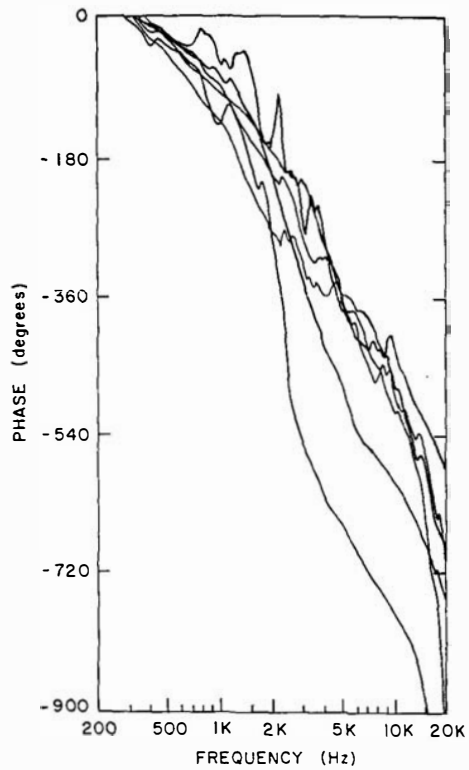
What we do not learn from these data are the priorities of these performance measures. Need they all be *equally* good? Neither do we see the potential effects of the listening room, although the effects of one particular room are included in the subjective ratings of these products.

7 LOUDSPEAKER PERFORMANCE IN THE LISTENING ROOM

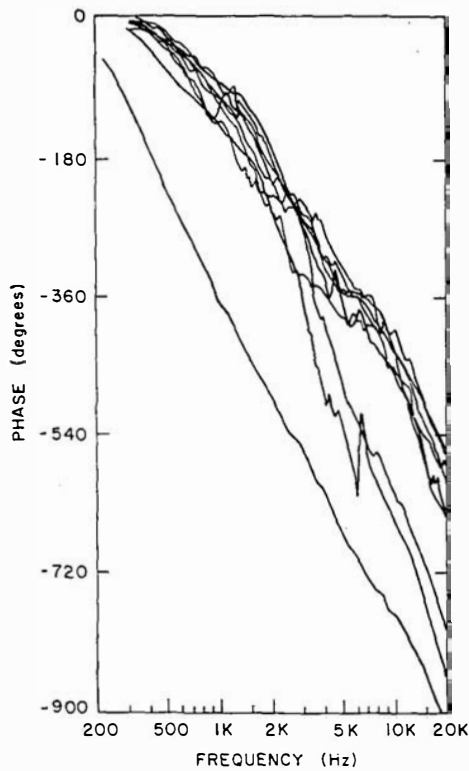
Sound propagation in rooms is a study in itself, and this presentation will not attempt to be comprehensive.



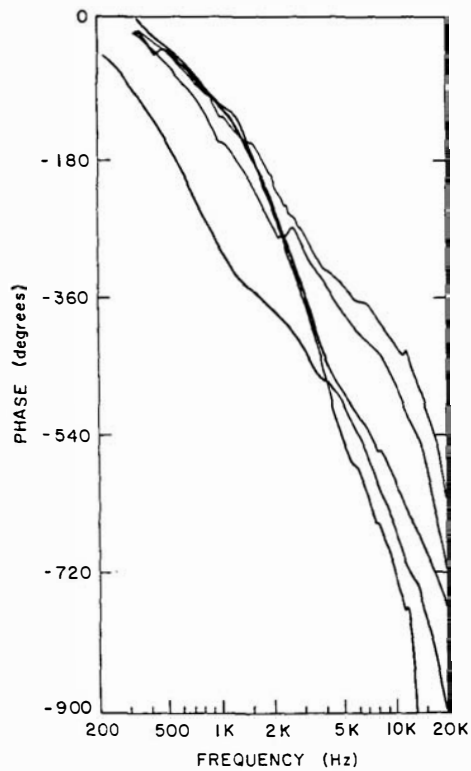
FIDELITY RATING 6.0 - 6.4



6.5 - 6.9



FIDELITY RATING 7.0 - 7.4



7.5 - 7.9

Fig. 13. Phase responses for loudspeakers in four preference groups. Some of the data may contain residual time-delay error. This will affect the overall slopes of the curves but not the contours and fine structure.

In general, though, it may be said that what is heard and what is measured in a listening room are combinations of the direct sound, the slightly delayed and attenuated early reflections, and the progressively delayed and attenuated sounds that comprise the decaying reverberant sound field.

While there is agreement that listening rooms influence what is heard, there are different hypotheses about the relative weightings of the three classes of sound described. A measuring microphone performs a simple transduction of the pressure summation at the diaphragm location, without regard for the direction or timing of the incident sounds. Two ears and a brain, however, are rather more elaborate in their processing. The external ear encodes the incoming sound with spectral modifications characteristic of the direction of incidence. The processing in the individual ears and in the binaural hearing mechanism contains directional and temporal masking effects that emphasize, in certain respects, the importance of early sounds over similar sounds that arrive later, and over others arriving from different directions. The literature on sound localization has extensive coverage of these phenomena. ([56] and [67] are good recent surveys of knowledge in this field.)

At low frequencies acoustical interference due to the room boundaries adjacent to the loudspeaker causes variations in the sound power supplied to the listening area [68]. This situation is further complicated by complex frequency-dependent standing-wave patterns in the room and the effect of adjacent boundaries on sounds at the listener or microphone location. Fortunately these effects tend to diminish in the high-density room modes above 200–300 Hz.

All of these factors affect what is heard and what is measured, but in different ways. The challenge is to find physical measures that are good representations of the important perceptions. In addition, it would be especially useful if these measures could be related to free-field data with their inherent high resolution and precision.

7.1 Listening-Room Measurements

The listening room used in the subjective measurements has been described in an earlier paper [69, App.]. It is 6.7 by 4.1 by 2.8 m (22 by 13.5 by 9.2 ft) with a reverberation time of 0.34 ± 0.085 s from 250 Hz to 4 kHz, rising to 0.85 s at 40 Hz and falling to 0.25 s at 10 kHz. It has a hard ceiling, carpet on the floor, cupboards and bookshelves on the walls beside and behind the listeners, drapes over the wall behind the loudspeakers, and upholstered chairs. The sidewalls between the loudspeakers and the listeners are flat and highly reflective at audio frequencies. It is, in fact, the prototype of the IEC-recommended listening room [59], intended to represent a good domestic listening environment in a style typical of much of North America and Europe (Fig. 14).

The measurements shown here were made with the microphone, a Brüel & Kjaer 4134, directed toward the ceiling. (A random-incidence microphone should

be oriented so that the direct sound is incident along an axis of flat amplitude response.) It was set at listener ear height (1 m) and positioned in turn at six locations within the listening area. The loudspeakers were placed, in turn, at locations A, B, and C, differing in distance from the sidewalls and the end walls and the floor.

Fig. 15 shows results of measurements of this kind on three loudspeakers that listeners rated between 7.5 and 7.9. Each of the middle and top curves is the combination of 18 stepped-tone measurements, incorporating spatial averaging at both source and receiver locations. This disguises most of the position-dependent acoustical interference effects at middle and high frequencies, but it includes the adjacent boundary interferences and standing-wave effects that are characteristic of the restricted positions imposed by the requirements for good stereo listening.

The middle curves show that spatial averaging alone removes most of the violent amplitude fluctuations

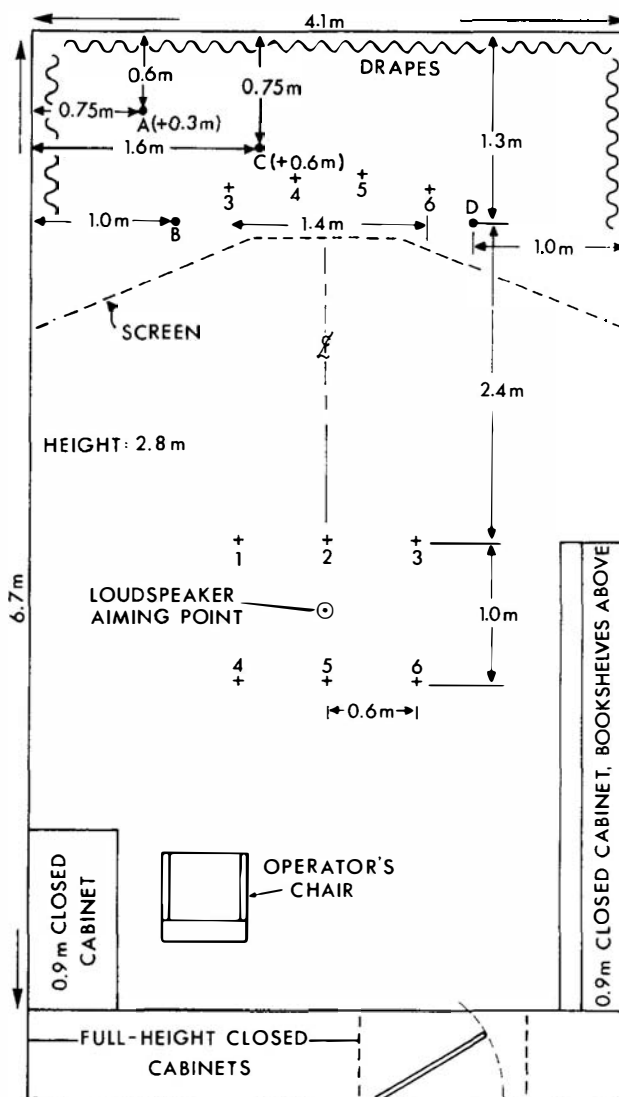


Fig. 14. Plan view of listening room showing loudspeaker and listener locations. In monophonic tests, loudspeakers were located at positions 3, 4, 5, and 6 and listeners at positions 1, 3, and 5. In stereophonic tests, loudspeakers were at positions B and D and listeners at positions 2 and 5.

caused by acoustical interference, normally seen in pure-tone room measurements. Constant factors such as low-frequency room resonances at 45, 52, 62, 74, 98, and 130 Hz are nevertheless clearly in evidence. These and other details are lost in the one-third-octave smoothed versions of the curves shown at the top of Fig. 15. The gently undulating curves retain some distinctive features but display a broadly similar overall trend that is better seen in the bottom curve, the mean of all three sets of measurements.

If room measurements were indeed the dominant measure of loudspeaker sound quality, this mean curve could be regarded as the design objective for high-quality loudspeakers measured in this room. However, the amount of variation seen in the individual performances (top curves) suggests a rather limited precision. The design objective is therefore not very demanding, and this would apply to the intrinsic design of the loudspeaker as well as to attempts at corrective equalization by electronic means. Performance evaluation by such measurements alone would be rather imprecise.

Figs. 16 and 17 show measurements on the same loudspeaker at the specific positions used in the mono and stereo tests. Two features should be noted. First, the large dip at 60 Hz, and smaller ones at multiples thereof, are caused by a combination of destructive interference between the direct sound and sound reflected from the rear wall, and of microphone (ear) locations that are close to a null in the vertical standing-wave system. Second, the shallow depression between 300 and 500 Hz is associated with the floor reflection which, because of the sound absorption by the carpeting, causes partial cancellation at the microphone locations. In Fig. 17 the small dip at about 150 Hz appears to be associated with the sidewall reflection for this specific loudspeaker location.

In all, it is fairly clear that at frequencies below about 500 Hz the room characteristics dominate, with the range of 300–500 Hz influenced by the strong floor

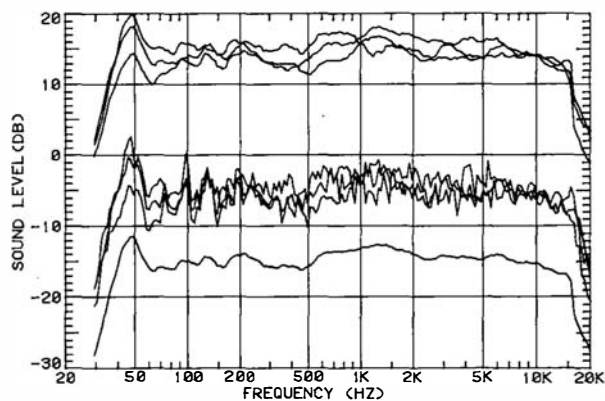


Fig. 15. Room amplitude-response measurements for three loudspeakers with fidelity ratings between 7.5 and 7.9. Each curve in top two groupings is the mean of measurements at listener positions 1 to 6 for each of the loudspeaker locations A, B, and C (18 measurements). Top curves show result of both spatial and one-third-octave spectral averaging. Middle curves have spatial averaging only. Bottom curve is the mean of all three sets of measurements with one-third-octave smoothing.

reflection and lower frequencies dictated by a combination of wall and ceiling reflections and room standing waves. None of this is new, but it is reassuring to be able to be analytical in the room situation without recourse to time-domain analyses. In this respect, measurements offering high resolution in the frequency domain as well as spatial averaging are distinctly beneficial.

7.2 Synthesis of Room Response from Free-Field Data

If cost, time, and complexity were disregarded, it would be possible to synthesize the sound field at a microphone location from a summation of sound "rays" emanating from the loudspeaker in all directions and individually followed through all the boundary interactions to arrive, finally, at the microphone. This is clearly not a practical procedure, so some simplification is in order.

A superficial examination of the specific situation involving positions A, B, and C in the room used here revealed that after the direct sound, there were a few reflections strong enough to be of individual importance, and then there were the remainder that contributed to the reverberant sound field. Using loudspeaker 7 as an

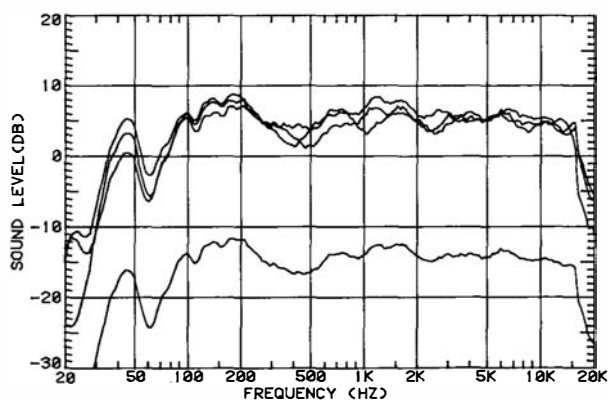


Fig. 16. Room amplitude-response measurements for three loudspeakers with fidelity ratings between 7.5 and 7.9. Each of the top curves is for loudspeaker locations 3, 4, 5, and 6 (12 measurements). Bottom curve is the mean of the three measurements. All curves are one-third-octave smoothed.

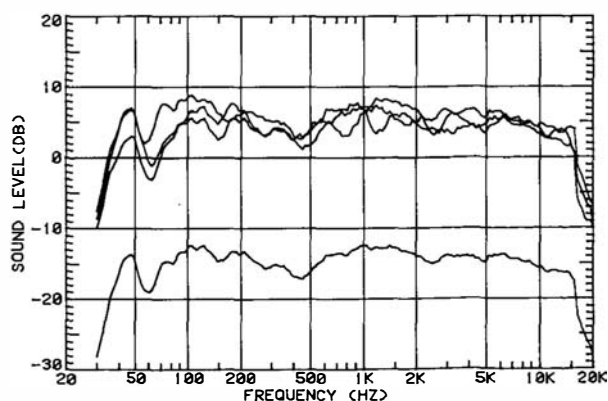


Fig. 17. As Fig. 16, but loudspeakers at position B (stereo left) measured at listener positions 1–6.

example, Fig. 18 illustrates a crude simulation of this argument. This loudspeaker is particularly interesting because it has a smooth and flat axial response, with progressive deterioration off axis (see Fig. 8).

The three superimposed curves in Fig. 18 have been adjusted to show the approximate relative strengths, in the listening area, of the direct sound (solid curve), first reflections (dashed curve), and reverberant sound (dotted curve). The direct sound is simply the on-axis anechoic response adjusted in level for the listening distance. The first reflections are the energy summation of sounds reflected from the floor, ceiling, and sidewalls, derived individually from the appropriate off-axis free-field measurements corrected, using the inverse-square law, for the distance traveled, assuming perfect reflection at the boundaries. The principal contributions in this example are from the floor and left sidewall reflections. The reverberant sound is derived from the total sound-power calculation, modified by the frequency-dependent sound absorption in the room [70]. The rising reverberation time at low frequencies results in increased levels, and the abundant high-frequency absorption in the room and in the air itself results in a severe attenuation at high frequencies.

This display illustrates clearly that for this loudspeaker, in this room, at this listening distance the low-frequency room measurement is mainly of reverberant energy. Through the middle frequencies there is a competition between the direct and reverberant sounds and a small number of energetic first reflections. At very high frequencies the direct sound prevails.

Performing an energy summation of the three components yields the top curve in Figure 18, which if the argument is reasonable, should bear a resemblance to real measurements of this loudspeaker in the room. It should be emphasized, however, that the simple synthesis did not take into account any of the acoustical wave effects, boundary interaction, and room resonances that are known to be significant. This estimate therefore should at best be an indicator of the measurement trend, around which variations will occur

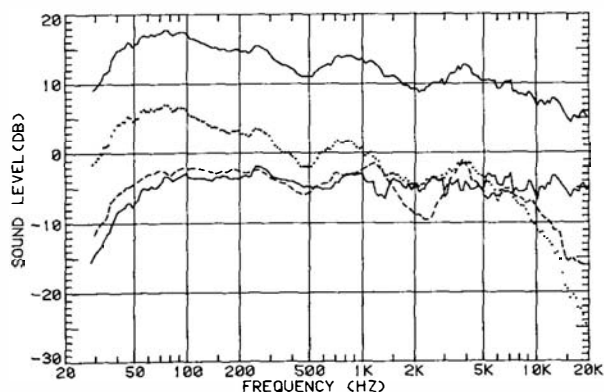


Fig. 18. Lower curves—estimates of direct (solid), early reflected (dashed), and reverberant (dotted) sounds at listener location 2 for loudspeaker 7 at position B (stereo left). Top curve—energy summation of three lower curves, shifted vertically for clarity. No spectral smoothing.

depending on source and microphone locations and the attendant boundary and standing-wave effects.

Fig. 19 shows a superimposition of 18 one-fourth-octave smoothed stepped-tone room measurements (source at A, B, and C, microphone at positions 1–6), showing the range of fluctuations associated with these specific locations. (One-fourth-octave smoothing was selected as a reasonable compromise between smoothing the interference effects and revealing the underlying details of room response.) Plotted above this collection for comparison is the simple synthesis from Fig. 18 which bears a good resemblance to the underlying trend of the measurements, notably at frequencies above about 500 Hz.

Combining the six measurements at each of the loudspeaker locations improves the comparison and shows more clearly the effects of loudspeaker position (Fig. 20). The spatial and one-fourth-octave spectral averaging have removed much of the interference clutter at high frequencies, showing the excellent fit of the predicted and measured curves. Below 500 Hz are the clear acoustical interference effects of the adjacent-boundary reflections. The calculated (dotted) curve thus

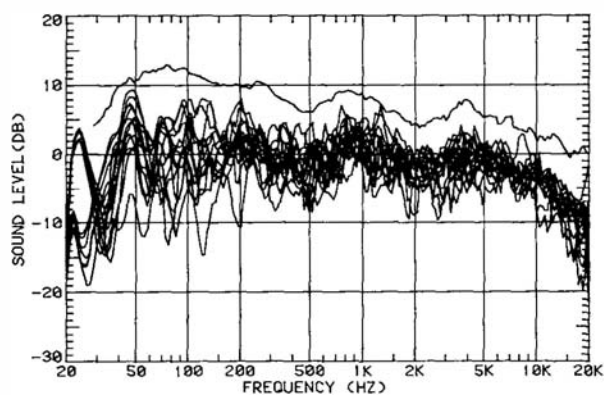


Fig. 19. Superimposition of one-fourth-octave smoothed room amplitude-response measurements of loudspeaker 7 located in turn at positions A, B, and C measured at listener locations 1–6. Curve plotted above this collection is the synthesis from Fig. 18.

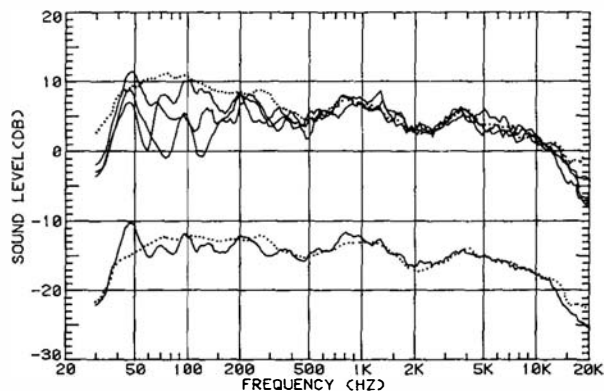


Fig. 20. Top—solid curves are the mean room amplitude responses of loudspeaker 7 at locations A, B, and C, averaged over listener positions 1–6. For comparison, dotted curve is the synthesis from Fig. 18. Bottom—solid curve is the mean of the three top curves; dotted curve is the mean amplitude response in front hemisphere.

is a good predictor of the listening-room measurements at frequencies above about 500 Hz, and below this frequency it defines an upper envelope beneath which room and boundary effects cause their inevitable attenuations at frequencies depending on the specific room configuration.

Combining all 18 measurements of Fig. 19 (or the top three in Fig. 20) yields the bottom curve in Fig. 20, the average of the averages. This artificial device describes a "generic" room response curve for this loudspeaker in a range of stereo-listening situations in this room. It would appear to describe a mean performance curve, above and below which the individual measurements fluctuate, with increasing tolerances at low frequencies.

The dotted curve shown weaving its way through this mean performance curve is another prediction from free-field data, but this time a rather more intuitively derived one. It is simply the mean amplitude response in the front hemisphere calculated as the energy average of the 25 measurements made at 15° intervals between ±90° horizontal and vertical (see Sec. 5.2). The spatial distribution of these measurements over the surface of the front hemisphere causes this curve to be the equivalent of an axially weighted hemispheric sound-power measure. (A true sound-power calculation involves weighting the individual measurements according to the solid angle represented by each one.) Thus the high frequencies will tend to be elevated slightly, compared to the true sound-power measurement, and because the exclusively (for these loudspeakers) low-frequency rear radiation is not included, the low-frequency portion of the curve will be lowered. The result, fortuitously, is a simple but effective estimate of the mean room response that seems to work well for loudspeakers in general, as long as they are of the conventional forward-facing variety.

Although the author came to this conclusion independently, 30 years ago Gee and Shorter [63], in describing their system for sound-power measurement, speculated that "it might be possible, by weighting the results obtained in different zones, to give more prominence to the front response, and so arrive at an empirical figure, intermediate between the axial and mean spherical responses, which would give a useful approximation to the overall effect obtained in the average listening room."

Fig. 21 shows the superimposed mean front-hemisphere amplitude response for the same three loudspeakers used for the room measurements shown in Fig. 15, and below these curves, the average. For comparison the average of the real room measurements shown in Fig. 15 is shown as a dotted curve. Above about 500 Hz the agreement is very good. Below this frequency some effects of the boundary interferences can still be seen, and the difference is similar to that observed for loudspeaker 7 in Fig. 20 (lower curve).

Comparing Fig. 20 with Figs. 15 and 21 one can see some evidence perhaps of why listeners rated loudspeaker 7 slightly below the other three products.

Even though actual room measurements were not available for all of the loudspeakers discussed here, it now is possible to simulate such curves from the anechoic data. Fig. 22 shows the mean amplitude response in the front hemisphere for the 20 highest rated loudspeakers in this group. Even though the room and the listening arrangement determine the form of actual room measurements below about 500 Hz, it is reasonable to view these data in a comparative sense, since the listening tests utilized a rigidly controlled standard physical arrangement. These data again confirm the importance of fine detail in the visual rating of loudspeaker performance; the lower rated loudspeakers tend to exhibit more evidence of resonances, and some of them have distinctly uneven responses when viewed in a more general sense. The highly rated loudspeakers appear, in general, to have the widest bandwidth and smoothest amplitude responses, both attributes that have been recognized in earlier analyses.

The overall forms of the curves, though, are not flat, even disregarding the undulations in the middle frequencies. In fact, straight-line approximations to the frequency responses of the eight loudspeakers with the highest ratings had downward tilts of 0.1 dB per octave (loudspeaker 6) to 0.6 dB per octave (loudspeakers 5 and 7). The calculated mean response for three loudspeakers shown at the bottom of Fig. 21 has a slope of -0.25 dB per octave. These variations are rather large compared to the detectable difference of 0.1–0.2 dB per octave reported by Kommamura and Mori [71]. The suggestion is that this may not be a particularly sensitive measurement.

7.3 A Caution about Loudspeaker Placement

In the monophonic listening comparisons loudspeakers were placed side by side at positions 3, 4, 5, and 6, as shown in Fig. 14. Between large loudspeakers there would, at times, be little space, although care was taken to keep them as far apart as possible within the designated zone. The question is, however, what

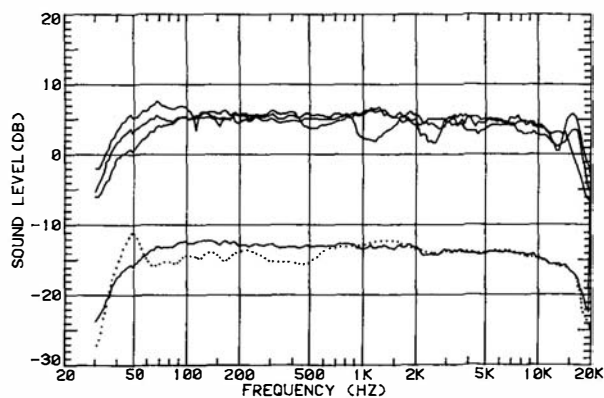


Fig. 21. Top—mean amplitude responses in front hemisphere for three loudspeakers with fidelity ratings between 7.5 and 7.9. Bottom—solid curve is the mean of the three top curves, derived from free-field data; dotted curve is the mean of room measurements on the same loudspeakers, from Fig. 15.

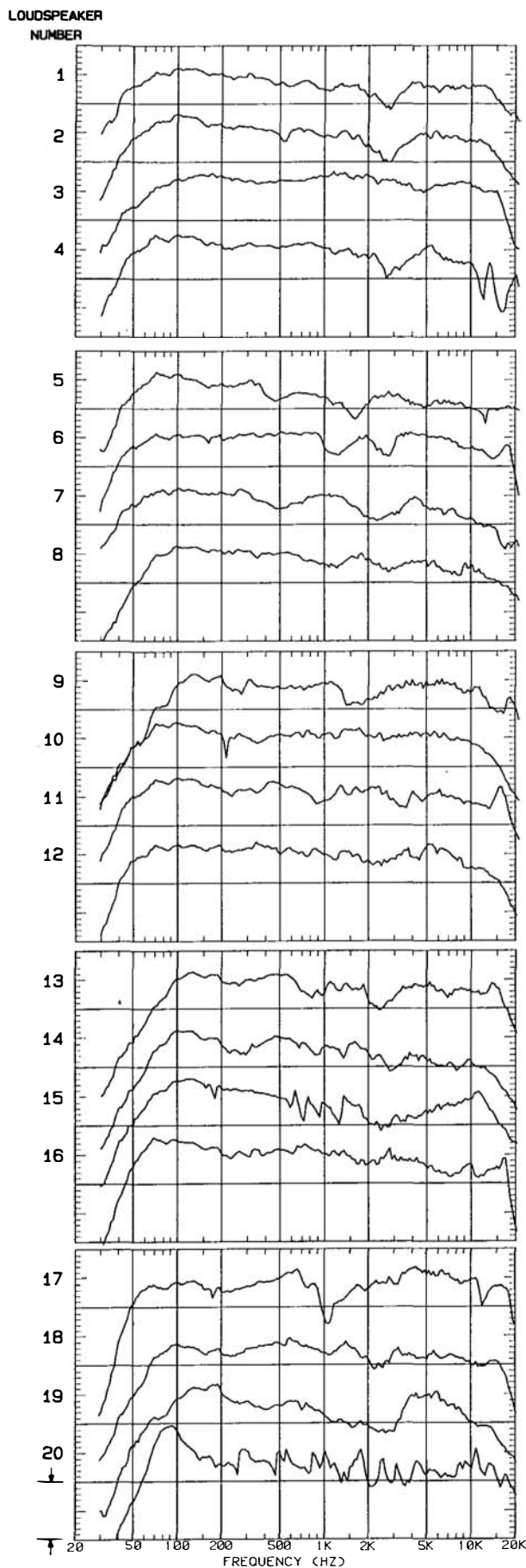


Fig. 22. Mean amplitude response in front hemisphere calculated from free-field measurements on 20 loudspeakers with fidelity ratings between 6.8 and 7.9.

spacing is sufficient to avoid seriously detrimental effects?

Fig. 23 shows measurements made at the three monophonic listening locations (1, 3, and 5). The test loudspeaker at position 5 was first measured alone, then with three others, all at spacings of 150 mm, and finally with all four loudspeakers almost touching. The result clearly shows the baffling effect of the closely adjacent loudspeakers, creating a half-space (2π) condition at frequencies above about 70 Hz. The lower difference curve shows a significant elevation in the sound level over about a two-octave frequency range. Increasing the spacing to 150 mm creates enough leakage to eliminate most of the baffle effect (top curve). Note also the downward shift in the rear-wall interference notch (60–70 Hz) as the baffle increased the distance from the woofer to the rear wall.

This result is consistent with the observations of Moir and Hands who tested the effect using free-field measurements [72].

7.4 Discussion

Without recourse to unwieldy techniques it is possible to show that steady-state measurements of amplitude response in the listening room are traceable to performance measurements made in the free field. More importantly, with a known room and listening configuration, the room response is, to a point, predictable from specific combinations of free-field data. The subject warrants a more penetrating analysis, but even this empirical exercise has, it is hoped, illustrated that the performance of a loudspeaker in a listening room is not, as some have argued, a matter totally divorced from measurements made in the free field.

It is evident that the use of sound-power measurements as substitutes for measurements in real listening rooms can be misleading. The example in Fig. 18 shows that

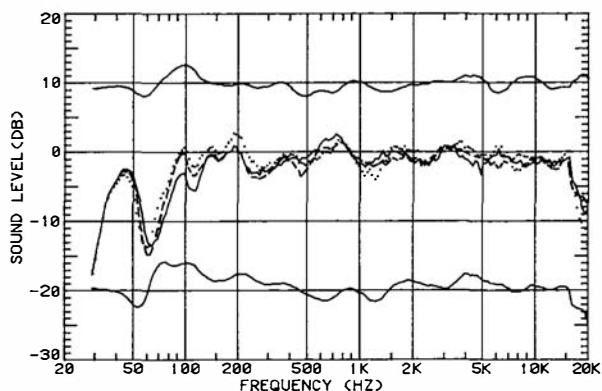


Fig. 23. Comparison of measurements averaged over three monophonic listening locations (1, 3, and 5) for one loudspeaker alone in position 5 (solid), with other loudspeakers in positions 3, 4, and 6, all separated from each other by 150 mm (dashed) and with no separation (dotted). Bottom curve is the difference between the zero separation condition and the loudspeaker alone (displaced -20 dB for clarity). Top curve is the difference between the 150-mm-spaced condition and the loudspeaker alone (displaced $+10$ dB for clarity). All curves are one-third-octave smoothed.

at moderate listening distances the reverberant sound field is truly dominant only at low frequencies. At these frequencies the response in an actual room will be modified by reflections from adjacent boundaries and by standing waves. Depending on the listening arrangement and the off-axis performance of the loudspeakers, first reflections from nearby room boundaries or surfaces can be a factor through the middle frequencies as well. The direct sound is a third ingredient in the midrange, and in some common listening situations, it can be the dominant factor at high frequencies.

It is therefore dangerous to attempt to generalize the meaning of steady-state measurements in listening rooms. They are not without meaning, but used in isolation, they can yield data which are very difficult to interpret correctly. With time-domain analysis, using impulse, correlation, or time-delay spectrometry methods, some resolution of the confusion is possible, but such methods are not straightforward, especially in the complex environment of a typical room.

Nevertheless room measurements do appear to contain some of the information related to the performance rankings, although without the precision of specific combinations of free-field data. Staffeldt and Rasmussen [73], [74] implied that some of this uncertainty could stem from the use of essentially omnidirectional microphones rather than an anatomically correct manikin for the room measurements. Measurements at the entrance to the ear canal do indeed incorporate the directional encoding of the external ear, head, and torso. These effects predominate at frequencies above about 1 kHz, however, and the uncertainties in room measurements appear to extend over the entire frequency range. Obviously, though, it is a matter worthy of consideration.

8 STEREOPHONIC IMAGING AND LOUDSPEAKER MEASUREMENTS

So far the discussions have concentrated on sound quality as assessed in monophonic listening tests. As pointed out earlier, listening in stereo has not been found to change the fidelity ratings of the high-ranking loudspeakers; however, there is the matter of spatial reproduction that must be addressed.

In [1] a good correlation was shown between the fidelity ratings and an overall spatial quality rating. This suggests that much or most of what is perceived to be a good reproduction of auditory images and spaciousness is due simply to accurate sound reproduction. A good monophonic loudspeaker used in pairs is, it seems, a good stereophonic loudspeaker.

Naturally there was the temptation to venture into this poorly charted territory to see whether, when examining the detailed structure of stereo imaging, there were some subtle effects that were being overlooked in the general assessments. There are high risks in looking at the results of a restricted test as we are about to do. The dangers derive from the fact that with loudspeakers of different design, there are differences in many parameters, all of which might influence the result.

The presentations and discussions are therefore restricted to only those aspects that the present evidence and past investigations suggest are relevant. Most of the arguments await a definitive test.

8.1 STEREO/MONO SERIES II

Readers are encouraged to examine [1, Sec. 3] for discussions and results that are not included here.

The three loudspeakers used in these interleaved stereo and mono comparison tests were selected because they exemplified three variations in contemporary domestic loudspeaker design. Loudspeaker AA (number 1 in the present tests) was a conventional two-way design, loudspeaker E (number 7 in the present tests) was a three-way phase-corrected design in a complex low-diffraction enclosure, and loudspeaker BB was a full-range electrostatic unit. In measured performance (Fig. 24) loudspeakers AA and E exhibited very good small-angle dispersion, showing little difference between the axial and the 0–15° listening-window curves. BB, because of a large tweeter, was rather more directional at high frequencies. In the midfrequencies AA showed the directionality of the woofer in measurements further off axis, E demonstrated the advantage of a midrange unit in a low-diffraction enclosure, and BB manifested the broadband directionality of a dipole radiator. (Rear radiation was suppressed above about 500 Hz by felt pads in the loudspeaker and by the drapes on the rear wall.)

In the assessments of overall spatial quality the greatest differences were apparent in the monophonic tests [1, Fig. 20], where BB was criticized for presenting a strong point-source illusion. It seems probable that this is due to its narrow dispersion, an observation made by Shorter [8] in his mono tests. This problem was not so serious in stereo listening, except with multimicrophone recordings that presented some essentially monophonic sounds panned to the left and right channels [1, Fig. 22], in which instances the criticisms were again of a small sound source. On occasion BB was also recognized as presenting a narrower sound stage, more abnormal effects, and a less satisfactory rendering of depth and ambience.

In the description of listening perspective for the music (choral, chamber, and jazz) recorded with a natural perspective, the modal listener response was "you are there" for AA and E and "close, but still looking on" for BB. According to the definitions of those phrases [1, App.], loudspeakers AA and E gave listeners some impression of being enveloped by the ambient sound of the recording environment, with BB tending to separate them from the performance.

All of these comments seem to be related to the relative strengths of the direct and reflected sounds wherein the loudspeakers with wide frontal dispersion, AA and E, presented listeners with images covering larger sound stages with more spatial involvement. This appears to have been achieved without any sacrifice in image definition or in the continuity of the sound stage [1, Fig. 22]. Listeners apparently judged the image sizes to be

appropriate to the musical experiences.

The effect of loudspeaker directivity on the size of auditory images has been noted before. Harwood [4], for example, pointed out that with speech or solo instruments highly directional loudspeakers are necessary to avoid diffuse images. This emphasis on compact images rather than a sense of spaciousness is appropriate, perhaps, in the context of a program engineer

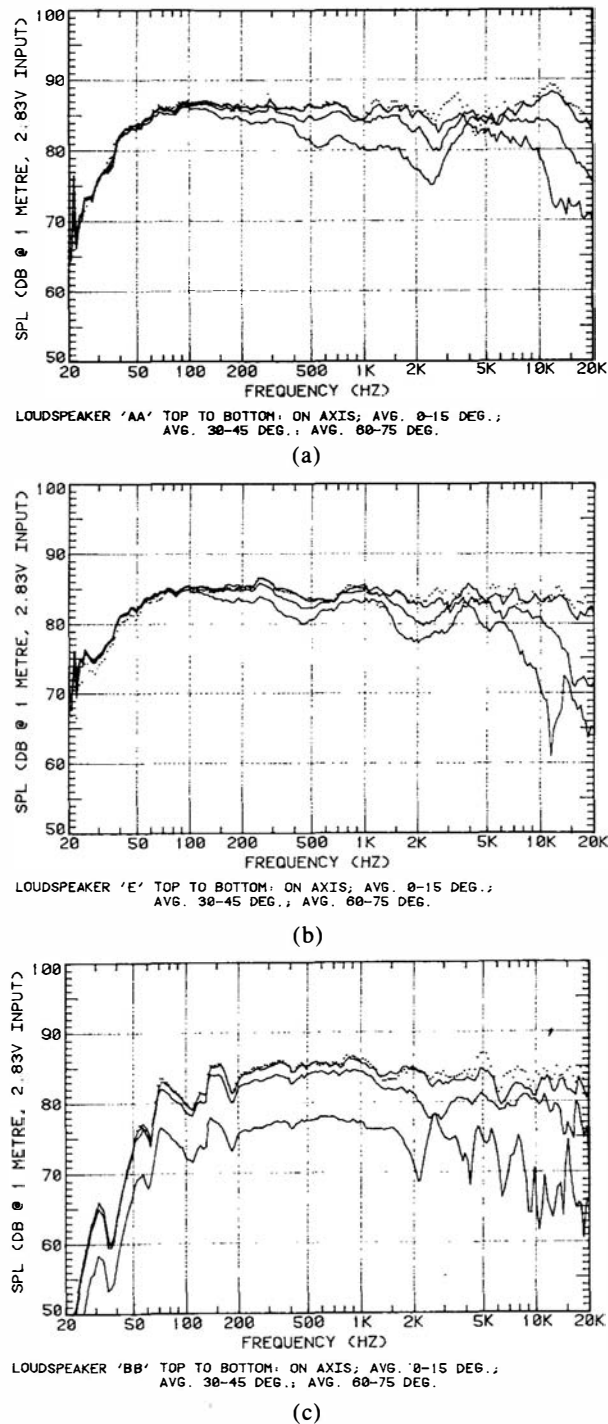


Fig. 24. Free-field measurements on three loudspeakers used in stereo/mono series II. Low-frequency dipole radiation characteristics of loudspeaker BB did not permit correction of anechoic-chamber defects below 200 Hz. These portions of the curves are shown because of the directional information therein.

in a broadcasting organization, a point well made by Kuhl and Plantz [75]. These investigators looked into the directional properties of loudspeakers that would be most suitable for control-room monitoring. Using only professional sound engineers as listeners, they found that narrow-dispersion loudspeakers were required for good reproduction of voices in radio dramas; dance and popular music was also desirably “aggressive” with “highly directed” loudspeakers. The majority of these same listeners, however, preferred wide-dispersion loudspeakers for the reproduction of symphonic music at home. In the control room, though, only about half of them felt that they could produce recordings with such loudspeakers.

Thus we come to the crux of a different problem, the matter of listener expectations and the motivations for listening. There is a trade-off, it seems, between the loudspeaker directivity required to preserve the illusion of truly compact sound sources in specifically localizable stereo images and that required to give the listener the impression of being immersed in another acoustic space. Bose [18], [19] consistently emphasized the importance of a strong reverberant sound field in the listening room, stating that “this spatial property of the sound incident upon a listener is a parameter ranking in importance with the frequency spectrum of the incident energy for the subjective appreciation of music” [18].

Queen [14] also concluded that widely dispersing loudspeakers were desirable so long as the uniformity of the spectral energy was preserved. The dilemma is summed up nicely by Brittain who recognized, in 1939, that “in order to reproduce reverberation naturally it is necessary for the loudspeaker to envelop the listener completely. If these loudspeakers are used to reproduce the direct sound as well as the reverberation the effect will be very unnatural, particularly when the source of sound is known to be small, as in speech” [2].

In the present experiments, however, all of the loudspeakers radiated the bulk of their energy toward the listener. AA and E were distinctive in that they radiated considerable energy at substantial angles off axis, but still in the front hemisphere. This form of directional sound output will energize the reverberant sound field but it will also, especially in the listening room used here, produce very energetic early reflections from the room boundaries. It is believed that these are responsible for the enhanced sense of space, while maintaining adequate stereo image quality.

Without going into details at this stage, it is interesting to note that in the concert hall there are precisely the same trade-offs where, as Rasch and Plomp put it, “in practice a compromise between requirements for definition and spaciousness is always necessary” [76]. Early reflections, especially lateral ones, have long been identified as providing pleasant spatial sensations [77]. There seem to be two separable aspects as explained by Bilson, 1) a broadening of the sound source that increases as a function of decreasing interaural correlation and increasing sound level, with low fre-

quencies being particularly influential; and 2) the sensation of being enveloped (involved) in the sound [78], [79].

With so much dilution of the direct sound by these early reflections it might seem that localization precision would suffer. Hartman found that indeed the early lateral reflections do tend to "delocalize" the source, but the early floor and ceiling reflections tend to reinforce the sense of localization. On balance, listeners in real environments listening to real sources of sound tended to regard the sense of surround as being more important than the localizability of the sound source [80].

Clearly there are some interesting parallels between listener preferences in the live performance and in listening to a stereo reproduction. There are also some differences, and depending on one's reasons for listening, loudspeaker directivity is a factor worthy of consideration. For domestic, recreational listening there seems to be some merit in loudspeakers with good dispersion in the frontal hemisphere, at least. The strong early reflections generated thereby do not appear to be unduly detrimental to the quality of individually localizable stereo images, while adding an apparently pleasant sense of proximity to the original recording environment. Again, though, the definitive tests remain to be done.

9 SUBJECTIVE MEASUREMENTS—A TEST FOR BIAS

In Sec. 5 a question was posed as to whether the results of the subjective measurements that formed the basis for this study contained any biases that could be identified in the objective measurements. The answer would appear to be a qualified "no."

From the point of view of the sound-quality assessments the evidence suggests that listeners simply preferred the loudspeakers with the fewest technical defects of those that were measured and are known to be audible. In this sense the subjective measurements appear to be neutral.

From the point of view of spatial quality, there seemed to be evidence of a factor relating the early reflected sound to sensations of spaciousness. This effect is likely to be dependent on the arrangement of loudspeakers and listeners in the specific room and on the disposition of sound-absorbing materials on the reflecting surfaces. It remains to be proved whether these particular listener responses can be generalized or whether they apply only to comparable listening conditions.

Further deterrents to complete generalization of these results exist in that the listeners were selected for listening experience and normal hearing and the loudspeaker systems were of the conventional multidriver forward-facing variety. Listeners with reduced hearing sensitivity tend to be less consistent in their opinions and to exhibit a preference for loudspeakers that may or may not be shared by listeners with normal hearing [1]. While it might be possible to design specific loudspeakers for specific hearing characteristics—a kind of prosthetic loudspeaker—it seems more realistic and

certainly more profitable to design for a high fidelity rating by normal listeners. In this way the product is likely to appeal to the majority of the purchasing population. Listeners with diminished hearing will simply be forced to select by listening only, as they do now.

The restriction of these results to conventional loudspeaker configurations is unfortunate and, possibly, temporary. It does, however, embrace the vast majority of products.

10 SUMMARY OF PRESENT WORK

At the outset of this protracted investigation there was some doubt that it would be possible, with any precision, to describe a good loudspeaker in purely technical terms. The problem was not a lack of competent measurements; it was a lack of agreement among listeners.

After the system for precise subjective measurements was developed, reliable listener responses provided the means by which measurement data could be classified. With this insight, consistent patterns emerged, indicating the principal parameters, and some less consistent relationships were apparent, identifying parameters of lesser importance.

Over many years of audio endeavor workers have expressed almost every point of view. Unfortunately those persuasive but untrustworthy personal opinions have caused some good ideas to be disregarded, only to be rediscovered later on, and in the present paper, once more. Although they did not and, under the circumstances, possibly could not develop the necessary scientific proofs, some of the very early workers appear to have had many of the right ideas. In truth, very little of real significance has been added in the past 25 years, but new equipment and techniques have enabled us to be rather more certain about the merits of established knowledge.

In many ways the outcome is a relief, in that it is a kind of reaffirmation of "motherhood." Given the proper circumstances, experienced listeners with normal hearing prefer loudspeakers with wide bandwidth, flat and smooth amplitude response, and uniformly wide dispersion. To arrive at this apparently simple conclusion one must define precisely what is meant by "amplitude response," and it is not what is commonly and currently popular.

Amplitude-response measurements should be made in the free field, at a distance that places the microphone in the far field (normally 2 m or more for domestic loudspeakers), with sufficient resolution in the frequency domain to reveal minor resonance effects (that is, probably 1/10 to 1/20 octave over most of the audible frequency ranges; the resolution at the frequency extremes may possibly be reduced). Measurements should be made at a sufficient number of orientations, including the reference axis, that it is possible to perform spatial averaging to remove the visual clutter of acoustical interference effects while retaining some directional information, at least in the frontal hemisphere.

Interpreting the data, after that much processing, is

relatively straightforward, involving looking for a flat wide-bandwidth on-axis response, and for consistently repeated patterns in the family of progressively off-axis measurements. The importance of the deviations from the underlying smooth contours must be weighted according to a set of rules that take account of the direction, shape, and magnitude of the irregularities, and the frequencies at which they occur. The result, with no further data, is a ranking of loudspeaker performance that has about the same resolution and order as listening tests conducted with the utmost of care, and with considerably more time and expense.

This is not to say that amplitude response is the only important measurement. It is, however, indicating that through this one dimension it is possible to indirectly observe behavior in another. To such an extent is this true that unless the amplitude response is of the proper form, there would appear to be little value in pursuing measurements in other dimensions.

On the other hand, if it is not possible to acquire amplitude-response information of the ideal kind, then additional information about phase and time-domain performances must be sought, and the appropriate rules for interpretation applied.

Given both the necessary amplitude response information and the additional data on the important linear and nonlinear distortions, a designer or reviewer should be in a position to recognize performance of the highest caliber with very little uncertainty. At more commonplace levels of performance, comprehensive data may help designers to achieve the compromises necessary to maximize listener ratings at a given price. Judging from many of the products evaluated in these tests, it is evident that some designers have been working with inadequate technical data of the most fundamental kind, and others have concentrated on perfecting the wrong parameters. For example, the need to concentrate on more than just the axial performance has implications for crossover design that have been largely avoided in the numerous academic exercises arguing the relative merits of networks [81], [82].

Nowadays free-field data can be acquired in normal semireverberant spaces by using time-delay spectrometry or impulse/fast Fourier transform analysis. Both techniques employ linear frequency bases which, when converted to the familiar logarithmic frequency scales, tend to reveal too much information at high frequencies and too little at low frequencies. The loss of resolution at medium and low frequencies and the usual total lack of information below a few hundred hertz are serious limitations of the methods, although there are techniques by which some restrictions can be relieved [83]. The linear frequency base is arguably advantageous for revealing certain kinds of behavior, such as comb filtering caused by interference between time-delayed signals. The experimental evidence suggests, however, that in some circumstances the eyes are more offended by these data than are the ears, with specific forms of this "defect" actually being desirable in stereophonic reproduction [41, Sec. 8.1].

Repeatedly it is necessary to modify visual interpretations of technical data with knowledge of how the physical phenomena will be perceived by two ears and a brain. In such cases, data processing rather than data acquisition can be of assistance. There is much more to be learned, but with the passage of time, orderly and useful patterns are emerging.

11 COMPARISON WITH EARLIER WORK

In the Part 1 review of previous published work on loudspeaker measurements it was clear that there were divergent opinions about certain loudspeaker measurements. The most prominent of these was in the measurement of amplitude response, where various workers advocated a flat on-axis free-field response as the design ideal, while others promoted this in conjunction with a consideration for off-axis and sound-power performance, in some kind of weighted combination, while still others put the emphasis on a flat sound-power or corrected listening-room measurement.

In the present results all versions of amplitude-response measurement were able to identify the loudspeakers with the lowest fidelity ratings, where the overall spectral output was very inconsistent. Higher rated loudspeakers presented the real challenges, especially when the fine details rather than the overall shape of the response curve seemed to convey the important information. Other factors being equal, it seems that the smoothness of a high-resolution measurement is an essential performance attribute.

Virtually all previous measurements of sound power and of amplitude response in the listening room utilized one-third-octave or wider bandwidth measurements. Such measurements cannot reveal the fine details of amplitude response that the present tests indicate are important to the identification of the highest rated loudspeakers. The addition of phase or time-domain performance data can be used to identify the presence of potentially audible colorations due to resonances. However, it would seem that the amplitude-response information is a relatively simple and clear indicator of listener reactions to the colorations. In this the present work supports the observations of Büchlein [45], Stevens [46], and Fryer [47] who emphasized the importance of coloration in loudspeaker evaluations.

Listening-room and sound-power measurements present the additional problem that a decision is required about the form of the "ideal" curve. This might be feasible for loudspeakers of similar design and directional properties, used in similar rooms. However, the methods would appear to be subject to error when applied to loudspeakers and rooms in general. None of the "ideal" curves shown in Fig. 1 is a good match for the room measurements shown in Figs. 15–17, which describe the performance of the same three top-rated products. Furthermore, the common steady-state measurement methods are not "analytical," denying access to directional and time-domain information about the loudspeaker. Perhaps the most serious weakness of the methods is that they can be fooled. To use an outrageous

example: a woofer under the sofa, a midrange in the ashtray, and a tweeter in the chandelier could be equalized to yield good sound-power or listening-room responses. The sound quality, however, would be abysmal. Consequently these methods cannot be used in isolation, but they can provide useful supporting data. The sound-power measurement has the advantage of being a calibrated measure, reproducible, within reasonable tolerances, in any laboratory. Listening-room measurements on the other hand are subject to the vagaries of a nonstandard environment, but they do convey some information about the combination of a specific loudspeaker and a specific room configuration.

From the sound-power data alone it was possible to recognize the worst loudspeakers, but there were ambiguities in ranking the better performers. This is essentially the nature of the caution issued by Brociner and von Recklinghausen [17], mentioned earlier. In summary, therefore, it would seem to be imprudent to consider either sound-power or steady-state listening-room measurements as the definitive measure of loudspeaker amplitude response.

At the other extreme, the anechoic on-axis response or any similar "direct-sound" criterion would appear to abandon important information about the nature of the sounds supplied to the early reflected and reverberant sound fields.

It should be noted, however, that reliance on the axial response alone would probably have identified a loudspeaker in one of the top two categories in Fig. 7, which is indeed useful. Even with good listening tests of the conventional kind, listeners might not be any more certain in their subjective rankings. On the other hand, the subjective measurements used in these tests [1] had enough resolution to make the distinction between those loudspeakers that were well designed in most respects that matter, and those where additional attention had been paid to aspects of performance that tend to be neglected.

Between these extreme views are those that acknowledge the importance of balanced performance, where several aspects of amplitude response must be examined, sometimes for different reasons. The final result is then the culmination of a careful trading off of performance in one aspect against another. The weight of the evidence in this study supports just such a view.

The on-axis response is very important, though not only for the obvious reason that it describes the first sound arrival at the listener's ears. In loudspeakers of domestic size, with low directivity indices, the axial response also conveys much of the basic pattern that is revealed in all other measurements of amplitude response. Observe, in Fig. 8, that the on-axis curves contain much of the character persisting through the spatial averaging of off-axis response and, ultimately, even to the sound-power measurement (Fig. 12), the result of combining 34 individual measurements. In progressively more subtle ways, the same effect can be seen in the higher rated loudspeakers. Note, however,

that this may not be as true for highly directional loudspeakers or loudspeaker arrays.

The only way to be certain about what is happening off axis is to measure it, not in the homogeneous sound field of a reverberation chamber, nor in the mysterious sound field of a listening room, but in the free field. It seems to be useful to observe the manner in which the loudspeaker reveals its directivity, as a function of frequency. It is of further assistance if these measurements incorporate a certain amount of spatial averaging, to remove the detailed irregularities caused by acoustical interference, and also if the measurements offer high resolution in the frequency domain so that abrupt changes in output can be seen. These local aberrations are helpful indicators of both the presence and the severity of resonances and the consequent audible colorations.

The directivity index is really just the difference between the shapes of the sound-power and on-axis responses. Spatial averaging smooths most of the sharp discontinuities in the sound-power curves so that steep fluctuations in the directivity index are usually the result of abrupt changes in the on-axis response. Since these can be caused by interference effects, it is possible for an erratic-looking directivity-index curve to contain false alarms. Replacing the on-axis curve with a spatially averaged listening window of, say, $\pm 15^\circ$ would produce directivity-index curves that are more easily interpreted while retaining high-frequency resolution. As a description of a specific aspect of loudspeaker performance, the directivity index has some limited use in, for example, public address applications. As a figure of merit it would appear to have little value.

In amplitude-response measurements in general, the evidence of the present work suggests that greater emphasis should be placed on spatial averaging rather than on spectral averaging. The latter yields smooth curves, but in doing so can discard important clues about loudspeaker performance.

The advocates of accurate waveform reproduction, implying both accurate phase and amplitude responses, are in a particularly awkward situation. In spite of the considerable engineering appeal of this concept, practical tests have yielded little evidence of listener sensitivity to this factor.

In the present work the limited results lend support to the popular view that the effects of phase are clearly subordinate to amplitude response. The two factors are related, however, in that, using the techniques advocated here, a smooth amplitude response will be achieved only if localized phase aberrations have been minimized. The gradual phase shifts appear to be relatively unimportant.

Nonlinear distortions were not a matter for specific study in these tests, although measurements of total harmonic distortion were made at two input power levels. The data were not shown, as they revealed no clear trends. In fact, with the exception of some evidence of excursion limiting at high sound levels by the small inexpensive loudspeakers, the distortion performances

were rather similar. Thus the conclusion is that this parameter was simply not a variable in these tests and remains, as mentioned earlier, a subject for further study.

12 CONCLUSION

There are two principal conclusions to this work. The first is that, given adequate preparation and experimental controls, listening tests can yield reliable subjective data and that listeners with good hearing performance agree closely on the relative merits of loudspeakers. The second is that the loudspeakers preferred by these listeners are those exhibiting measured performances that are superior in certain well-defined respects.

The original objective of the study appears to have been met in that it has been demonstrated that there are clearly parallel systems of subjective and objective measurements and it is possible to arrive at similar conclusions from measurements of either kind. Considering the relative costs of the methods, the technical assessments seem to have the long-term advantage in that, once the measurement facility is established, the measurements themselves can be performed quickly, as a matter of routine. Listening tests, on the other hand, require considerable labor and time if the results are to be trusted. The material and technical costs can be comparatively modest, however, involving the initial meticulous preparation of a listening room and ancillary sound playback apparatus.

In most situations a combination of both methods will be used, but it is hazardous to employ anything but the most precise of either technique. For example, not all technical measurement methods in common use are equally sensitive to performance imperfections. At present it would seem that almost any form of amplitude-response measurement can identify really poor loudspeakers. More specific measures and perhaps some processing are necessary in order to reliably identify good loudspeakers. But only very specific and possibly special measurements in a combination of domains can identify truly excellent products as reliably as thoroughly controlled subjective measurements. Conventional listening evaluations are essentially uncalibrated measurements; they are likely very often to be in error.

A review of published opinion on loudspeaker measurements indicated that not all workers are employing the most sensitive tests. Manufacturers are responsible only for their own products; however, product reviewers have a much wider responsibility. Theirs is a particularly difficult position. Purely subjective reviews are easily challenged unless extraordinary care has been taken in the preparation and conduct of the listening tests. The measurements seen in some publications appear to be a form of ornamentation for a review, implying but not confirming analytical ability. Others show fairly comprehensive data, but the interpretations have tended to be more descriptive than conclusive.

Obviously it is not a simple problem to evaluate loudspeakers for the benefit of a broad audience. How-

ever, in principle it is the same problem that is faced by the designers. It is hoped that the present work will be a stimulus to both groups.

13 ACKNOWLEDGMENT

The author wishes to thank René St. Denis for his assistance in creating hardware and software, and in collecting and organizing the masses of data used in this study. Special thanks are due to Bob Donato who has been most generous in sharing his knowledge of interfacing idiosyncrasies and data processing.

14 REFERENCES—PART 2

- [1] F. E. Toole, "Subjective Measurements of Loudspeaker Sound Quality and Listener Performance," *J. Audio Eng. Soc.*, vol. 33, pp. 2–32 (1985 Jan./Feb.).
- [2] F. H. Brittain, "The Appraisal of Loudspeakers," *GEC J.*, pt. 1, vol. 7, pp. 266–276 (1936 Nov.); pt. 2, vol. 8, pp. 1212–1230 (1937 May).
- [3] J.-C. Hentsch, "La Fidélité des Haut-Parleurs dans la Reproduction des Phénomènes Transitoires," *Tech. Mitt.-PTT* (Bern), no. 6 (1951 June).
- [4] H. D. Harwood, "The Hows, Wheres and Whys of Testing High Quality Loudspeakers," *Audio*, vol. 55, pt. 1, pp. 16–20 (1971 Aug.); pt. 2, pp. 18–21 (1971 Sept.).
- [8] D. E. L. Shorter, "A Survey of Performance Criteria and Design Considerations for High-Quality Monitoring Loudspeakers," *Proc. IEE*, vol. 105, pt. B, pp. 607–621 (1958 Nov.); reprinted in *J. Audio Eng. Soc.*, vol. 7, pp. 13–28 (1959 Jan.).
- [14] D. Queen, "The Effect of Loudspeaker Radiation Patterns on Stereo Imaging and Clarity," *J. Audio Eng. Soc.*, vol. 27, pp. 368–379 (1979 May).
- [17] V. Brociner and D. R. von Recklinghausen, "Speaker System Design Using a Reverberation Chamber," presented at the 34th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 16, p. 342 (1968 Jul.), preprint 579.
- [18] A. G. Bose, "On the Design, Measurement and Evaluation of Loudspeakers," presented at the 35th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 17, p. 86 (1969 Jan.), preprint 622.
- [19] A. G. Bose, "Sounding Recording and Reproduction, Part 1: Devices, Measurements, and Perception," *Technol. Rev. (MIT)*, pp. 19–25 (1973 June); "Part 2: Spatial and Temporal Dimensions," *ibid.*, pp. 25–33 (1973 July/Aug.).
- [41] D. Clark, "Measuring Audible Effects of Time Delays in Listening Rooms," presented at the 74th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 31, p. 972 (1983 Dec.), preprint 2012.
- [45] R. Bücklein, "The Audibility of Frequency Response Irregularities," *J. Audio Eng. Soc.*, vol. 29, pp. 126–131 (1981 Mar.), a translation of a 1962 original paper in German.
- [46] W. R. Stevens, "Loudspeakers—Cabinet Effects," *Hi-Fi News Rec. Rev.*, vol. 21, pp. 87–93, 97 (1976 Sept.).
- [47] P. Fryer, "Loudspeaker Distortions—Can We Hear Them?," *Hi-Fi News Rec. Rev.*, vol. 22, pp. 51–

56 (1977 July).

[56] J. Blauert, *Spatial Hearing* (1974); transl. by J. S. Allen (M.I.T. Press, Cambridge, MA, 1983).

[59] "Sound System Equipment, Part 13, Listening Tests on Loudspeakers," International Electrotechnical Commission, Publ. 268-13 (in press).

[60] F. E. Toole, "Selecting Loudspeakers, Part 2, Interpreting Specifications and Measurements," *Electron*, vol. 8, pp. 18-21, 40 (1971 July).

[61] F. E. Toole, "Speakers, Part 2," *AudioScene Canada*, vol. 14, pp. 41-51 (1977 Apr.).

[62] F. E. Toole, "Loudspeaker Measurements: Establishing a New Order," *Sound Canada*, vol. 14, pp. 14-30 (1983 Nov.).

[63] A. Gee and D. E. L. Shorter, "An Automatic Integrator for Determining the Mean Spherical Response of Loudspeakers and Microphones," British Broadcasting Corp., Engineering Monograph 8 (1955 Aug.).

[64] "Sound System Equipment, Part 5, Loudspeakers," International Electrotechnical Commission, Publication 268-5 (1972).

[65] F. E. Toole, "Subjective Measurements of Loudspeaker Sound Quality," presented at the 72nd Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 30, pp. 942, 944 (1982 Dec.), preprint 1900.

[66] F. E. Toole, "Loudness, Applications and Implications to Audio," *dB—The Sound Eng. Mag.*, vol. 7, pp. 27-30 (1973 May); pp. 25-28 (1973 June).

[67] N. I. Durlach and H. S. Colburn, "Binaural Phenomena," chap. 10 in E. C. Carterette and M. P. Friedman, Eds., *Handbook of Perception*, vol. 4 (Academic Press, New York, 1978).

[68] R. F. Allison, "The Influence of Room Boundaries on Loudspeaker Power Output," *J. Audio Eng. Soc.*, vol. 22, pp. 314-320 (1974 June).

[69] F. E. Toole, "Listening Tests—Turning Opinion into Fact," *J. Audio Eng. Soc. (Engineering Reports)*, vol. 30, pp. 431-445 (1982 June).

[70] L. L. Beranek, *Acoustics* (McGraw-Hill, New York, 1954), chap. 10.

[71] M. Kommamura and S. Mori, "New Measurement of Frequency Response Flatness Limen," *J. Acoust. Soc. Jpn.*, vol. 39, p. 463 (1983).

[72] J. Moir and E. Hands, "The Unwanted Speaker Effect," *HiFi News Rec. Rev.*, vol. 27, pp. 40-41, 45

(1982 Oct.).

[73] H. Staffeldt and E. Rasmussen, "The Subjectively-Perceived Frequency Response in Small and Medium-Sized Rooms," *J. Soc. Mot. Pict. Telev. Eng.*, vol. 91, pp. 638-643 (1982 July).

[74] H. Staffeldt, "Measurement and Prediction of the Timbre of Sound Reproduction," *J. Audio Eng. Soc.*, vol. 32, pp. 410-414 (1984 June).

[75] W. Kuhl and R. Plantz, "The Significance of the Diffuse Sound Radiated from Loudspeakers for the Subjective Hearing Event," *Acustica*, vol. 40, pp. 182-190 (1978 July).

[76] R. A. Rasch and R. Plomp, "The Listener and the Acoustic Environment," chap. 5 in D. Deutsch, Ed., *The Psychology of Music* (Academic Press, New York, 1982).

[77] M. Barron, "The Subjective Effects of First Reflections in Concert Halls—The Need for Lateral Reflections," *J. Sound Vib.*, vol. 15, pp. 475-494 (1971).

[78] F. Bilsen, "Psychophysical Observation on Spaciousness," *Proc. 10th Int. Cong. Acoustics* (Sydney, 1980), paper E-7.2.

[79] F. Bilsen and F. Brinkman, "On the Perception of Spaciousness with Artificial Head," *Proc. 11th Int. Cong. Acoustics* (Paris, 1983), pp. 135-138.

[80] W. M. Hartman, "Localization of Sound in Rooms," *J. Acoust. Soc. Am.*, vol. 74, pp. 1380-1391 (1983 Nov.).

[81] J. Vanderkooy and S. P. Lipshitz, "Power Response of Loudspeakers with Non-Coincident Drivers: The Influence of Crossover Design," presented at the 74th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 31, pp. 968, 969 (1983 Dec.), preprint 2049.

[82] H. Suzuki, "Mutual Radiation Impedance of a Double-Disk Source and Its Effect on Radiated Power," presented at the 76th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 32, p. 1008 (1984 Dec.), preprint 2144.

[83] L. R. Fincham, "Refinements in the Impulse Testing of Loudspeakers," *J. Audio Eng. Soc.*, vol. 33, pp. 133-140 (1985 Mar.).

Dr. Toole's biography was published in the April issue.