Measuring Loudspeakers, Part One

John Atkinson, November 1998

This series of articles was initially written (in slightly different form), as a paper presented at the 103rd Audio Engineering Society Convention, New York, September 1997. The preprint, "Loudspeakers: What Measurements Can Tell Us—And What They Can't Tell Us!," AES Preprint 4608, is available from the AES, 60 East 42nd Street, Room 2520, New York, NY 10165-0075. The AES internet site, *offers a secure transaction page for credit-card orders.*

What's the point of measuring the performance of an audio component if what is owners do is listen to it? Of the 20 or so regularly published English-language consumer magazines on audio, only three routinely publish rigorous loudspeaker measurements: *Hi-Fi News & Record Review* in the UK; *Audio* and *Stereophile* in the US (footnote 1). It would seem that supporting a reviewer's subjective impression of how a loudspeaker sounds with measured data is an idea that is out of fashion these days. But without measuring a loudspeaker's performance, how can a "Golden Ears" reviewer avoid falling into error? I believe that it is the responsibility of any magazine that publishes judgments on audio components to fully describe their performance. And that cannot be done without measurements.

For electronic components, measuring their effect on an input signal is relatively straightforward (though correlating those effects with perceived sound quality is not a job for the fainthearted). But for many decades, getting a true idea of how a loudspeaker performed was only possible using swept sinewaves in large, expensive anechoic chambers. However, the relatively recent advent of low-cost, PC-based measuring equipment capable of assessing acoustic performance in normal rooms has made it possible for more people to measure more loudspeakers than ever before. However, simply placing a microphone in front of a loudspeaker and pressing *ENTER* on the computer keyboard will still not produce a measurement that is likely to predict how that loudspeaker will sound in a room.

If you examine the vocabulary typically used to describe loudspeaker performance by audiophiles and in magazine reviews [1, 2, 3, 4], it covers the following areas:

• Musical and technical accuracy—how good? How close is the loudspeaker to reproducing an overall sound that will convince its listeners they're in the presence of live sounds?

• Frequency range—bass and treble extension. Does the sound have its full complement at the frequency extremes? Or has, say, the designer discarded ultimate bass extension, perhaps to maximize midrange clarity or overall power handling?

• Frequency balance—how neutral? Does the speaker have a natural timbre? If you play a recording of someone's voice, does it actually sound like that person's voice, or does it sound more like a combination of bass boom and a bat squeak? Is the presentation tilted-up at one or both frequency extremes? Is there a shelf effect present at one or more frequencies?

• Response and balance anomalies—coloration. Do voices sound like someone is speaking through a megaphone or pinching their noses. Is there excessive sibilance on a female voice or too much chest tone on male voice?Can the sounds of similar but different musical instruments be distinguished? Or does the loudspeaker's own character make a violin sound like a viola, an oboe like an English Horn, a Fender Stratocaster like a Gibson Les Paul? With some chronic loudspeaker cabinet problems, it can almost sound as though someone is banging a xylophone along with the sound of the recorded instruments.

• Clarity and transparency. How much detail can you hear? I remember taking part in a blind test in the early '80s where we A/B'd two speakers using a live microphone feed of a bunch of keys being rattled in an anechoic chamber. One speaker made it sound like there was just one key. With another loudspeaker, you could almost hear how many keys there were. Do individual images within the soundstage sound sufficiently separate from each other, or do they seem more like spatial highlights emerging from a background of sonic soup?

• Grain, hardness, and distortion. Does the loudspeaker sound significantly more distorted at high playback levels? Does the loudspeaker make listeners want to run screaming from the room? Do the listeners feel as if their eardrums are being gently caressed with sandpaper?

• Stereo imaging precision. When a dual-mono recording is played, do listeners perceive a narrow sonic object precisely midway between the loudspeakers, or do they just hear a vague, amorphous blob? Do central images stay centrally located at some frequencies but not others?

• Soundstage width and depth. With appropriate recordings, can listeners hear acoustic objects precisely positioned anywhere on a two-dimensional grid defined by the loudspeaker and listener positions?

• Dynamics (micro and macro). Are loud musical passages appropriately louder than the quiet passages (macrodynamics)? Can listeners hear subtle changes in one acoustic object when something else is playing very loudly (microdynamics)? Does it all blur at high playback levels, or does it take being played loudly to make the music "come to life"?

• Pace'n'rhythm. First defined by *Stereophile* contributor and UK reviewer Martin Colloms [5]. Some loudspeakers clearly make the music sound like it's going slowly, while others make it sound like it's going faster, even though the recording's playing time and the music's tempo obviously cannot be affected by anything that a loudspeaker does.

Footnote 1: Julian Hirsch published descriptions of how his review speakers measured throughout his tenure at *Stereo Review*, but as nary a graph was to be seen in that magazine's pages, I didn't feel it appropriate to include it in this list. However, with the policy changes announced in its September 1998 issue, *Stereo Review* would appear to be taking a step forward.

Subjective loudspeaker performance is thus a multidimensional phenomenon. However, to make objective measurements that are both meaningful and practicable involves a subjective choice about what parameter to plot against one, or at most two, other parameters. All other parameters have then to be held constant. If you plot, say, a loudspeaker's sound-pressure level against frequency for a given input voltage, the result is the typical amplitude or "frequency" response. But this measured response will only be valid on the chosen axis in an anechoic chamber at the chosen sound-pressure level at one instant of time. How typical will it be of what the loudspeaker does with music in a real room played at widely differing spls? It is important, therefore, to keep in the back of your mind that to make "objective" measurements involves subjective choices!

A list of measurements that I typically perform in connection with the loudspeaker reviews published in *Stereophile* includes:

• Voltage sensitivity on the chosen axis.

• Electrical impedance (magnitude and phase).

• Impulse and step responses.

• Amplitude and phase response on the chosen axis in the farfield.

• Nearfield amplitude response (at low frequencies).

• Polar behavior—Dispersion in horizontal and vertical planes.

• Power and other in-room responses.

• Nonlinear distortions of various kinds.

• Delayed acoustic resonances.

• Cabinet vibrational behavior.

It should be obvious that not one of the parameters in this second list appears to bear any direct correlation with one of the subjective attributes in the first list. If, for example, an engineer needs to measure a loudspeaker's perceived "transparency," there isn't any single two- or three-dimensional graph that can be plotted to show "objective" performance parameters that correlate with the subjective attribute. *Everything* a loudspeaker does affects it to some degree or other.

Of course, there are some performance parameters that correlate significantly—perceived bass extension with measured low-frequency extension, for example—but it is important to remember that there will always be other aspects of measured performance that also contribute. Anyone who looks at published measurements should never assume that one measurement—a frequency response, or an impedance curve, or a dispersion pattern—fully or even partially describes the sound that they will hear. It's only the totality of all possible measurements looked at simultaneously that will give the reader any idea of what's going on. What you hear always depends on more than one measurement. Ergo, no one measurement can tell the whole story.

And given half a chance, all measurements will tell lies. It's very easy to assume that if you get a piece of test gear, turn it on, and hook it up to the device-under-test, that the resultant graph is meaningful. It's never safe to assume that a) the graph is correctly plotted, or b) that you are actually measuring what you think you're measuring. You still need another source of data, much as in pre-calculator days someone using a slide-rule needed to know approximately what the answer would be before they did a calculation. When you measure a loudspeaker's complex impedance, for example, it is helpful to look at the waveform of the signal present at the speaker terminals with an oscilloscope and to listen to the speaker's acoustic output. The test set might still produce a nice-looking graph, even if the speaker isn't making a sound!

In the following sections of this three-part article are discussions of how I prefer to perform standard measurements and how they should be interpreted. How each measured area of performance affects the areas of subjective performance is examined, with particular attention paid to measured characteristics that appear to correlate strongly with very good or very poor perceived sound quality.

Voltage Sensitivity

A loudspeaker's sensitivity appears to be universally confused with its efficiency. Efficiency is strictly defined [6, 7] as how much *acoustic* power the loudspeaker puts out for how much *electrical* power it is being driven with. If you feed a loudspeaker with 100 electrical watts, how many acoustic watts of sound does it produce? The answer is "not many," a typical moving-coil loudspeaker being about 1% efficient.

Efficiency is usually expressed in the form of a sound-pressure level produced by a speaker at a specific distance, 1m, for 1W input; *ie*, in dB/W/m. This is problematic, however, as there is no simple way of determining, for a given loudspeaker, what actually is a 1W input—it depends on both impedance and frequency. Feed a loudspeaker with an impedance of 8 ohms at 1kHz with a signal at the same frequency at a 2.83V level, and yes, you are feeding it 1W of electrical energy. (By Ohm's Law, Power = V *x* V/R = (2.83 *x* 2.83)/8 = 1W.) But if, as is very often the case, the loudspeaker has a much lower impedance at 200Hz—2 ohms, say—the loudspeaker fed the same 2.83V at 200Hz will now suck four times as much power from the amplifier. For the same sound pressure level, the speaker is four times as *efficient* at 1kHz as it is at 200Hz.

And why anyway does efficiency matter? While audio engineers a half-century ago were interested in the transfer of power (and telecommunications engineers still are), since the advent of solid-state devices, audio amplifiers act more-or-less as voltage-source devices—they maintain the same output voltage no matter what the load and the current drawn. What is important, therefore, is not efficiency but *voltage sensitivity* [8]: how loud a loudspeaker plays for a given voltage level from the amplifier. It is generally defined as the sound-pressure level produced by a loudspeaker at 1m by an input voltage of 2.83V (the voltage necessary to produce 1W dissipation in an 8 ohm resistor).

The advantage of specifying sensitivity rather than efficiency is that it remains unchanged no matter what the impedance of the loudspeaker, as it is assumed that the amplifier will always be able to provide the necessary current to maintain the 2.83V. The nearer a loudspeaker's modulus of impedance approaches that of a pure 8 ohm resistor, the closer the equivalence between the two criteria; but when a speaker has an impedance that differs significantly from 8 ohms, they can be very different, as in the case I mentioned above.

A classic example of this difference between the two terms is an electrostatic speaker I measured some years ago. It had an impedance in the bass of over 100 ohms. Its *sensitivity* was very low, around 79dB/2.83V/m. But consider its impedance: by comparison with a typical 8 ohm dynamic speaker, the 'stat is drawing almost no current. It is therefore *very* efficient at transforming electrical power into acoustic power.

A loudspeaker's voltage sensitivity is particularly important when matching amplifiers and loudspeakers. If you have a 20W amplifier, you had better use a very sensitive loudspeaker or else your music will not play very loud. Conversely, if you choose a loudspeaker with a sensitivity which is high—say, 100dB/2.83V/m—you can probably get away with a 5W amplifier, meaning that spending $10,000 on a 600W design would be a waste of money. However, even if it is agreed that voltage sensitivity is the appropriate parameter, there appears to be wild disagreement about how it should be assessed.

The problem is that loudspeakers tend not to have flat response. It is very tempting, therefore, for a speaker company's marketing department to look for a peak in that unflat response and say that, because the speaker gets that loud at that frequency, that's the sensitivity. The bandwidth of a loudspeaker will also affect the measured sensitivity if wide-band noise is used as a test signal. Two speakers may sound equally loud on music, but on noise, the model with better extension at the frequency extremes will measure as having a higher sensitivity. What is needed, therefore, is a means of producing a measured sensitivity that correlates with perceived loudness.

In 1990 Ronald Aarts of Philips [9] carried out a study of the effect of loudspeaker response on perceived loudness. He concluded that weighting the spectral balance with the popular noise A-weighting curve gave limited correlation with subjective loudness, instead calculating the loudness in phons based on critical-band analysis (ISO932B). Critical-band analyzers not being easy to come by, at *Stereophile*, I feed the loudspeaker with 20kHz-bandwidth noise at a standard level, capture the output waveform with the DRA Labs MLSSA system used in its storage-oscilloscope mode [10], and apply B-weighting to the 1/10-octave-smoothed power response to reduce the effect of bandwidth differences. *Stereophile* has performed four sets of single-blind tests involving 30 different loudspeaker models with the loudnesses equalized on the basis of the B-weighted sensitivity assessed in this manner [11, 12, 13, 14]. These tests indicate a reasonable connection with perceived loudness. ss on music. Only with loudspeakers that had a *grossly* unflat on-axis response did it still prove somewhat unreliable. Analysis of the archived data using an ISO932B figure is a future project.

What is a typical loudspeaker sensitivity? Fig.1 tabulates the calculated B-weighted sensitivities for 261 of the loudspeakers reviewed in *Stereophile* between January 1991 and June 1997. The mean measured sensitivity is 88dB(B)/2.83V/m; the median is 85dB(B). Almost 40% of the models measured had B-weighted sensitivities falling between 84.5dB and 87.4dB. The distribution seems roughly "Normal," with only a few models falling below 80dB(B) or above 90dB(B). The low-sensitivity models tend to be panel speakers of various kinds, while all the speakers of 95dB(B) sensitivity or higher are professional monitors (and one musical instrument speaker measured for curiosity's sake).

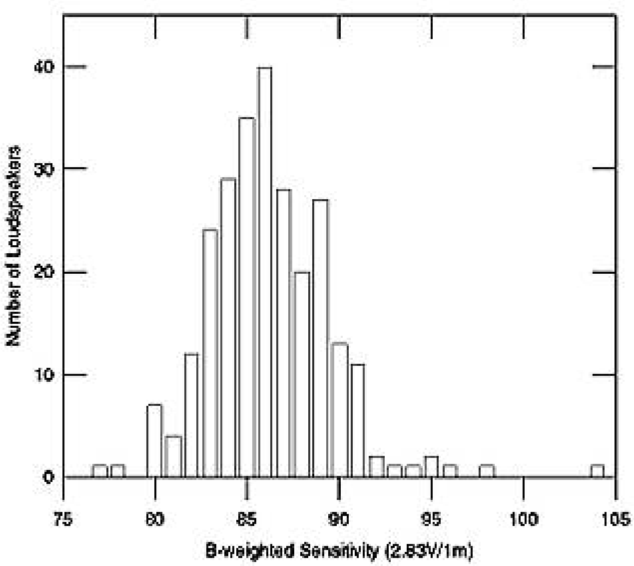

Fig.1 261 loudspeakers, B-weighted voltage sensitivities.

The relatively narrow spread of impedances should not come as a surprise. To achieve a high sensitivity requires heroic engineering, which will always be expensive. All speaker designers are constrained by budget constraints, which in turn mean that they will tend to settle on similar compromises involving magnet size, voice-coil design, and cone area.

In general, my measured sensitivities are slightly low compared with those published by the speakers' manufacturers. Mostly, I believe this is due to the optimistic nature of published sensitivity specifications. But it does need to be pointed out that *Stereophile*'s office is located in Santa Fe, New Mexico at an altitude of 2150m (7000'). A series of experiments, measuring the same loudspeaker samples in Santa Fe and at sea level [15, 16, 17], showed that the only substantial effect on loudspeaker performance due to the altitude was a reduction in sensitivity. Rather than measure sensitivity directly, therefore, I compare each speaker's spl with that produced by a reference speaker (a BBC LS3/5A) that I have measured both at altitude and at sea level. It is possible, therefore, for a systematic error to have crept into my measurements. However, I don't believe this to be more than ±1dB.

Electrical Impedance
If the amplifier that drives a loudspeaker behaves as a voltage source, the loudspeaker's impedance will indicate how much current it will suck from that amplifier. Impedance is both reactive (*ie*, it has a nonzero electrical phase angle) and varies with frequency [18].

It can be measured in two ways. First is to measure the electrical impulse response of a voltage divider formed by the loudspeaker and a standard resistor and calculate the transfer function by performing a Fast (Discrete) Fourier Transform on the time-domain data [19]. Second is to use a swept sinewave signal or a series of discrete-frequency sinewaves to drive the loudspeaker-under-test, keeping the current constant and measuring the magnitude and phase independently [20]. In practice, the constant-current source is replaced by a constant-voltage source and a series high-value resistor, which means that error creeps in for high-value impedance maxima [21].

Then there is the fact that the first method drives the speaker-under-test with a wide-band noise signal while the second drives it with one frequency at a time. Some engineers conjecture that speakers will behave differently to the two stimuli. Having used both methods, I have only found minor differences between the results, if at all. However, for practical reasons, I use the second method, using an Audio Precision System One to drive the loudspeaker in 240 frequency steps from 10Hz to 50kHz with 6V via a series 600 ohm resistor (actually the unit's source impedance). While a true plot of a loudspeaker's complex impedance would be a three-dimensional "pigtail" plotted against magnitude, phase, and frequency, I plot a conventional Bode plot of magnitude and phase against frequency (fig.2), two of the two-dimensional "shadows" cast by the three-dimensional pigtail plot. I find this more informative than a plot of the complex impedance (the end view of the pigtail). I also use! a linear impedance scale rather than the logarithmic scale preferred by some engineers, which I also find more informative. To ensure consistency and accuracy, I regularly measure the impedance of a reference 0.5% tolerance 10-ohm resistor.

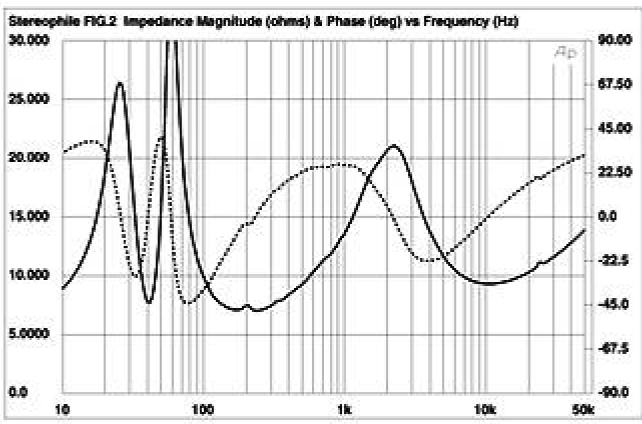

Fig.2 Typical two-way loudspeaker, measured electrical impedance magnitude (solid trace) and phase (dashed trace) plotted against frequency in Hz.

In private communications back in 1991, both Fred Davis and Don Keele pointed out that *Stereophile*'s loudspeaker impedance phase curves published between late 1990 and early 1991 were upside down, in that the positive and negative phase angles were reversed [22]. (Negative phase angle, the current leading the voltage, is due to the load being capacitive; positive phase angle, with the current lagging the voltage, is due to the load being inductive.) This appeared to be due to a software bug; *Stereophile*'s published curves since then have been corrected, with negative phase angles plotted below the frequency axis. However, looking at impedance plots published in books and other magazines reveals wide disagreement on this. Some writers still appear to confuse capacitive and inductive phase angle, while others conventionally plot negative phase angle *above* the frequency axis.

What is a typical loudspeaker impedance? Fig.3 tabulates the mean impedance magnitudes of 330 of the loudspeakers measured from January 1991 through June 1997 (six speakers had integral amplifiers or were intended to be used with external line-level crossovers; the impedance of the others were not measured for various reasons). Only a small number of models feature mean impedances below 4 ohms or above 15 ohms, with the overall mean equal to 8.6 ohms. The median value is 9.25 ohms. According to Martin Colloms [18], the German DIN standard requires that a loudspeaker's impedance magnitude not vary more than ±20% from its nominal value. However, I have found loudspeaker impedances tend to range higher than this, the average standard deviation on the mean 8.6 ohms being 3.7 ohms or 43%. As long as the driving amplifier has a low source impedance, this variation in impedance magnitude will not introduce any sonic effects. But when a tube amplifier is used, which can have a so! urce impedance of perhaps a few ohms, the result can be a significant variation of the loudspeaker's perceived frequency balance, as has been pointed out in *Stereophile* [23].

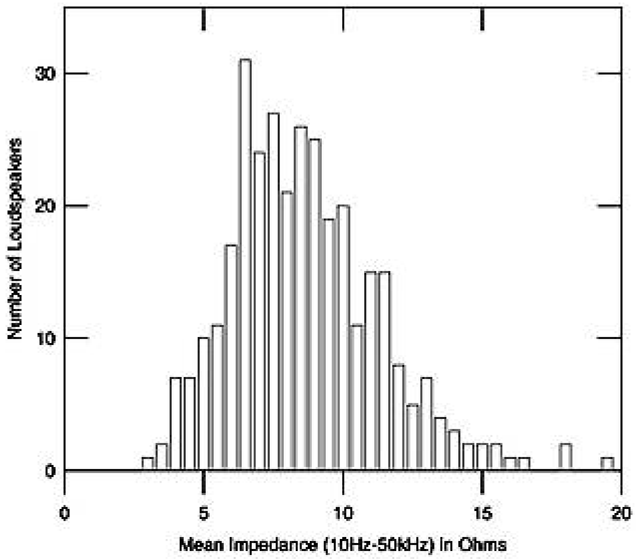

Fig.3 330 loudspeakers, mean impedance magnitudes.

Fig.4 tabulates the *minimum* impedance magnitudes of the same 330 loudspeakers. It can be seen that there is much less variation than with the mean values, with the majority of models having a lowest impedance between 2.76 and 4.26 ohms. The mean and median were 4.3 ohms. In fact, disregarding an exotic electrostatic design that resembled a short circuit above 30kHz, and early samples of two speakers that featured an exotic crossover and dropped below 0.5 ohms above 100kHz, only 6 models had a lowest measured magnitude below 2 ohms. For interest's sake, the loudspeaker with the highest minimum value—8.33 ohms at 171Hz, with a phase angle of -3.1 degrees—was a 1977-vintage BBC LS3/5A. One interesting national difference emerged: moving-coil speakers designed by British engineers tended to have their minimum impedance in the high treble, around 10kHz, while those from US engineers tended to have it in the low midrange.

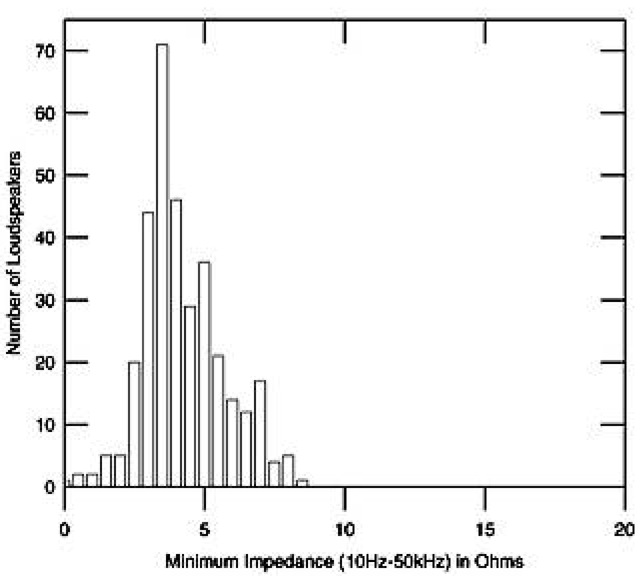

Fig.4 330 loudspeakers, minimum impedance magnitudes.

The impedance measurement is a major diagnostic tool. It is possible to find out a lot about how a loudspeaker is going to behave just from looking at its electrical impedance. Without even seeing the speaker, the number of "ways" will almost always be apparent from the impedance plot, as will whether it is a sealed-box design (one hump in the bass), a reflex or a transmission line (two humps in the bass), or a horn of some kind (a series of regularly spaced peaks). The approximate low-frequency extension will be apparent from the shape of the plots in the bass, as will its Q. The "saddle" between the twin low-frequency impedance peaks typical of a reflex design indicates the tuning of the port, which is generally the bass frequency where the loudspeaker's response has dropped by 6dB, approximately half the reference loudness.

You can also find out whether there are resonances present in the system and what kind they are. For example, if you perform the measurement with sufficient frequency resolution, small glitches appear in the plots due to cabinet resonances of various kinds [24]. Those present at 200Hz and 450Hz in fig.2 are due to cabinet vibrational modes, for example, while the glitch at 27kHz is due to the tweeter's first dome or "oil-can" resonance, where the central region of the dome is moving in the opposite direction to the surrounding annulus.

Fig.5 shows the impedance magnitude and phase for a different model where the designer decided he would throw out all the damping inside the cabinet because to him it made it sound better. Well, it probably made the speaker sound more "exciting," but there were both major air-space resonances and panel resonances present all through the midrange; predictably, they showed up strongly in the speaker's impedance plot as easily visible glitches. Before I had even auditioned the speaker, I could predict from fig.5 that it would have severe midrange coloration problems. The reviewer (Tom Norton) had already noted these problems in his auditioning well before he had seen this measurement.

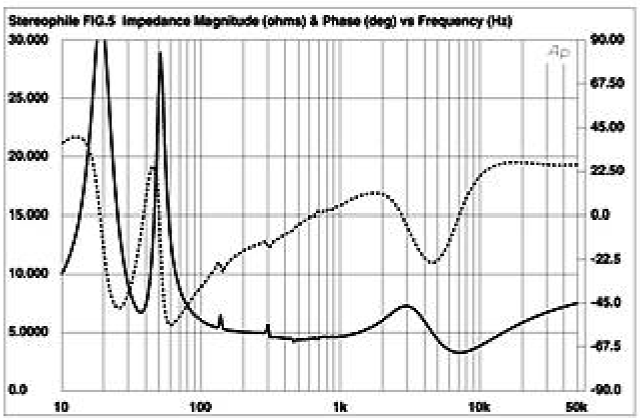

Fig.5 Loudspeaker with undamped cabinet panels, measured electrical impedance magnitude (solid trace) and phase (dashed trace) plotted against frequency in Hz.

Perhaps most important, you can predict from its impedance plot how hard the loudspeaker is for an amplifier to drive. Because a loudspeaker's impedance is reactive, the current will lag or lead the signal voltage by the phase angle [25]. In the worst case—when the phase angle is 90 degrees—the amplifier is required to source the maximum current at the same time as the signal voltage approaches zero. Simply to specify a loudspeaker as having an 8 ohm nominal impedance, therefore, can be misleading. Depending on the phase angle of the impedance, which will be *different* at every frequency, the loudspeaker could look to the amplifier as having a much lower impedance. However, as the late Peter Baxandall pointed out in an Audio Engineering Society presentation in 1987 [26], the maximum phase angle *never* occurs when the impedance has its lowest amplitude. As both are two-dimensional projections of a three-dimensional phenomenon, they're mathematically related.

There has been much conjecture on this subject for the past 20 years in audiophile circles. Eric Benjamin of Dolby examined the matter in detail in 1992 [27]. (The "References" section of his paper is comprehensive.) In particular, he looked at the theoretical power dissipation in an amplifier's class-B output stage, which will depend on the speaker impedance magnitude and phase angle. While he found that the maximum current drawn by a loudspeaker almost never exceeds that predicted by the minimum impedance value, the calculated dissipation in the amplifier output devices was between 120% and 270% of that predicted, depending on the impedance and the drive signal.

It appears, therefore, that amplifier output stages need to be overspecified from slightly to considerably, depending on the loudspeakers they are required to drive. That most amplifiers *don't* appear to be significantly stressed is due to the fact that, as E. Brad Meyer has pointed out [28], at typical listening levels in a typical room with typical loudspeakers, a consumer amplifier is never required to deliver more than a few watts.

Nevertheless, a speaker's impedance behavior can have a primary effect on its sound quality. Figs.6 & 7, for example, show two of the lowest-impedance loudspeakers that I have measured, yet their impedance curves vary significantly, particularly regarding phase angle. Fig.6 shows a speaker that remains above 5 ohms below 2kHz, but plunges to 1 ohm above 12kHz, with an amplifier-crushing combination of 2.4 ohms and -70 degrees at 3.3kHz. By contrast, fig.7 shows a loudspeaker that has an amplitude of 3 ohms or below over almost all the band, yet the phase angle is benign. It is not easy to predict just from looking at these graphs which would be the harder speaker for an amplifier to drive. A future *Stereophile* project is to calculate Benjamin's "peak drive difficulty" data for the loudspeakers I have measured for the magazine in order to develop an index figure that can be quoted in reviews.

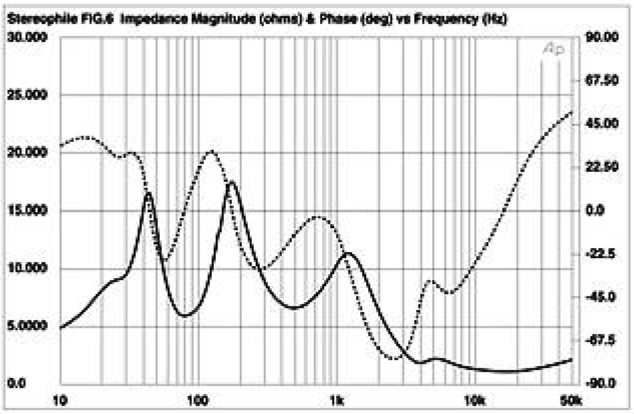

Fig.6 Bad amplifier load 1, measured electrical impedance magnitude (solid trace) and phase (dashed trace) plotted against frequency in Hz.

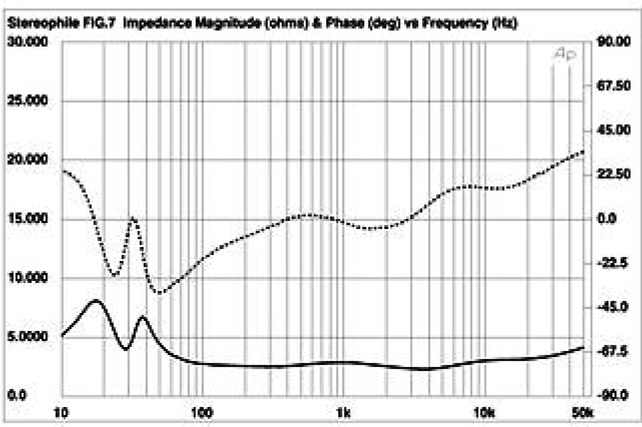

Fig.7 Bad amplifier load 2, measured electrical impedance magnitude (solid trace) and phase (dashed trace) plotted against frequency in Hz.

Part 2 of this article examines the concepts of a loudspeaker's time-domain behavior and nonlinear distortions of all kinds. Part 3 will examine all the various types of frequency response.

REFERENCES

[1] M. Colloms, *High Performance Loudspeakers*, Fourth Edition, pp.391-395 (John Wiley & Sons, 1991).

[2] F.E. Toole, "Listening Tests—Turning Opinion into Fact," presented at the 69th Convention of the Audio Engineering Society (1981 May). Preprint 1766. *J. Aud. Eng. Soc.* Vol.33, pp.439-440 (1982 June).

[3] F.E. Toole, "Subjective Measurements of Loudspeaker Sound Quality and Listener Performance," *J. Aud. Eng. Soc.* Vol.33, pp.9-12 (1985 January/February).

[4] *Perception of Reproduced Sound 1987*, Symposium Proceedings edited by S. Bech & O. Juhl Pederson (1987).

[5] M. Colloms, P. van Willenswaard, "Pace, Rhythm, & Dynamics," *Stereophile* Vol.15 No.11, pp.76-97 (1992 November).

[6] M. Colloms, *High Performance Loudspeakers*, Fourth Edition, p.366 (John Wiley & Sons, 1991).

[7] K. Satoh, "Sound Reproduction Devices and Systems," *Audio Engineering Handbook*, pp.7.20-7.23, edited by K. Blair Benson (McGraw-Hill Book Company, 1988).

[8] J. Atkinson, "Loudspeaker Sensitivity or Efficiency," *Stereophile* Vol.17 No.6, p.83 (1994 June).

[9] R.M. Aarts, "The Calculation of Loudness of Loudspeakers During Listening Tests," presented at the 88th Convention of the Audio Engineering Society (1990 March). Preprint 2928. *J. Aud. Eng. Soc.* Vol.39, p.27 (1991 January/February).

[10] D.D. Rife, "Microphone Calibration, Absolute SPL Measurements," DRA Labs *MLSSA Reference Manual v.10.0A*, pp.96-101 (1996).

[11] T.J. Norton, J. Atkinson, et al, "9 Loudspeakers," *Stereophile* Vol.14 No.7, pp.111-148 (1991 July).

[12] T.J. Norton, J. Atkinson, et al, "8 Loudspeakers," *Stereophile* Vol.15 No.5, pp.137-177 (1992 May).

[13] T.J. Norton, J. Atkinson, et al, "5 Loudspeakers," *Stereophile* Vol.16 No.1, pp.192-237 (1993 January).

[14] T.J. Norton, J. Atkinson, et al, "8 Loudspeakers," *Stereophile* Vol.16 No.9, pp.197-156 (1993 September).

[15] J. Atkinson, "As We See It: Under Pressure," *Stereophile* Vol.13 No.5, pp.5-9 (1990 May).

[16] J. Atkinson, "The Effect of 2150m Altitude on Measured Loudspeaker Performance," unpublished report (1992).

[17] J. Atkinson, "Acoustic Measurements at High Altitude (HI-ALT.ATK)," CompuServe CEAUDIO Library, Section 4 (Audio Issues/BAS) (1994 November).

[18] M. Colloms, *High Performance Loudspeakers*, Fourth Edition, pp.369-371 (John Wiley & Sons, 1991).

[19] D.D. Rife, "Loudspeaker Impedance," DRA Labs *MLSSA Reference Manual v.10.0A*, pp.186-189 (1996).

[20] R. Metzler, "Loudspeaker Testing With System One," *Audio Precision Application Note No.3*, pp.14-15 (1989).

[21] V. Dickason, *The Loudspeaker Design Cookbook*, Fourth Edition, pp.126-129 (Audio Amateur Press, 1991).

[22] J. Atkinson, "Magneplanar MG2.6/R," *Stereophile* Vol.14 No.11, p.167, footnote 5 (1991 November).

[23] T.J. Norton, "Questions of Impedance Interaction," *Stereophile* Vol.17 No.1, p.109 (1994 January).

[24] K. Satoh, "Sound Reproduction Devices and Systems," *Audio Engineering Handbook*, p.7.19-7.20, edited by K. Blair Benson (McGraw-Hill Book Company, 1988).

[25] *Basic Electricity*, Second Edition, compiled by the Bureau of Naval Personnel, pp.206-214 (Dover, 1969).

[26] P.J. Baxandall, "A Technique for Displaying the Current and Voltage Output Capability of Amplifiers and Relating This to the Demands of Loudspeakers," presented at the 82nd Convention of the Audio Engineering Society (1987 March). Preprint 2421. *J. Aud. Eng. Soc.* Vol.36, p.3 (1988 January/February).

[27] E. Benjamin, "Audio Amplifiers for Loudspeaker Loads," presented at the 93rd Convention of the Audio Engineering Society (1992 October). Preprint 3432. *J. Aud. Eng. Soc.* Vol.42, pp.670-683 (1994 September).

[28] E.B. Meyer, "Amplifier Power Requirements Explained (POWER.EBM)," CompuServe CEAUDIO Library, Section 4 (Audio Issues/BAS) (1994 November).

Click **here** to return to this article.

Measuring Loudspeakers, Part Two

John Atkinson, December 1998

This series of articles was initially written (in slightly different form), as a paper presented at the 103rd Audio Engineering Society Convention, New York, September 1997. The preprint, "Loudspeakers: What Measurements Can Tell Us—And What They Can't Tell Us!," AES Preprint 4608, is available from the AES, 60 East 42nd Street, Room 2520, New York, NY 10165-0075. The AES internet site, http://www.aes.org/ *, offers a secure transaction page for credit-card orders.*

In the first part of this series of articles, I examined why I feel a review magazine should include measurements in its loudspeaker reviews. I also went into some detail on the concepts of electrical impedance and voltage sensitivity and what they mean to audiophiles. Now it's time to turn to the basics of a speaker's acoustic behavior.

Impulse & Step Responses

Before the days of digital storage oscilloscopes, it was difficult to examine the time-domain behavior of loudspeakers. Capturing the time record of a component's output is almost trivially easy these days, however. This puts a powerful tool in the hands of audio engineers—and even audio reviewers!

The fundamental test signal in the time domain is the impulse response—the output of the component when presented with a narrow rectangular voltage pulse—which carries within it a complete characterization of a component's linear performance. However, there are practical difficulties with using pulses to test loudspeakers—the very high dynamic range leads to a compromised Signal/Noise Ratio—so I use a method based on Maximum-Length Sequence test signals [29], realized in a commercial piece of test equipment, the MLSSA measuring system from DRA Labs [30].

MLSSA feeds the loudspeaker a pseudo-random sequence of rectangular voltage pulses. By performing a cross-correlation function between the test signal and the signal picked up by a microphone, the host PC can calculate the system's impulse response. It is important to remember that this is the calculated impulse response, *not* the response of the speaker-under-test to an electrical impulse. As pointed out to me in 1991 in a private communication by Verity's Graham Bank, then with Celestion, the two are not the same. The loudspeaker might well behave differently when presented with the low-dynamic-range MLS pulse train than it would when driven by a high-dynamic-range single impulse. The assumption is also made that a linear system is being tested and that is not necessarily the case. The presence of nonlinear behavior in the loudspeaker being tested leads to spurious "echoes" or reflections in the calculated impulse response [31].

Fig.8 shows a perfect impulse response. The voltage instantaneously rises from the graph's time axis, remains at a fixed DC level for a small period of time, then instantaneously drops back to the time axis. The width of this rectangular pulse is inversely proportional to the frequency bandwidth of the signal. If the impulse were infinitely narrow on this graph's time axis, it would require a system with infinite bandwidth—"from DC to light," was how one of my college professors used to describe it—to be reproduced with its shape intact.

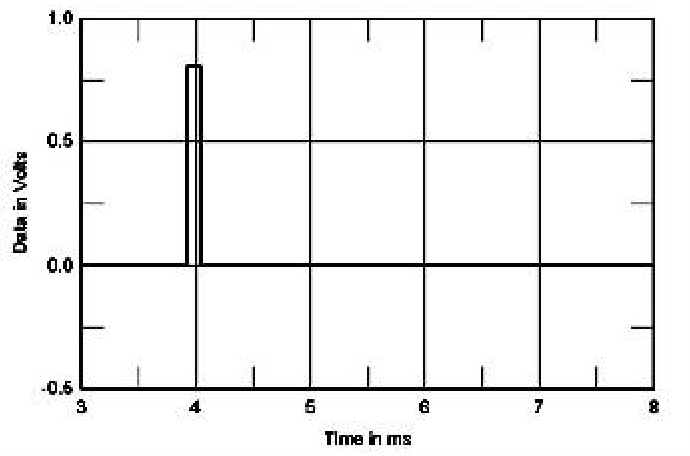

Fig.8 Perfect impulse response.

An analogy can be drawn with the Heisenberg Uncertainty Principle: the more tightly you confine something in one domain, the more uncertain it becomes in an alternate domain. If you confine the rectangular pulse in the time domain to an infinitely narrow time, its spectrum in the frequency domain becomes infinite. Conversely, if you confine the frequency-domain representation of a signal, its spectrum, to a single line—*ie*, there is just one frequency present—the time-domain representation of the signal extends from the beginning of all time to the end of all time.

A typical loudspeaker's impulse response is shown in fig.9, this acquired with a 30kHz bandwidth and with the MLSSA system's antialiasing filter set to a Thomson low-pass function, which doesn't ring. The sharp up-spike and down-spike is the tweeter's output, followed by lower-frequency information, a few ripples and reflections of the sound wave from the speaker's baffle and cabinet, and finally the reflections of the sound from the room boundaries starting at the 7.5ms mark. An impulse response is extremely hard to interpret, not least because, with a loudspeaker, it is visually dominated by the tweeter's output. Its shape doesn't really tell you much in itself about woofer and midrange-unit behavior. It is also important to note that the impulse response and any information derived from it are only relevant to the loudspeaker's output at the specific microphone position used to capture it.

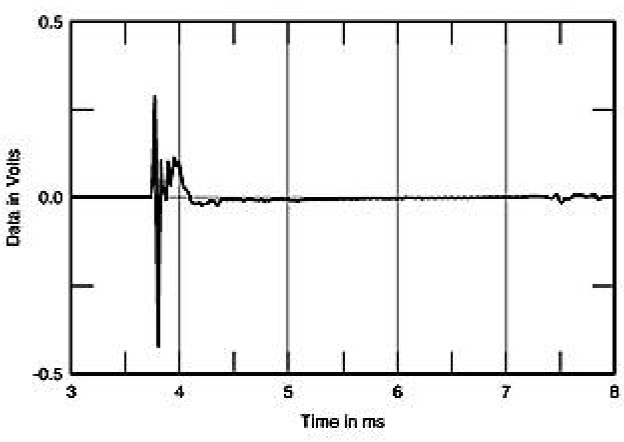

Fig.9 Typical loudspeaker impulse response.

What I have found more useful as a diagnostic tool is the speaker's step response. Instead of driving the loudspeaker with a single rectangular pulse, you feed it a voltage which instantaneously rises from zero to a positive value and stays there (fig.10). This has practical certain problems, so I use the MLSSA software to calculate a loudspeaker's step response from its impulse response. A plot of the step response appears to give more-or-less equal visual weighting to the outputs of all of a speaker's drive-units. You can now glean a lot more information about the lower-frequency drivers, as well as getting a good idea of how time-coherent the speaker is.

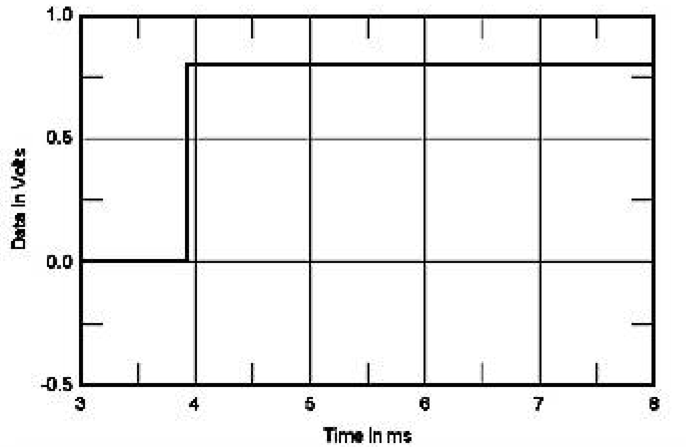

Fig.10 Perfect step response.

The acoustic pressure in a speaker's step response cannot rise instantaneously because the risetime is limited by the tweeter's restricted ultrasonic response. But as long as the speaker has a bandwidth greater than the upper audio limit of 20kHz, the difference in risetime is trivial. More importantly, because a loudspeaker cannot reproduce the acoustic equivalent of a DC voltage—all loudspeakers can be modeled as a high-pass filter of some kind—the sound decays back to the time axis in an exponential manner. It always crosses that axis but eventually returns to it so that the areas on both sides of the time axis are equal, ensuring that there is no DC component present. An ideal step response should therefore resemble a slightly concave right triangle.

Fig.11 shows a good step response produced by a time-coherent, three-way loudspeaker, with the outputs of the three drive-units adding in-phase at the microphone position. There are not that many speakers that produce this good a step response. Of the speakers I have measured for *Stereophile*, only about 10—models from Quad, Thiel, Dunlavy, Spica, and Vandersteen—have step responses this good.

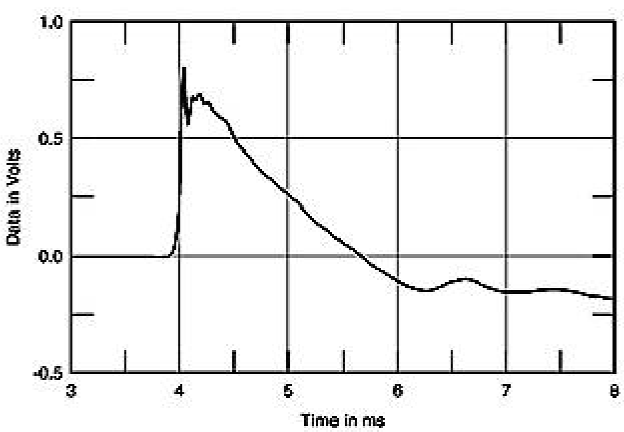

Fig.11 Good loudspeaker step response.

Fig.12 shows a more typical step response, again of a three-way loudspeaker. This time there are actually three step responses apparent in the graph: a narrow, positive-going step response from the tweeter; the next, negative-going step is the midrange unit (as will be seen, it's connected with opposite polarity to the tweeter); with finally a slow, wide positive pulse from the woofer. To confirm this analysis, figs.13, 14, & 15 show the step response of the three units, measured individually. The tweeter's and woofer's step responses (figs.13 & 15) initially depart from the time axis in the positive direction, meaning that both units are electrically connected with positive absolute polarity: a positive electrical voltage results in a positive acoustic pressure. However, the midrange unit's step response (fig.14) initially moves away from the time axis in the negative direction, showing that this drive-unit is electrically connected in inverted absolute polarity.

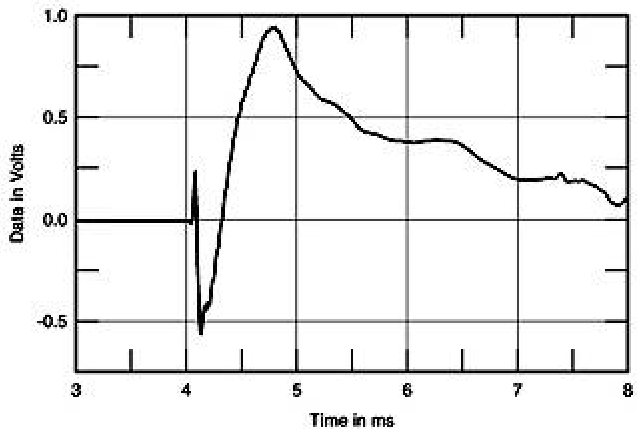

Fig.12 Step response calculated from the time information in fig.9.

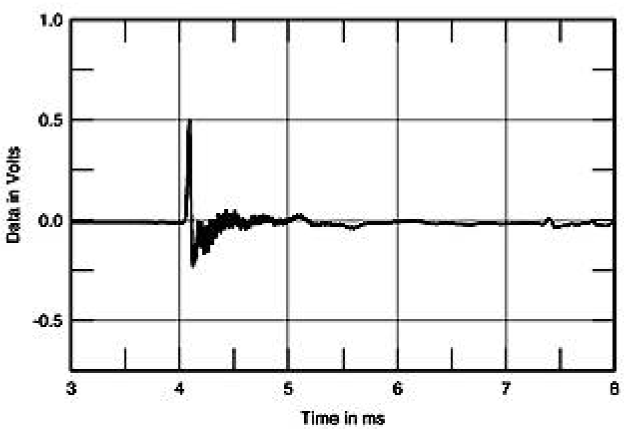

Fig.13 Tweeter step response from fig.12.

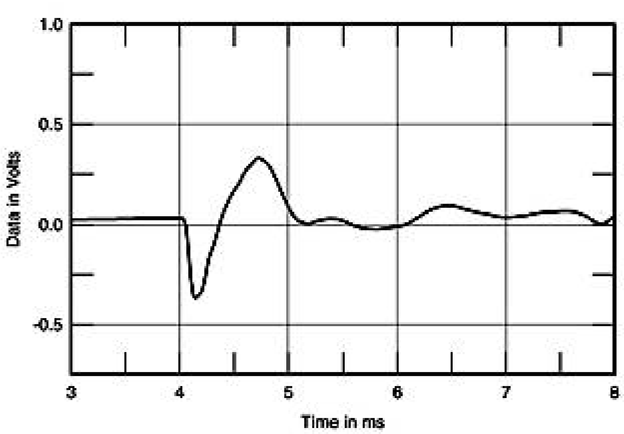

Fig.14 Midrange unit step response from fig.12.

Note that the woofer's step response (fig.15) features a very slow risetime as well as a slow decay. Many audiophiles talk about a loudspeaker having a "fast woofer." Fig.15 reveals that there can be no such thing. A woofer's risetime is dominated by the crossover low-pass filter, which discards the high-frequency information associated with a quick move away from the time axis. A "fast woofer" is therefore an oxymoron. However, I believe that when people talk about "fast woofers," what they're really referring to is after-the-event behavior associated with the Q or "Quality Factor" of the speaker system's low-frequency tuning. Does the woofer stop quickly after the exciting signal has passed? Or does it keep moving, adding low-frequency ringing—"boom"—to the speaker's sound?

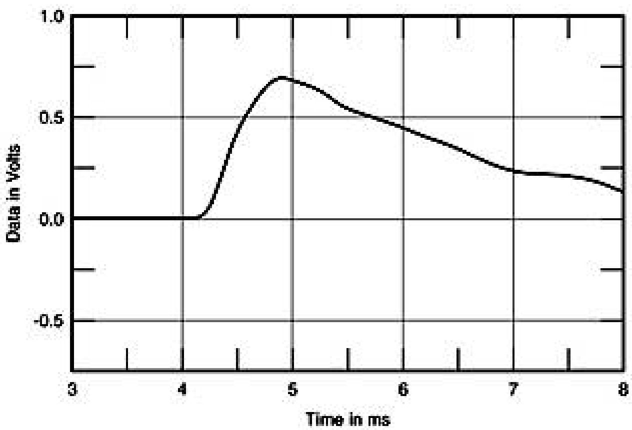

Fig.15 Woofer step response from fig.12.

It appears that the designer of the speaker featured in these four graphs has chosen to use high-order crossover filters of some kind, which necessarily introduce significant (180 degrees or greater) phase shift in the crossover region. To this must be added the phase shift due to the time delay between the units, and the additional 180 degrees phase shift due to the inversion of the midrange's electrical polarity. The result is an on-axis amplitude response in which the drive-units add in-phase to give (we hope) a flat response. The tradeoff is that the system's time coherence is sacrificed.

Many loudspeakers are claimed by their manufacturers' marketing departments to be time-coherent. There are also a number of speakers that have sloped front baffles, implying that they are time-coherent. However, its step response immediately gives you an indication of whether or not a loudspeaker is time-coherent (on the chosen measurement axis). And almost all loudspeakers are not. Along with false claims of high sensitivity, false claims of time coherence are among the commonest lies in high-end audio.

Acoustic Phase Responses

Does a loudspeaker's time coherence matter? A "perfect" speaker, of course, would have both a perfect impulse response and a perfect frequency response (at one point in space). Another way of looking at a loudspeaker's time-domain performance is to examine its acoustic phase response, the phase angle between the pressure and velocity components of the sound plotted against frequency.

Again, this is an aspect of loudspeaker behavior that has proved controversial. One school of thought holds that it is very important to perceived quality; another, which includes almost all loudspeaker engineers, finds it unimportant. Floyd Toole, now with Harman International but then with Canada's National Research Council, in his summary of research at the NRC into loudspeaker performance that is described in two classic 1986 papers [32, 33], concluded thusly: "The advocates of accurate waveform reproduction, implying both accurate amplitude and phase 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...the limited results lend support for the popular view that the effects of phase are clearly subordinate to amplitude response."

This is also my view. Of the 350 or so loudspeakers I have measured, there is no correlation between whether or not they are time-coherent and whether or not they are recommended by a *Stereophile* reviewer. However, I feel that if other factors have been optimized—on-axis response, off-axis dispersion, absence of resonance-related problems, and good linearity—like a little bit of chicken soup, time coherence (hence minimal acoustic phase error) cannot hurt. In my admittedly anecdotal experience, a speaker that is time-coherent (on the listening axis) does have a small edge when it comes to presenting a stereo soundstage, in terms of image focus and image depth. But time coherence does not compensate for coloration, poor presentation of instrumental timbres, a perverse frequency balance, or high levels of nonlinear distortion.

In 1990, Rodney Greenfield and Malcolm Omar Hawksford [34] used DSP-based digital filters to try to separate the audible effects of a loudspeaker's phase error from its amplitude response error. The point was made that a semi-reverberant environment will tend to mask phase effects. In addition, when typical recordings are played, which may have undergone many phase-altering stages during production, the audibility of phase differences becomes moot: "one is simply detecting a change in phase distortion and not a correction of it and as such preferences would most likely be personal." Nevertheless, the authors "very tentatively" concluded that equalizing a loudspeaker's excess phase error modified listeners' perception of the apparent soundstage.

It is important to note that there are phase responses and phase responses. Not only is phase error associated with a lack of time coherence, phase error is introduced by *any* departure from a flat amplitude response. The phase response of what is called a "minimum-phase" system is related to the amplitude response by a mathematical function called the Hilbert Transform. So, for any nonflat system like a loudspeaker, it is important to distinguish between the two sources of phase error. The MLSSA system allows its operator to subtract the calculated minimum-phase, amplitude error-related phase response from the measured acoustic phase response. The result is what is called the "excess" phase response mentioned above. (Note that the MLSSA system does measure the actual acoustic phase. Some competing measurement systems calculate just the amplitude-related minimum phase response.)

Fig.16 shows a typical two-way loudspeaker's excess-phase response on its reference axis, plotted from the lower midrange upward in frequency. (Note that the time of flight of the sound from the loudspeaker to the measuring microphone has to be windowed out of the impulse response before the actual phase response is calculated.) It should be clear that this is *not* a time-coherent design. There's a negative or leading phase error that increases above 2kHz in a linear manner with logarithmic frequency.

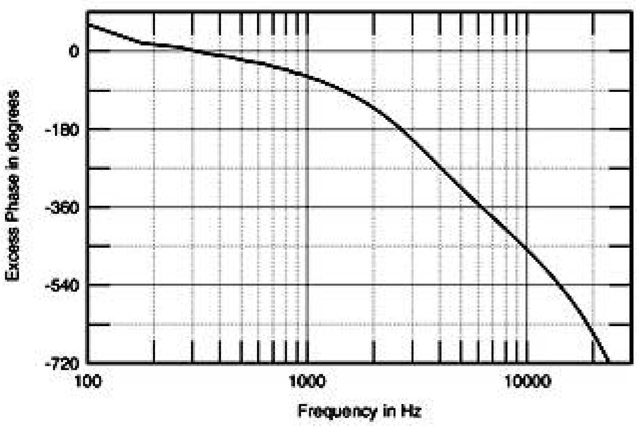

Fig.16 Acoustic excess phase response of typical loudspeaker.

This negative-slope, approximately straight line in the frequency region covered by the tweeter means that there is a simple time delay between the tweeter's output and that of the woofer. The drive-units are mounted on a flat baffle, but a tweeter is physically less deep than a woofer, meaning that its acoustic center is closer to the microphone. Unless some time delay is introduced in the tweeter drive signal, a flat-baffle speaker will never have a flat excess-phase response.

By contrast, fig.17 shows the excess-phase response of the speaker whose step response was shown in fig.11. This model uses a sloped front baffle and a crossover with first-order acoustic slopes to give a time-coherent performance on the listening axis. There's a little bit of positive error at high frequencies, meaning that the microphone is just a little high of the optimal axis.

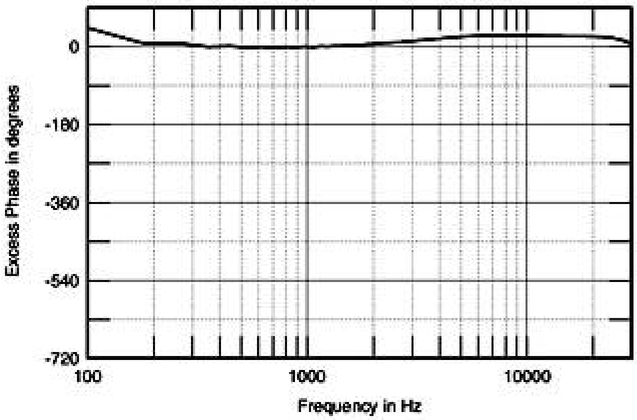

Fig.17 Acoustic excess phase response of time-acoherent loudspeaker.

But, as I said above, the fact that almost no loudspeakers perform in this manner does not stop many of them from getting good reviews, either in this magazine or in others.

Delayed Resonances

By taking a loudspeaker's impulse response, applying a window to the time data, using the FFT operation described in the sidebar to produce an amplitude response plot [35], then repeatedly moving the window along in time by an arbitrary number of samples and again taking an FFT, a three-dimensional graph can be produced [36]. This Cumulative Spectral-Decay (CSD) or "waterfall" plot—amplitude is plotted against frequency on the *x* axis and against time on the *z* axis—can be revealing of resonant problems in a loudspeaker's acoustic output that might not be associated with a peak in the amplitude response.

As mentioned in the sidebar, the shorter the time record examined, the higher the frequency at which meaningful data starts and the larger the gap between the frequency bins in the plot. With the physical conditions in the room in which I measure loudspeakers for *Stereophile*, I can get a reflection-free time window of around 3.5-4ms, meaning that the lowest frequency at which meaningful data exists is (1000/3.5)Hz or (1000/4)Hz = 250-285Hz. In the graphs for *Stereophile*, therefore, I don't plot data below 300Hz. The resolution in these plots is sufficient to reveal woofer and midrange cone problems at the top of their passbands, as well as tweeter resonances. The FFT window chosen has an effect on the display: the shorter the window's risetime, the more true the CSD plot is to the loudspeaker's transient behavior; however, by choosing a longer risetime, frequency-domain aspects are more easily seen.

Fig.18, for example, shows such a plot calculated for a small two-way loudspeaker, with the data windowed to exclude both the time of flight and any reflections from room boundaries and microphone stand hardware that might have been present. The FFT window was a Blackman-Harris type with a risetime of 0.15ms. The first dome resonance from the metal-diaphragm tweeter can be seen as a ridge of delayed energy at 27kHz. A second, less severe resonance-induced ridge can be seen at 3kHz, associated with a small peak in the amplitude response. It dies away relatively quickly, but its presence will make this speaker somewhat intolerant of recordings that have a little too much energy in the same region. (Some listeners might call this character "analytical.")

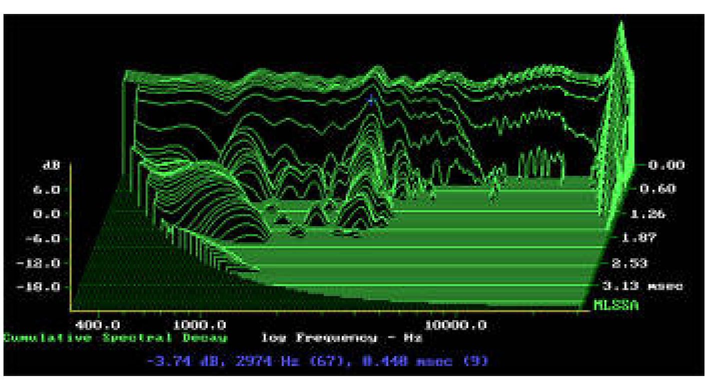

Fig.18 Typical cumulative spectral-decay plot (0.15ms risetime).

By contrast, fig.19 shows the CSD of a loudspeaker whose on-axis frequency response is severely unflat. The graph is dominated by a severe mode at 3750Hz. In addition to the speaker's timbral problems, this resonance could be heard as adding a hard "zinginess" to the overall sound. However, loudspeakers with such audible resonant problems appear to be very rare these days.

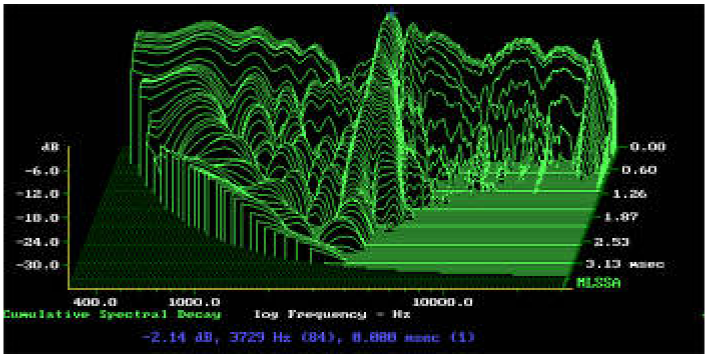

Fig.19 Poor cumulative spectral-decay plot (0.15ms risetime).

To produce a meaningful CSD, the time data need to be free both from noise—it helps to average as many separate impulse response measurements as the host PC can manage—and from environmental reflections, such as those from the microphone stand and its associated hardware, and the speaker support. If not, such reflections produce spurious ridges in the plot that might be interpreted as indicating the presence of resonances. Again, it is also important not to aggressively window the data and produce too short a time record. While this can produce smooth-looking plots [37], they are misleading. Version 10.0A of the MLSSA software flags the area in a CSD plot with dots where the data are invalid due to an inadequate time record (shown in the bottom-left corners of figs.18 & 19). Hawksford [38] has also suggested modifying the CSD plot by compensating for the loudspeaker's minimum-phase behavior. This should make lower-frequency resonances easier to see, but I have yet to try it.

Floyd Toole and his associate Sean Olive did considerable work on the audibility of resonances [39, 40]. It is generally held that high-Q, high-but-narrow peak resonances are less objectionable than low-Q, low-but-broad peak resonances. It is also held that dips in the amplitude response that might also be associated with resonant behavior are less audible than peaks. In my experience, the cleaner-looking a loudspeaker's CSD plot in the upper midrange and treble—above 1kHz, say—the better the chance it will receive a positive review. Loudspeakers that are praised by listeners for "good clarity," "low grain," or "excellent transparency" tend to have clean-looking CSD plots. Conversely, loudspeakers that are referred to as being "grainy" or "harsh" have hashy-looking CSD plots (although, of course, such parameters as nonlinear distortion and frequency balance also contribute to such comments).

Panel Vibrational Behavior

Generations of audiophiles have tried rapping speaker cabinets with their knuckles to see how "dead" the enclosure is. Some enclosures sound like a block of stone, others sound more like a xylophone. Lipshitz, Heal, and Vanderkooy [41] concluded from calculations of the total radiated energy that the sound of the cabinets of the loudspeakers with which they were experimenting would be audible or close to the borderline of audibility. A loudspeaker I reviewed in 1997 [42] had cabinet resonances that were so severe that if you played music through it then paused the CD player, you'd hear audible reverberation at the listening chair as the excited resonances died away.

I have not yet found it practicable to produce quantitative information on cabinet resonant behavior. However, to look at the behavior of loudspeaker cabinets in a more rigorous way than the simple "knuckle-rap" test, I make use of MLS excitation and CSD plots. An inexpensive piezoelectric-tape (polyvinylidene fluoride) accelerometer, 4" long by 1" wide and similar to an acoustic guitar transducer [43], is taped to the cabinet walls at various places, the cabinet is excited with a 2kHz-bandwidth test signal from the DRA Labs MLSSA system at a standard level, and an impulse response is calculated/captured.

Stanley Lipshitz and his colleagues noted that the accelerometer measurements of a loudspeaker cabinet's walls varied tremendously according to how the speaker was supported while the measurement was being performed. For *Stereophile* reviews, I support each loudspeaker with three upturned metal cones that contact the base of the speaker in the center at the rear and at the two front corners. This allows resonant modes to develop to their fullest, according to the results of a series of experiments I carried out examining this subject in detail [44]. These tests also showed that the best means of coupling a speaker to its stand—"best" in the sense of maximally reducing the amplitudes of cabinet vibrations—was to use a "lossy" coupling material, such as Blu-Tack.

Fig.20 shows a typical impulse response calculated from the accelerometer's output. The ringing overlaying the decay tail of the impulse is obvious. Fig.21 shows a cumulative spectral-decay plot calculated from the time-domain data in fig.20. Four or five resonant modes can be seen.

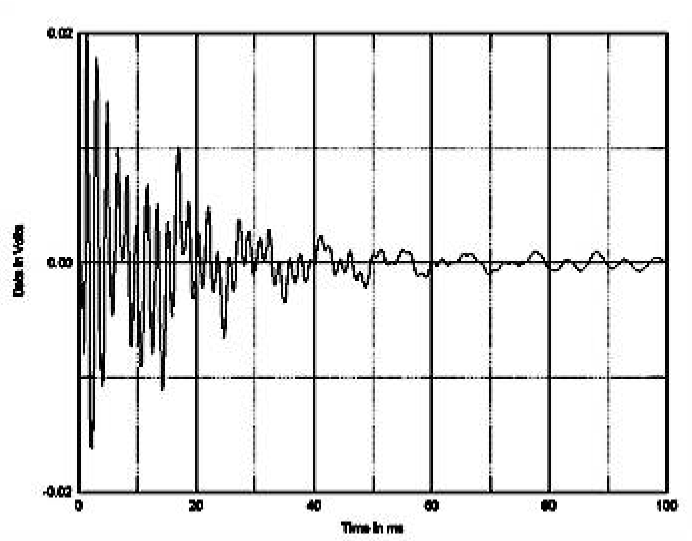

Fig.20 Impulse response calculated from output of PVDF tape accelerometer fastened to center of loudspeaker cabinet sidewall (100ms time window).

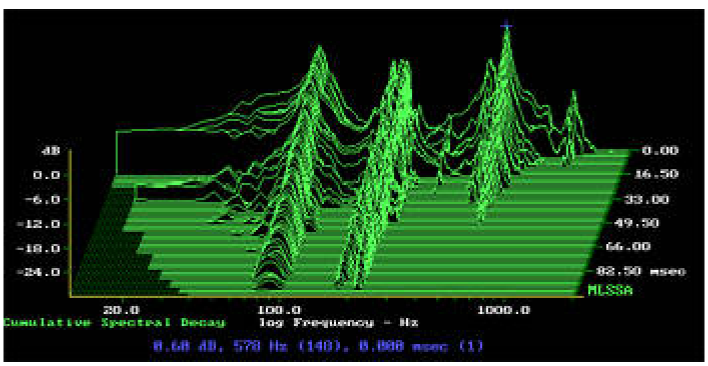

Fig.21 Cumulative spectral-decay plot (0.15ms risetime) calculated from the time data in fig.20.

It is hard to predict the effect of such behavior on perceived sound quality. The amplitude of the modes might be small, but a loudspeaker cabinet can represent a much larger radiating area than its drive-unit(s). A panel may be very lively, but if it faces away from the listener, its subjective effect may be minimal.

It must be a good idea in general to reduce the amplitude of resonant modes, but how audible a high-Q, narrow resonance will be depends not only on its level and the area of the radiating panel, but also its frequency. If the resonance lies in the upper bass, say between 100 and 200Hz, it will likely coincide with specific musical notes as well as being excited most of the time. An octave or so higher, and a narrow-bandwidth, high-Q resonance might well "fall in the cracks" between the discrete frequencies of musical notes. (With equal-temperament Western music, the notes are farther apart in frequency the higher they lie on the musical staff.) In addition, the higher in frequency these resonances are, the more subjectively benign they appear to be, probably because they decay faster.

While a higher-frequency resonance would still be excited by wide-band, noise-type signals like drums, the resonance might remain lurking undetected for most of the time, in which case it might as well not exist. On the other hand, if it coincides with the frequency of a musical note, it can be excited continually. Thus if you examine two similar loudspeakers, each of which has a strong panel resonance present in the midrange, one might sound colored while the other sounds clean.

This also suggest a mechanism for some "tweaks" that are claimed to improve the sound of loudspeakers. By mass-loading the panel and adding damping, the frequency and Q of any resonances present may be shifted by only a little, but enough to move them into the musical gaps [45].

But what if you damp such a resonance only a little? This might have the paradoxical effect of making it more audible, as it will now be more likely to be excited more often, due to its lower Q. (A broad, shallow peak covers more frequencies than a narrow, sharp one.) However, any detectable change in sound tends to be reported by audiophiles as being an improvement!

However, looking at the behavior of the 300 or so loudspeakers that I have measured, several common factors emerge from the auditioning that correlate with the presence of strong cabinet resonances between 100Hz and 500Hz. (Remember that other "objective" factors will also contribute to the same subjective perceptions.) The clarity in the lower midrange can be disappointing. Tenor instruments like cello or trombone lack clarity or acquire a "woody" character. The bass can sound muddy, diffuse, one-note, blurred, or lacking in power, rather than tight, articulate, and extended, as it does in real life. Music can seem to drag, in rhythmic terms. Male voices can "boom" and female voices "hoot" at some frequencies and not others, with the result that the little inflections of tone that are characteristic of real voices become diluted. Centrally placed images, particularly of voices, can smear toward the speaker positions at some frequencies.

Harmonic Distortion

There is considerable discussion in the literature of nonlinear (harmonic) distortion in loudspeaker behavior [46, 47, 48]. All loudspeakers have nonlinear distortion, and small, inexpensive loudspeakers tend to have more nonlinear distortion than large, expensive loudspeakers. Perversely, I don't think this is that important a factor in loudspeaker performance. I have measured loudspeaker harmonic distortion spectra when listening tests had suggested that it was unusually high or low [49, 50]. I have also investigated distortion when I have found a loudspeaker producing audible *sub*-harmonics, tones whose frequencies are an integral fraction, one half, one third, one quarter, of the fundamental [51]. In a presentation at the 1989 Audio Engineering Society Convention in New York, the mathematician Manfred Schroeder postulated that the production of subharmonics is often related to the presence of chaotic behavior in a diaphragm. This latter phenomenon can be heard on *Stereophile*'s *Test CD 2*, Track 25.

But of all the loudspeakers that have been reviewed in *Stereophile* in the past eight years, there are only a few in which noticeable levels of harmonic distortion have been associated with negative review findings. However, I do conjecture that listeners use overall distortion to set a comfortable playback level. If a loudspeaker has high intrinsic distortion, hence a limited dynamic range, it won't be played as loud. Once the level of harmonic distortion rises above a threshold (probably one that is different for each listener), the listener reaches for the volume-control knob. I realize, of course, that my opinions on this subject will be controversial.

Part 3 of this series will examine what is meant by a loudspeaker's frequency response and how loudspeakers behave in rooms.

References

[29] D.D. Rife, J. Vanderkooy, "Transfer Function Measurement Using Maximum-Length Sequences," presented at the 83rd Convention of the Audio Engineering Society (1987 October). Preprint 2502. *J. Aud. Eng. Soc.* Vol.37, pp.323-348 (1989 June).

[30] D.D. Rife, "Introduction, Theory, & Loudspeaker Time Coherence," *DRA Labs MLSSA Reference Manual v.10.0A*, pp.7-9, pp.145-150, & pp.159-163 (1996).

[31] J. Vanderkooy, "Aspects of MLS Measuring Systems," *J. Aud. Eng. Soc.* Vol.42, pp.219-231 (1994 April).

[32] F.E. Toole, "Loudspeaker Measurements and their Relationship to Listener Preferences: Part 1," *J. Aud. Eng. Soc.* Vol.34, p.231 (1986 April).

[33] F.E. Toole, "Loudspeaker Measurements and their Relationship to Listener Preferences: Part 2," *J. Aud. Eng. Soc.* Vol.34, pp.335-336 (1986 May).

[34] R. Greenfield, M.O. Hawksford, "The Audibility of Loudspeaker Phase Distortion," presented at the 88th Convention of the Audio Engineering Society (1990 March). Preprint 2927.

[35] R.N. Bracewell, *The Fourier Transform and Its Applications*, Second Edition, Revised (McGraw-Hill, 1986).

[36] D.D. Rife, "Waterfall & Cumulative Spectral Decay Plots," *DRA Labs MLSSA Reference Manual v.10.0A*, pp.77-82 &

pp.158-159 (1996).

[37] *La Passion du Haut-Parleur Audax*, Polydax Speaker Catalog (1994).

[38] M.O. Hawksford, "Minimum Phase Correction of Loudspeaker Cumulative Spectral Decay Plots," presented to the London Section of the Audio Engineering Society (1995 February); M. Colloms, "Industry Update: Digital filters correct speaker response," *Stereophile* Vol.18 No.5, pp.35-37 (1995 May).

[39] F.E. Toole, S.E. Olive, "The Perception of Sound Coloration Due to Resonances in Loudspeakers and Other Audio Components," presented at the 81st Convention of the Audio Engineering Society (1986 November). Preprint 2406.

[40] F.E. Toole, S.E. Olive, "The Modification of Timbre by Resonances; Perception and Measurement," presented at the 83rd Convention of the Audio Engineering Society (1987 October). Preprint 2487.

[41] S. Lipshitz, M.K. Heal, J. Vanderkooy, "An Investigation of Sound Radiation by Loudspeaker Cabinets," presented at the 90th Convention of the Audio Engineering Society (1991 March). Preprint 3704.

[42] J. Atkinson, "Total Music Systems Adiabat: Measurements by JA," *Stereophile Guide to Home Theater* Vol.3 No.3, pp.128-129 (1997 Fall).

[43] V. Dickason, *Voice Coil* Vol.5 No.1, pp.1-2 (1991 November).

[44] J. Atkinson, "The Sound of Surprise," *Stereophile* Vol.15 No.9, pp.162-170 (1992 September).

[45] J. Atkinson, J. English, "Microscan D-8 Vibration Damper," *Stereophile* Vol.16 No.8, pp.150-156 (1993 August).

[46] K. Satoh, "Sound Reproduction Devices and Systems," *Audio Engineering Handbook*, pp.7.23-7.31, edited by K. Blair Benson (McGraw-Hill Book Company, 1988).

[47] M. Colloms, *High Performance Loudspeakers*, Fourth Edition, pp.338-339 (John Wiley & Sons, 1991).

[48] F.E. Toole, "Loudspeaker Measurements and their Relationship to Listener Preferences: Part 1," *J. Aud. Eng. Soc.* Vol.34, pp.233 (1986 April).

[49] J. Atkinson, "Avalon Eclipse," *Stereophile* Vol.14 No.1, p.225-234 (1991 January).

[50] J. Atkinson, "B&W Silver Signature & Velodyne DF-661," *Stereophile* Vol.17 No.6, p.83 (1994 June).

[51] J. Atkinson, "Magneplanar MG2.6/R," *Stereophile* Vol.14 No.11, p.164-171 (1991 November).

Sidebar: The Fast Fourier Transform

As I have implied in this article, a representation of a signal in the time domain has a related representation in the frequency domain. There is a mathematical operation, the Fourier Transform, that can be used to calculate the frequency-domain representation—the signal's spectrum—from its time-domain representation—the signal's impulse response. In practice, a mathematical shortcut is used, leading to the designation of the mathematical operation as the *Fast* Fourier Transform or FFT (footnote 1).

The basic premise behind the application of the FFT to a loudspeaker's time-domain behavior is that impulse response be first "windowed" to reject data that are not of interest (room reflections, for example). It is then assumed that the impulse response is part of a continuous repetitive waveform, the end of the data window being spliced to its beginning to produce an infinite series of copies of the impulse response. Applying the FFT produces a line spectrum with its first component having a frequency equal to the reciprocal of the time window.: *ie*, if the impulse response is windowed to have a length of 4ms, the fundamental frequency of the transformed line spectrum is (1/0.004)Hz = 250Hz. The spectral lines present at integer multiples of that 250Hz fundamental—500Hz, 750Hz, 1000kHz, 1250Hz, 1500Hz, etc.—paint out the frequency spectrum of the signal.

It is important to note that no information is present in that spectrum below the fundamental frequency. Nor is any information present between the spectral lines. The windowing operation necessary to perform the FFT eliminates real information. And the shorter the time window, the higher the fundamental frequency in the resultant spectrum and the larger the gaps in that spectrum.—**John Atkinson**

Footnote 1: In the PC dark ages of the early '80s, I wrote a mixed Basic/assembler FFT program that took two *minutes* to calculate a 1024-point FFT on a 6502-based computer. But in these days of 300MHz Pentiums and G3 Macs, the FFT operation is truly fast, that same 1024-point FFT taking milliseconds, if not microseconds. Even 32,768-point FFTs, which would have taken hours in 1982, are now commonplace. (For the jitter spectral analyses that accompany *Stereophile*'s CD player reviews, for example, I average 64 32,768-point FFT calculations in a matter of seconds.)—**John Atkinson**

Click **here** to return to this article.

Measuring Loudspeakers, Part Three
John Atkinson, January 1999

This series of articles is based on a paper presented at the 103rd Audio Engineering Society Convention, New York, September 1997. The preprint, "Loudspeakers: What Measurements Can Tell Us—And What They Can't Tell Us!," AES Preprint 4608, is available from the AES, 60 East 42nd Street, Room 2520, New York, NY 10165-0075. The AES internet site, http://www.aes.org/, offers a secure transaction page for credit-card orders.

In the first two articles in this series, I examined the loudspeaker's electrical behavior and how it behaves in the time domain. But the on-axis amplitude response is the most commonly seen loudspeaker measurement and the one that most people assume best correlates with sound quality. Whether it does or not will be examined in this final article.

There's only one method to assess a speaker's "frequency response" with complete accuracy from 20Hz upward: using a very slowly swept sinewave in a large anechoic chamber. This is a room where sound-absorbing materials on every surface soak up every sound emitted by the speaker. The room is therefore removed from the picture and the only sound that reaches the measuring microphone is therefore that from the speaker. Anechoic chambers are expensive, both in terms of real estate and in their construction, so those of us with limited budgets have to look elsewhere. Three testing strategies can be adopted: LMS (gated toneburst), TEF (swept sinewave with time-delayed swept bandpass filter), and MLS (Maximum Length Sequence) testing. All of these methods allow a loudspeaker's anechoic response to be assessed, after a fashion, in a real room. All my measurements for *Stereophile* are done using the MLS technique, using a commercial instrument, the DRA Labs MLSSA system, a combination of a full-length PC card and a DOS program running on a PC.

To examine a loudspeaker's anechoic amplitude response at one point in space, you first capture or calculate its impulse response. The impulse response carries almost all the information to describe the speaker's behavior. (It doesn't describe the speaker's nonlinear behavior, so it must be remembered that all quasi-anechoic frequency-response measurements assume the speaker being tested produces zero distortion—which is never true!) That impulse response is windowed to eliminate room-boundary reflections in the impulse tail. Then, using the Fast (Discrete) Fourier Transform [52], those time-domain data can be transformed into frequency-domain information—see the sidebar on the Fast Fourier Transform that accompanied the second article in this series (*Stereophile*, December 1998, p.85). However, there are a number of things to consider, some practical, some theoretical.

And there is also the issue of repeatability *vs.* absolute accuracy. My first goal in establishing a loudspeaker measurement regime at *Stereophile* was to be able to reliably produce repeatable measurements. Standardization of techniques and rigor applied to setup is important here. I routinely measure the same sample speakers at regular intervals (a year or so). The measured response is effectively unchanged, meaning that this goal of internal consistency has been achieved. Absolute accuracy is more difficult to achieve and, in a sense, is a never-ending quest. However, when I have been able to compare my measurements of loudspeakers with those of reviewers for other magazines or with those of manufacturers, there is generally a good correspondence, taking into consideration different samples, microphones, techniques, and test conditions. It is obvious that we are all measuring the same model.

The first practical aspect of performing the measurement is the microphone used. For economic reasons—my measuring microphone has to double as a recording microphone—I use a B&K Type 4006, ½", phantom-powered omnidirectional microphone. While it has a very flat amplitude response for a recording microphone, its on-axis response still varies by +2.1dB, -1.5dB between 1kHz and 20kHz [53]. In 1991 it was compared in an anechoic chamber against a B&K instrumentation microphone to generate a response error curve (fig.22, top trace). As long as I always use the same measurement distance and the same microphone axis, it is reasonable to subtract the microphone's amplitude error data from a loudspeaker's measured response. However, I don't compensate for the microphone's phase response, meaning that there's always just a little bit of error in my published phase measurements. But as I said in the second article in this series, I hardly ever find a speaker that has a good enough time response to make it worth showing. The lower trace in fig.22 shows the response error of an inexpensive measuring microphone from Old Colony Sound Lab, the Mitey Mike II (footnote 1), that I have also been using for the past year or so. Designed by Joe D'Appolito and using a small omnidirectional back-electret capsule from Panasonic, the MM II is significantly flatter in response than the B&K.

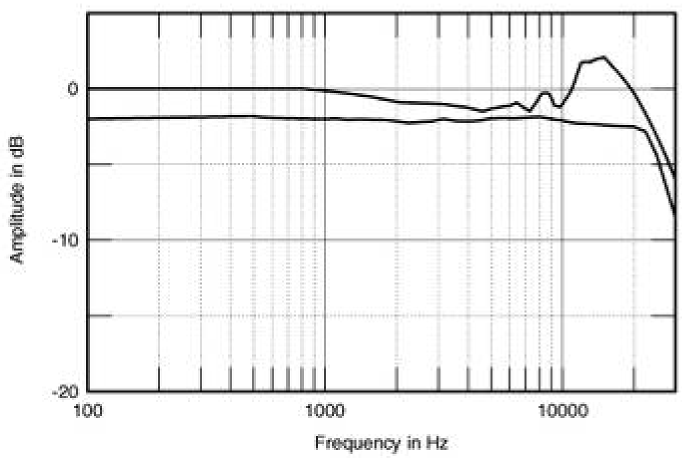

Fig.22 On-axis response of B&K 4006 omnidirectional microphone used for *Stereophile* loudspeaker reviews since 1989 (top). Response of Mitey Mike II also used for reviews since 1997 (bottom).

Footnote 1: Described in full in the Four/97 issue of *Speaker Builder*, the Mitey Mike II is available calibrated or uncalibrated, with either a pre-assembled or a DIY head amplifier from Old Colony Sound Lab, P.O. Box 576, Peterborough, NH 03458. Tel: (603) 924-6371. Fax: (603) 924-9467.

Second is the microphone mounting arrangement. The MLS technique is very sensitive to reflections of the sound emitted by the loudspeaker. It is important, therefore, to mount the microphone so that any reflections from the stand and its associated hardware occur after or at the same time as the inevitable reflections from the room boundaries. The B&K microphone is held endwise by a female XLR plug that is flush-mounted into a 2m length of ½" plumbing pipe. This pipe is held in a conventional microphone stand, but with the upright and clamp 1.5m away from the microphone capsule. Fig.23 shows the difference in the measured response of a BBC LS3/5A loudspeaker at a distance of 50" (1.27m) made by replacing this custom stand with a commercial boom stand and a conventional spring clamp. (The FFT window and time data length were identical for both measurements.) Ripples in the treble response averaging ±0.5dB in amplitude can be seen, these resulting from acoustic reflections from the stand and mounting hardware.

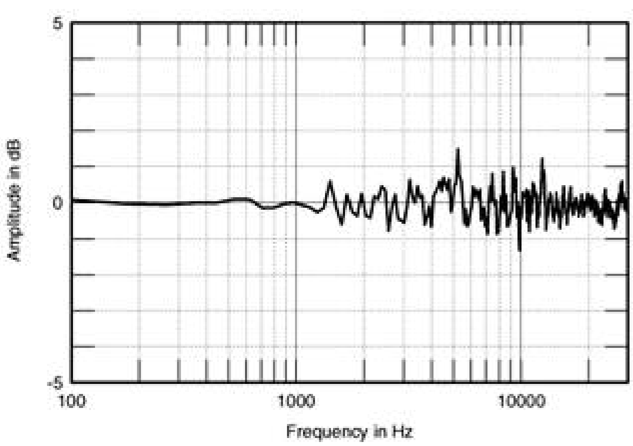

Fig.23 Difference in the measured response of a BBC LS3/5A loudspeaker at a distance of 50" made by replacing custom stand with a commercial boom stand and a conventional spring clamp.

The measurement system's antialiasing filter also introduces a small amount of response error. This too is subtracted from the calculated response for the graphs published in *Stereophile*. If you're not careful, you also get ripples in the measured response from the windowing used to remove room reflections. With a rectangular window, which gives the best frequency resolution, it is important to set the markers where you wish to truncate the impulse-response data at amplitude values as close as possible to zero. Otherwise the result will be ripples in the FFT-derived amplitude response. Tapered windows, where the farther away a datum point is from the center of the time window, the more it is attenuated, can also be used, but this is at the expense of frequency resolution [54].

It is typical Zen wisdom that the ideal room for measuring a loudspeaker using MLS techniques is an anechoic chamber. Not for quietness—the MLSSA system has good immunity to environmental noise, and I can measure loudspeakers in *Stereophile*'s listening room, averaging up to 128 separate measurements to maximize S/N Ratio. However, remember that you have to window out room reflections in order to extract a loudspeaker's anechoic response. Truncating the time data in that manner will reduce both the plot's frequency resolution and the lowest frequency at which valid data will be produced. With a time window truncated to 5ms, for example, the response will only be accurate down to 200Hz. The farther away the room walls, or the more absorptive the room walls, the longer the time window that can be analyzed and the more accurate the measurement. Alternatively, the microphone can be placed farther away from the speaker for the same bandwidth and frequency definition.

In my room, I build an "acoustic black hole" on the floor between the speaker and the microphone stand using a 10" mat of several kinds of absorbent foam. This does a good job of absorbing floor reflections in the midrange and treble, meaning the main room reflection is now that from the ceiling. (This can be seen in a speaker's impulse response—fig.9, *Stereophile*, December 1998, p.79—as the small ripple near the right-hand side of the graph.)

Older engineers are probably used to measuring a speaker's frequency response at 1m. This, I feel, is too close for all but very small speakers (see later). I therefore use a standard microphone distance of 50" (1.27m), which is the farthest I can get back in my room and still preserve reasonable measurement bandwidth. Ideally, I would like to measure at 2m, where a large loudspeaker's drive-units will be better-integrated and where there will be less proximity effect (more on this later).

There is a problem with taking the response at just one point in that there is almost *too* much information. Some of the fine detail will be specific to just that one point in space. With a loudspeaker whose drive-units are mounted in some kind of vertical array, it seemed sensible to implement some kind of spatial averaging. This would smooth out any position-dependent wrinkles in the measured response, while leaving the significant information intact. Accordingly, my published responses are the average of seven measured responses, taken at 5 degree intervals across a 30 degrees horizontal window on the reference axis.

A final matter should be discussed. It is a hidden assumption when measuring a loudspeaker's amplitude response that the microphone is in the far field; ie, is more than a couple of wavelengths away at the lowest frequency of interest. An alternate way of looking at the matter is that the microphone should be at least as far away as the largest dimension of the loudspeaker to be measured. With my standard microphone distance of 50", this assumption will no longer be true for large loudspeakers. With big speakers, such as the various kinds of panel speakers, there will be a proximity effect [55] that tilts up the response at low frequencies. This, of course, will also be true when the loudspeaker is listened to at the same distance.

Assessing the acoustic performance of big panel speakers is therefore an undertaking fraught with difficulty. Some years ago, for example, I had to measure a loudspeaker that had a small dome tweeter that radiated sound only in the forward direction, a large ribbon midrange unit that behaved as a dipole, and an omnidirectional woofer. Both the measured response and the perceived balance of this speaker varied according to how far away the listener and microphone were, rendering meaningless any discussion of this speaker's "frequency response."

How meaningful *is* a loudspeaker's on-axis amplitude response? *Stereophile*'s founder J. Gordon Holt argued over a decade ago [56] that a loudspeaker with a measured flat on-axis response won't sound correct: "Many times in past years I have been impressed by the incredible flatness of the measured high-end response of some speakers...In every such case I have been equally amazed at how positively awful those loudspeakers sounded—so tipped-up at the high end that I could not enjoy listening to them," he wrote, adding that "Audiophile loudspeakers which measure nearly flat through the lower middle range seem to have a penchant for sounding sucked-out and gutless through that region...loudspeakers that measure flat in my own listening room sound thin at the low end, while those sounding flat at the bottom measure as having a low-end rise."

However, another *Stereophile* reviewer, Martin Colloms, disagrees [57], citing the results of single-blind listening tests: "The favored speakers were those which possessed very even axial responses over 100Hz to 10kHz when measured by third octave and octave averaging." As a design goal, he felt that the engineer "should be aiming for a deviation of ±0.25dB or less in the forward directed response, while according rather less importance to narrow band deviations of greater amplitude..." *Audio* reviewer Don Keele [58] adds that "the on-axis response should be smooth, because it defines the spectral balance of the sound that first arrives at the listener and so is of greatest subjective importance in judging timbre."

An analysis of 74 loudspeakers that I performed in 1991 [59] also showed a good correspondence between flatness of measured on-axis response and listener preference. Grouping loudspeakers by the log-frequency-weighted standard deviation of their response between 170Hz and 17kHz—the weighting was to compensate for the linear spacing of the frequency bins produced by the FFT process—I discovered a clear correlation between flat on-axis quasi-anechoic response and the tendency for the loudspeaker to get a positive review in *Stereophile*. This correlation also appeared when the overall results of blind listening tests performed by the magazine were analyzed [60].

I suspect that the measured responses referred to by Holt were taken in-room. In this case, the measured data will include contributions from both the speaker's anechoic axial response and its power response (see later). The presence of absorbent room furnishings and drapes will dramatically affect the balance at high frequencies; the power response of a typical forward-firing loudspeaker tends to slope down with increasing frequency. It is likely, therefore, that a flat in-room response can be achieved only by equalizing the on-axis response to tilt up. This will be much more significant in large than in small rooms (see later).

Fig.24 shows the anechoic or quasi-anechoic response of a good-sounding loudspeaker taken on its tweeter axis with the MLS technique described above, averaged across a 30 degrees horizontal window centered on the tweeter axis and corrected for both the microphone's departure from flatness and the response error introduced by the system's anti-aliasing filter. The curve is actually a composite, consisting of the seven spatially averaged responses taken with a 30kHz bandwidth from 1kHz to 30kHz, a separate on-axis measurement taken with a 5kHz bandwidth from 312.5Hz to 1kHz (this makes the graph look more presentable, but the true frequency resolution is unchanged), and the complex sum of the speaker's nearfield port and woofer responses below 312.5Hz (the complex sum adds the two amplitude responses taking the phase responses into account). A big, high-Q, hence narrow, peak can be seen at the metal-dome tweeter's ultrasonic resonant frequency, but the response through the midrange and treble is otherwise pretty flat: it fits within ±2dB limits over almost the entire audio band.

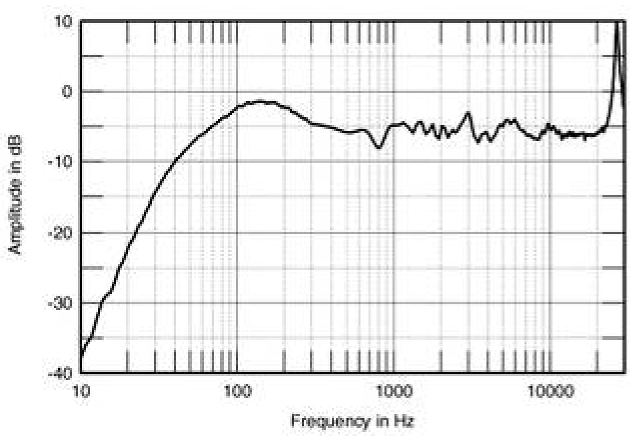

Fig.24 Typical MLS-derived response of a good-sounding loudspeaker at 50", averaged across a 30 degrees horizontal window on the tweeter axis, and spliced to the nearfield LF response.

Note that the frequency resolution between 300Hz and 1kHz is limited, the data points being spaced quite wide apart. But the graph is still informative. A small and probably benign dip can be seen just below 1kHz. There's a small peak at 3kHz that may add a little "bite" to the sound, or it might add a little bit of presence-band emphasis that will make the sound "vivid" without it being so much so that the sound will become aggressive. On the other hand, the peak may be compensating for something else the speaker does, meaning that the perceived sound balance might actually be flat (see later).

By contrast, Fig.25 shows a similar measurement for a loudspeaker that a *Stereophile* blind listening panel thought sounded quite colored, with a skewed presentation of instrumental timbres. The speaker is quite flat in the midrange and has reasonably good bass extension. However, there is a severe suckout in the mid-treble. In all but large rooms, this will produce a characteristic, hollow-sounding coloration. There's some peakiness in the mid-treble region above the suckout that will add a little bit of "zing" to the perceived balance. The designer presumably was trying to compensate for one with the other, as the mean tweeter level is about 5dB below that of the woofer. I conjecture that the designer was probably bothered by the peak at 3.5kHz, which would make the speaker sound bright, all things being equal. Thinking "brightness" correlated with too much tweeter energy, he padded down the tweeter. This was to no end, however, because the real problem appears to be a vicious resonance or breakup mode in the woofer cone. (This can be seen in fig.19 on p.85 of the December '98 issue.)

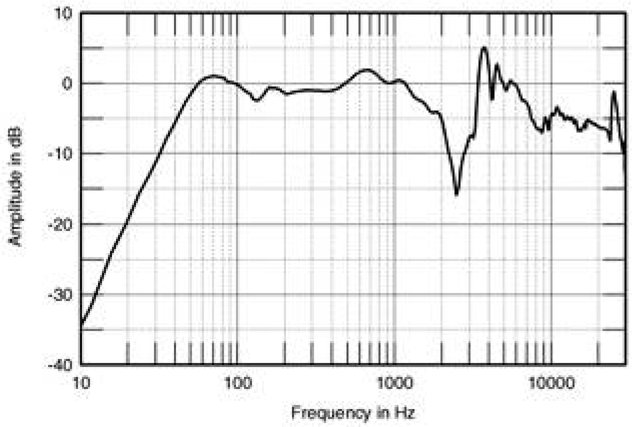

Fig.25 MLS-derived response of a poor-sounding loudspeaker at 50", averaged across a 30 degrees horizontal window on the tweeter axis, and spliced to the nearfield LF response.

The bargraph in Fig.26 tabulates the quasi-anechoic response measurements of 320 of the loudspeakers that I have measured since that 1991 article. Again they are grouped by the standard deviation of their responses over the two decades from 170Hz to 17kHz, weighted to compensate for the fact that an FFT-derived amplitude response has a linear frequency spacing, but is conventionally plotted on a logarithmic frequency scale. The mean deviation from a flat response is 3dB, with only a few loudspeakers having a weighted standard deviation of 1dB or less. On the other hand, only a few loudspeakers have weighted standard deviations greater than 5dB. The overall level of attainment featured by loudspeaker designers is quite good, at least regarding flatness of on-axis response.

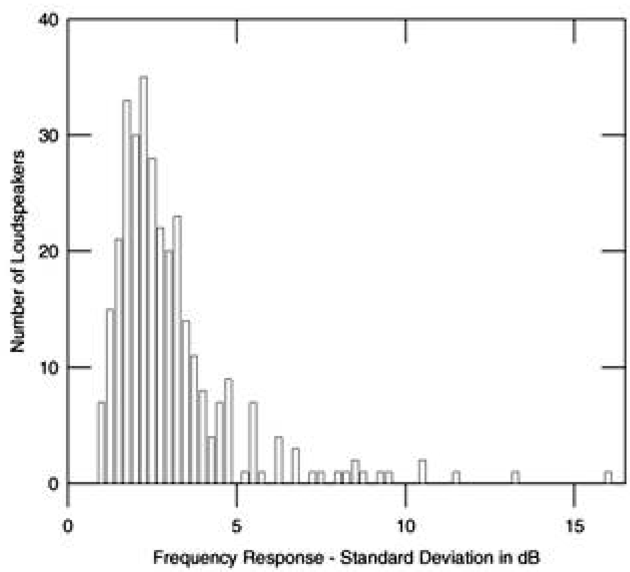

Fig.26 320 loudspeakers, range of frequency-weighted standard deviations, 170Hz-17kHz.

Table 1: Frequency-weighted response standard deviation *vs* recommendation in *Stereophile*'s "Recommended Components."

Response	Number	Number Not	
1.0	15	1	7%
1.5	39	2	5%

1.5	39	2	5%
2.0	73	8	11%
2.5	56	8	14%
3.0-4.0	70	20	29%
4.0-5.0	32	14	44%
>5.0	33	14	42%

Is there a correlation between this arbitrarily chosen measure of response flatness and whether or not a loudspeaker sounds good—or at least gets a good review in *Stereophile*? Looking at the raw data used to compile Table 1, there appears to be no correspondence between flatness of on-axis response and price. But Table 1 does demonstrate that it appears that the chance of being recommended by this magazine increases the flatter a loudspeaker's on-axis response becomes.

However, while flatness of midrange and treble response is a good thing to have in a loudspeaker, it doesn't in itself mean that the speaker will sound good. The model with the overall flattest response in fig.26 was not recommended by the magazine because of problems in other areas of performance. On the other hand, once the response flatness deviates above a certain level—a frequency-weighted standard deviation between 170Hz and 17kHz of approximately 3.5dB, for example—it's unlikely that the speaker will either sound good or be recommended. Exceptions are: a) if the speaker is very cheap or its errors are in a relatively innocuous part of the frequency spectrum or b) if the standard deviation is high because of an overall smoothly tilted response; or c) if the speaker is otherwise superb in almost every respect.

It's also apparent that a speaker with a basically smooth, flat response but a single area of unevenness—due perhaps to a strong woofer-cone breakup mode or a cabinet resonance—will be downgraded accordingly. The same holds true for crossover or drive-unit integration problems. Similarly, if a speaker has a flat response but limited high-frequency dispersion, its appeal will be limited due to the room reverberant field tending to sound too lifeless. And no matter how flat a speaker's midband, a somewhat loose, underdamped bass is an almost unforgivable fault for many listeners, except when balanced by a degree of HF emphasis or a tilted-up treble. Conversely, the worst thing a speaker designer can do with a speaker that has an overdamped, lightweight bass alignment is to give his brainchild a tilted-up treble balance. The lack of across-the-band balance is musically disturbing, and the emphasized treble exaggerates problems elsewhere in the system.

Interestingly, if a speaker is rolled-off at both extremes, it seems to sound more acceptable than might be expected from its flatness quotient as calculated here. If this somewhat midrange-forward balance is not smooth, however, the ear seems to detect the unevenness as a cupped-hands coloration. And the opposite response trend, a lack of energy in the upper midrange, can lead to a musically uninvolving presentation.

Fig.27 shows the responses of the 20 loudspeakers with the lowest standard deviation, 1.375dB or less. These can all be seen to be superbly flat (justifying the use of the log-frequency-weighted standard deviation to characterize the speakers). The best performers are, in fact, flatter than my B&K measurement microphone! By contrast, Fig.28 is the "rogues gallery" of the speakers with the largest response deviations. (It doesn't include panel speakers, which tend to have large standard deviations because their measured responses slope down with frequency, due to the proximity effect described earlier). The high standard deviation is affected by large dips and peaks in the response, or by overall tilts up or down. One small satellite loudspeaker from Bose that I measured for *Stereophile Guide to Home Theater*, for example, smoothly but persistently slopes up by 10dB from 300Hz to 10kHz.

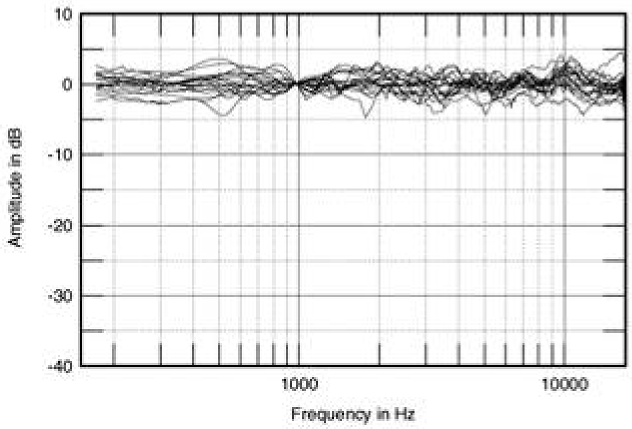

Fig.27 MLS-derived responses at 50" of 20 loudspeakers with lowest standard deviations, averaged across a 30 degrees horizontal window on the tweeter axis.

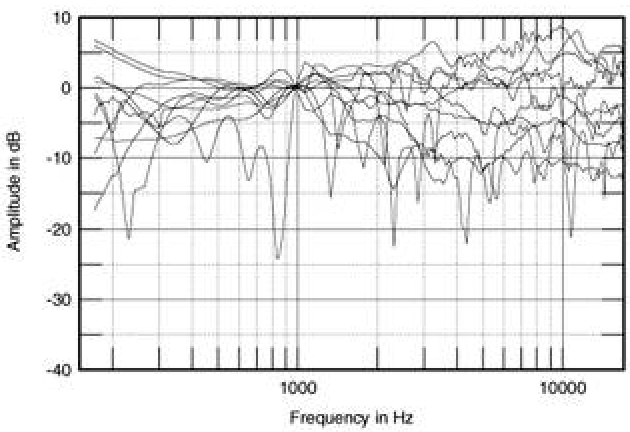

Fig.28 MLS-derived responses at 50" of 13 loudspeakers with highest standard deviations, averaged across a 30 degrees horizontal window on the tweeter axis (not including panel speakers).

The responses of the panel speakers reviewed in the time period are shown in fig.29. The graph is a little confusing, but it should be apparent that the 50" microphone distance does result in a tendency for the responses to slope down from low to high frequencies. This, as mentioned earlier, is due to the microphone still effectively being in the nearfield at low frequencies but in the farfield at high frequencies.

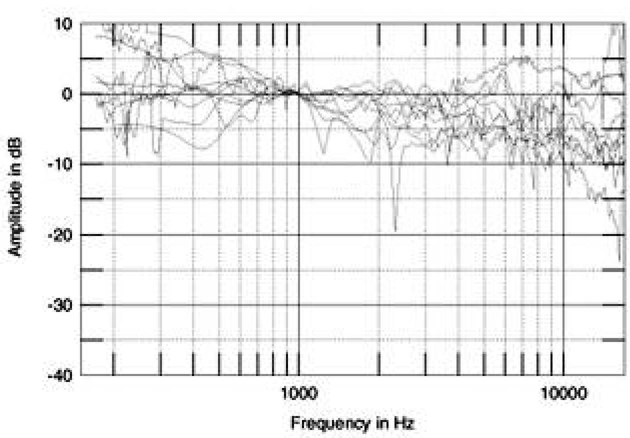

Fig.29 MLS-derived responses at 50" of 11 panel loudspeakers, averaged across a 30 degrees horizontal window on the tweeter axis.

The data for seven professional monitors are included in this survey. Their responses are plotted in fig.30, and it is interesting to note that their on-axis responses are no more flat than the group average. Studio monitors have other tasks to fulfill, however, including offering a much higher dynamic range than is necessary for a domestic design. Nevertheless, one of the flattest overall responses was that of a small active studio monitor from Finnish manufacturer Genelec included in this group.

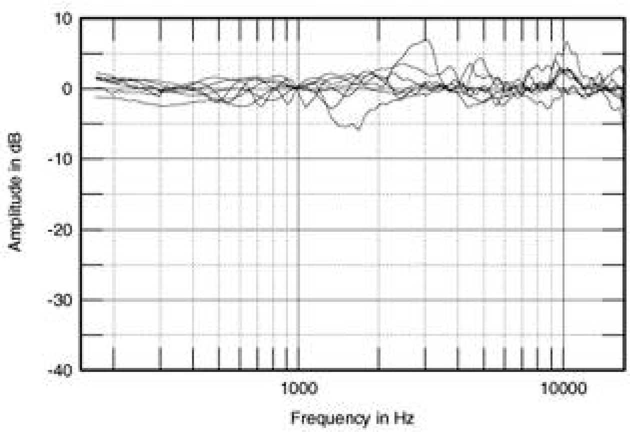

Fig.30 MLS-derived responses at 50" of 7 studio monitors, averaged across a 30 degrees horizontal window on the tweeter axis.

There are five speakers included in fig.26 where the designer has used DSP chips to equalize the on-axis response [61, 62, 63] and in one case to implement the crossover filters in the digital domain. Somewhat surprisingly, while all had low-response standard deviations, there were many conventional passive designs that were as flat or even flatter.

Nearfield Responses

Because the time data need to be truncated to eliminate room reflections, a farfield MLS-derived amplitude response is of little use in characterizing a loudspeaker's behavior at low frequencies. However, a classic 1974 paper by Don Keele [64] discusses the technique of taking a loudspeaker's response in the nearfield, with the microphone capsule placed very close to the radiating diaphragm. This appears to give a response that accurately reflects a loudspeaker's low-frequency output as assessed in the farfield, a conclusion more recently confirmed by Struck and Temme of B&K [65]. If there are several diaphragms, the nearfield responses can be taken individually, then summed in the ratio of the square roots of their radiating diameters [66]. Note that the sum needs to be complex, taking into account both acoustic phase and the phase shift associated with the different distances of each of the diaphragms to a nominal far-field listening position. The measured nearfield response is only valid while the wavelength of the sound is very much larger than the diaphragm size. However, as we are only interested in what happens below 300Hz or so, that is not a factor unless the speaker is very large—again measuring panel speakers proves problematic!

Fig.31 shows an example of a typical ported or reflex design measured in the nearfield. The individual responses of the woofer and port are shown, with their complex sum. The woofer output drops to a minimum at the port tuning or counter-resonance frequency. This graph has 1Hz resolution; with sufficient resolution, a true notch could be seen because when the port has its maximum output, there's so much back pressure on its diaphragm that the woofer can't move. At this point, almost all the speaker's output is coming from the port. Below that frequency, both woofer and port roll out with a second-order, 12dB/octave slope. However, as they are out of phase with one another, their summed response rolls out with an ultimate fourth-order 24dB/octave slope.

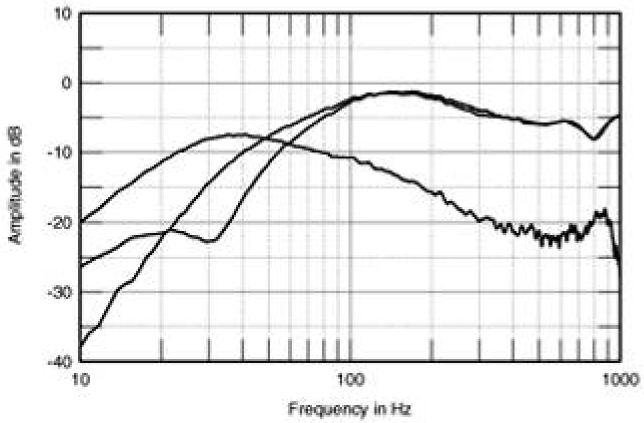

Fig.31 Nearfield low-frequency responses of typical reflex loudspeaker.

The effect of this on a speaker's in-room performance is interesting. Up to the middle of 1997, the vast majority of the 360 speakers I measured were reflex designs—300 models, or 83%—the designer using the port to extend the design's anechoic low-frequency performance. Yet in an actual listening room, the increased rate of low-frequency rolloff of a reflex design leads to less low-bass output than with an equivalent sealed-box design, with its 12dB/octave rollout. However, it is fair to point out that reflex loading doesn't just increase the anechoic LF extension but can also increase power handling and dynamic range in the bass. Similarly, having measured many speakers with exotic LF alignments, ranging from the so-called "transmission line" to multiported, multidriver monstrosities, it is my considered opinion that in almost every case, the same or better bass performance could be achieved with an equivalent sealed-box alignment.

For published graphs, the loudspeaker's nearfield response is spliced to the farfield response in the 300Hz region. However, as pointed out in the Keele paper, the nearfield response assumes a 2-pi or half-space loading for the drive-units—close coupling to the room boundaries. This results in an apparent low-frequency boost in the resultant composite graph compared with a true anechoic response made of the same speaker. Given that a loudspeaker's woofer and port are always within a fraction of a wavelength from one boundary—the floor—and almost always less than one wavelength from three other boundaries—the walls and ceiling—below 100Hz or so, my experience has been that this does give a truer representation of a loudspeaker's real low-frequency performance than the anechoic response in all but extremely large rooms. Certainly, the loudspeakers I have auditioned that have true, flat anechoic extension to very low frequencies sound as if they have a somewhat exaggerated bass response—which is how they appear with a nearfield measurement.

Polar Response/Dispersion

Looking at the anechoic response at one point in space can be instructional, but it must not be forgotten that loudspeakers emit sound in a full sphere [67]. I examine how a loudspeaker's amplitude response changes in two planes, vertical and horizontal, using the DRA Labs system's ability to plot an arbitrary number of FFT-derived response plots against a third, arbitrary variable [68]. The loudspeaker is rotated on a commercial stepper motor-driven turntable (made by the Italian company Outline and available from Old Colony Sound Lab) in 5 degree steps, an impulse response being captured/calculated at each position. There is a practical problem, which is that a loudspeaker's acoustic center almost *never* coincides with its center of mass. Performing polar response measurements in this manner therefore involves some complicated balancing acts: as a result, the more off-axis the measurement, the more error will be introduced.

Other than in the case of multi-directional designs, I limit the range covered from 90 degrees on either side of the reference axis and from 45 degrees below that axis to 45 degrees above. With conventional forward-firing designs, the response to the rear is predictable and correlated with the physical size of the loudspeaker. All such designs are omnidirectional at low frequencies.

Fig.32 shows the resultant forward dispersion in the horizontal plane for a typical two-way design—it uses a woofer/midrange unit with a 180mm or 200mm (8") chassis and a 1" (25mm) dome tweeter—centered on the tweeter axis. The microphone distance was 50" (of necessity a little farther away for the extreme off-axis angles). The on-axis response of the speaker has been subtracted from each of the curves to highlight the manner in which that response changes to the speaker's sides.

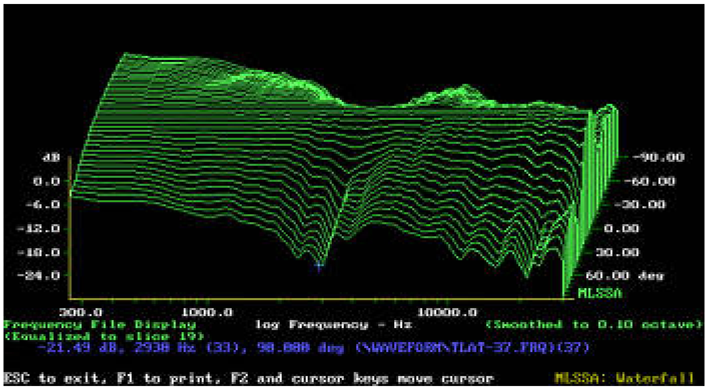

Fig.32 Horizontal dispersion of typical two-way loudspeaker (from 90 degrees on one side to 90 degrees on the other side), normalized to response on tweeter axis.

Below 300Hz the dispersion is approaching omnidirectional, the physical size of the speaker being significantly smaller than the sound's wavelength. However, the loudspeaker's output drops quite rapidly to its sides in the 2-3kHz region. This is because the 8" woofer is a little big to be crossed-over this high in frequency. It starts to beam when the radiating diaphragm starts to become of the same order as the wavelength. From 3kHz to 10kHz, the speaker system's dispersion becomes very wide, due to the tweeter being much smaller than the wavelength of the sound it is emitting. The tweeter's dispersion subsequently narrows above 10kHz when *its* size becomes significant compared to the wavelength.

It is hard to predict the effect of this discontinuity in the dispersion in the crossover region. If the speaker's on-axis response is flat, unless the listener's room is heavily damped with drapes and carpets, the speaker in fig.32 will probably sound bright. There will be too much energy in the room in the mid-treble region compared with the low treble. However, this particular loudspeaker has a slight peak at 3kHz in its on-axis response (fig.24). As a loudspeaker's perceived balance in a room will have contributions from both the on-axis and off-axis responses (see later), the on-axis peak will to some extent be compensated by the off-axis lack of energy in the same region. I believe that much of the fine-tuning performed by loudspeaker designers—commonly referred to as "voicing" a design—involves balancing the on-axis and off-axis responses to give an overall flat perceived in-room response.

Fig.33 shows a very different horizontal dispersion plot. This speaker uses a smaller woofer and therefore has wider dispersion in the crossover region. There is a little bit of excess energy off-axis at the base of the tweeter's passband, and there is a ridge of off-axis energy in the top audio-band octave because the tweeter has a plastic "phase-plate" over the diaphragm. Otherwise the overall dispersion is remarkably even, the design gradually and uniformly becoming slightly more directional with increasing frequency. If its on-axis frequency response was flat (and it doesn't suffer from audible colorations due to resonances), then this loudspeaker will tend to sound neutrally balanced in a typical listening room. I'd be surprised if it had an identifiable character. Reflections of its sound from room sidewalls will not vary significantly from its on-axis sound, other than a reduced high-treble content. All things being equal—pair-matching in particular—this well-controlled way in which a loudspeaker's off-axis sound changes is always associated with good, precisely defined stereo imaging. I haven't come across a speaker I've measured like this that didn't have good stereo imaging.

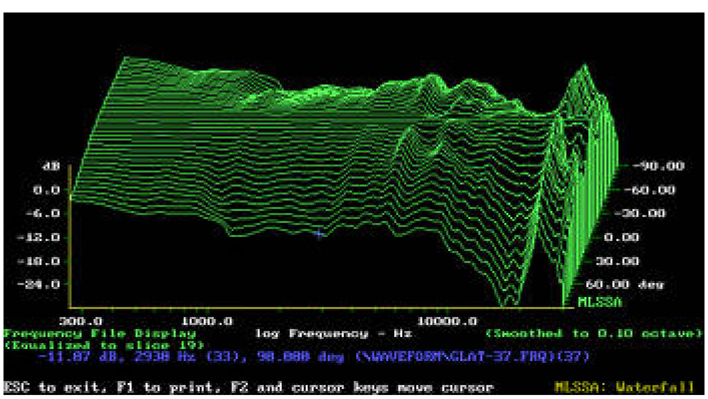

Fig.33 Horizontal dispersion of loudspeaker (from 90 degrees on one side to 90 degrees on the other side) recognized as having good stereo imaging, normalized to response on tweeter axis.

The fact that almost all designs use a vertical array of drive-units means that vertical dispersion is strongly affected by the crossover used. Fig.34 shows the vertical dispersion family for a small two-way design, again normalized to the tweeter-axis response. The crossover-induced lobing can be seen: there's a lack of energy in the crossover region between the woofer and tweeter above the reference axis. This design needs to be used with reasonably high stands if it is not going to sound hollow.

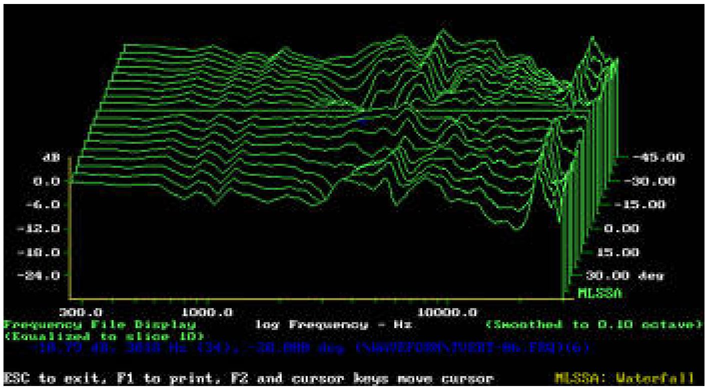

Fig.34 Vertical dispersion of loudspeaker (from 45 degrees below the HF axis, front, to 45 degrees above, back) recognized as having good stereo imaging, normalized to response on tweeter axis.

In-Room Amplitude Response

Unlike other audio components, the perceived sound of a loudspeaker is affected by factors other than its intrinsic performance. All the previous measured acoustic parameters that have been discussed examine the loudspeaker in isolation. But loudspeakers are used in reverberant rooms rather than anechoic chambers, and the interaction between the two is complicated [69, 70, 71, 72, 73, 74]. However, my experience has been that different models of loudspeakers sound different in a consistent manner when similarly set-up and auditioned in different rooms. But it is clear that the sound perceived by the listener consists of a combination of the direct sound of the loudspeaker (the first-arrival sound, which correlates with the anechoic response, at least in the midrange and treble) and the room's reverberant field (which is affected by the loudspeaker's off-axis behavior or overall power response) [75, 76].

The question is, What is the balance between the two factors in that combination? For many years, the reviewers at *Consumer Reports* have held that the power response—the overall power put out by the speaker, summed over a complete sphere—dominates the perceived balance. However, this will only be true in a very large room with the listener sitting a long distance away from the speaker. However, several loudspeakers that I have measured have had the tweeter and midrange units electrically connected so that the units were 180 degrees out of phase in the crossover region between them. This results in a large suckout in that region in the measured anechoic response; this is also audible as a "hollow"-sounding coloration when the loudspeakers are listened to at a relatively close distance. However, upon investigation of some of the circumstances underlying some of the designs, it appeared that the designers were listening to their speakers at least 15' away in very large rooms. Their perception of the loudspeaker's balance was therefore mainly related to its overall power response in the room.

I have measured the in-room response of a subset of 60 loudspeakers. For historical reasons, but also to act as a check on the MLS measurements, I use both a different technique and a different test setup. The rectangular room is my own dedicated listening room. This measures around 19' by 15.5', with a 9.5' ceiling broken up by 9" *vigas*—raw pine logs. The room is carpeted, and there are patches of Sonex foam on the ceiling to damp primary reflections of the sound. The other wall has RPG Abffusors behind the listening seat to absorb and diffuse what would otherwise be early rear-wall reflections of the sound that might blur the stereo imaging precision. ASC Tube Traps are used in the room corners to even out the effect of the room's upper-bass resonant modes, the result being a relatively uniform reverberation time of around 200ms from the upper bass to the middle treble, falling to 150ms above 10kHz.

For this in-room spectral analysis I use an Audio Control Industrial SA-3050A spectrum analyzer with its own microphone. I average six measurements at each of 10 separate microphone positions for left and right speakers individually. These positions are arranged in a rectangular grid 8' wide by 18" high, centered on the position of my ears in the listening chair, 36" from the floor and around 9' from the loudspeaker positions. The 120 original spectra are averaged to reduce the effect of room resonant modes. What you're left with is basically a snapshot of the balance that the listener hears. Fig.35 shows a typical curve (again it's the loudspeaker whose anechoic response was shown in fig.24). This measurement has proved to give a good correlation with a loudspeaker's perceived balance in my room.

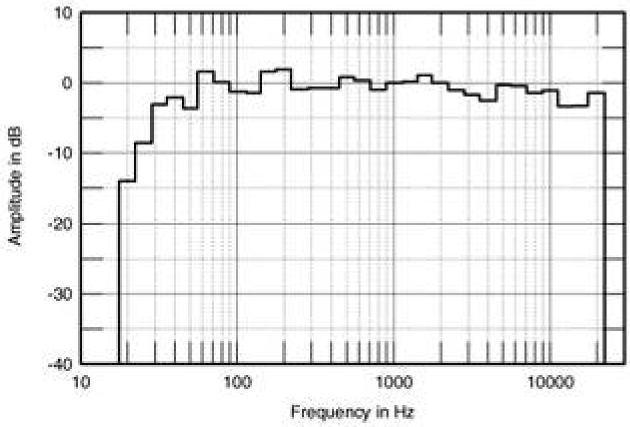

Fig.35 Typical spatially averaged, 1/3-octave response of loudspeaker in JA listening room.

What Makes a Good-Sounding Loudspeaker?
Vance Dickason [77] offers some discussion of this question, but the definitive answers are to be found in Floyd Toole's comprehensive 1986 papers [78, 79]. Nothing that I can conclude from my past eight years' work, at least when it comes to conventional forward-firing, moving-coil designs, is in serious conflict with his findings. As I wrote in 1991 [80], "The best-sounding loudspeakers, in my opinion, combine a flat on-axis midrange and treble with an absence of resonant colorations, a well-controlled high-frequency dispersion, excellent imaging precision, an optimally tuned bass, and also play loud and clean without obtrusive compression."

Overall conclusions
While each measurement of a specific area of loudspeaker performance gives important information regarding possible sound behavior, it emerges that there is no direct mapping between any specific area of measured performance and any specific subjective attribute. As a result:

• Any sound quality attribute always depends on more than one measurement.

• No one measurement tells the whole story about a speaker's sound quality.

• Measuring the performance of a loudspeaker involves subjective choices.

• All measurements tell lies.

• Most important, while measurements can tell you how a loudspeaker sounds, they can't tell you how good it is. If you carefully look at a complete set of measurements, you can actually work out a reasonably accurate prediction of how a loudspeaker will sound. However, the measured performance will not tell you if it's a good speaker or a great speaker, or if it's a good speaker or a rather boring-sounding speaker. To assess quality, the educated ear is still the only reliable judge.

And no matter how good any one measurement, if the beginning of the third movement of Beethoven's Fifth Symphony, where the composer introduces the trombones for the first time, or Jimi Hendrix's hammered-on tremolo at the start of "Voodoo Chile" on *Electric Ladyland*, doesn't send shivers down your spine, the loudspeaker is still doing something, somewhere, wrong.

Acknowledgments: When it comes to measuring loudspeakers, the following individuals have all freely shared their knowledge, experience, and opinions: Larry Archibald, Graham Bank, Paul Barton, John Bau, Martin Colloms, John Crabbe, John Dunlavy, Malcolm Omar Hawksford, Ken Hecht, J. Gordon Holt, David Inman, Ken Kantor, Don Keele, Thomas J. Norton, Wes Phillips, Doug Rife, David Smith, Bob Stuart, Jim Thiel, Kevin Voecks, and Peter Walker. Any errors, misunderstandings, or misstatements in this paper are my own. I would also like to take the opportunity to thank Martin Colloms and the late James Moir for allowing me to take part in many, many single-blind loudspeaker listening tests from 1977 though 1985. The experience was an education in how to assess the sound of a loudspeaker by listening to it.

REFERENCES

[52] D.D. Rife, "Loudspeaker 'Anechoic' Response," *DRA Labs MLSSA Reference Manual v.10.0A*, pp.153-158 (1996).

[53] "Types 4003 & 4006 Specifications," *Series 4000 Professional Microphones User's Manual* (Brüel & Kjaer 1989).

[54] M. Colloms, *High Performance Loudspeakers*, Fourth Edition, pp.147-151 (John Wiley & Sons, 1991).

[55] G. Sanders, "Manufacturer's Comment: Martin-Logan," *Stereophile* Vol.20 No.5, p.225 (1997 May).

[56] J.G. Holt, "Down With Flat!" *Stereophile* Vol.8 No.4, pp.5-8 (1985 August).

[57] M. Colloms, *High Performance Loudspeakers*, Fourth Edition, pp.317-319 (John Wiley & Sons, 1991).

[58] D.B. Keele Jr., " Equipment Profile: JBL L7 loudspeaker," *Audio* Vol.45 No.12 (1992 December).

[59] J. Atkinson, "JA on Loudspeakers," *Stereophile* Vol.14 No.10, pp.205-212 (1991 October).

[60] J. English, "Postscript," *Stereophile* Vol.16 No.1, pp.237-238 (1993 January).

[61] J. Atkinson, "Celestion DLP600 loudspeaker processor," *Stereophile* Vol.15 No.8, pp.145-158 (1992 August).

[62] R. Deutsch, J. Atkinson, "Audio Physic DSP digital speaker equalizer," *Stereophile* Vol.17 No.11, pp.169-174 (1994 November).

[63] J.G. Holt, J. Atkinson, T.J. Norton, "Meridian Digital Home Theatre," *Stereophile* Vol.18 No.6, pp.73-85 (1995 June).

[64] D.B. Keele Jr., "Low-Frequency Loudspeaker Assessment by Nearfield Sound-Pressure Measurement," *J. Aud. Eng. Soc.* Vol.22, p.154 (1974 March), reprinted in *Loudspeakers Vol.I*, pp.330-338, edited by R.E. Cooke (Audio Engineering Society, 1978).

[65] C.J. Struck, S.F. Temme, "Simulated Free Field Measurements," *J. Aud. Eng. Soc.* Vol.42, pp.467-482 (1994 June).

[66] D.D. Rife, "Loudspeaker Near-Field Measurements," *DRA Labs MLSSA Reference Manual v.10.0A*, pp.176-180 (1996).

[67] F.E. Toole, "Loudspeaker Measurements and their Relationship to Listener Preferences: Part 2," *J. Aud. Eng. Soc.* Vol.34, p.327 (1986 May).

[68] D.D. Rife, "Loudspeaker Polar Response," *DRA Labs MLSSA Reference Manual v.10.0A*, pp.180-185 (1996).

[69] R.F. Allison, R. Berkovitz, "The Sound Field in Home Listening Rooms (Part 1)," *J. Aud. Eng. Soc.* Vol.20, p.459-469 (1972 June), reprinted in *Loudspeakers Vol.I*, pp.330-338, edited by R.E. Cooke (Audio Engineering Society, 1978).

[70] R.F. Allison, "The Influence of Room Boundaries on Loudspeaker Power Output," *J. Aud. Eng. Soc.* Vol.22, p.314 (1974 May), reprinted in *Loudspeakers Vol.I*, pp.339-345, edited by R.E. Cooke (Audio Engineering Society, 1978).

[71] T. Holl, "Speakers, Rooms and Responses," *Hi-Fi News & Record Review* Vol.28 No.5, pp.40-43 (1983 May).

[72] D.R. Moran, "Down With Stands: Quantifying the Allison Effect," presented at the 91st Convention of the Audio Engineering Society (1991 October). Preprint 3206.

[73] S.E. Olive, P.L. Schuck, S.L. Sally, M.E. Bonneville, "The Effects of Loudspeaker Placement on Listener Preference Ratings," *J. Aud. Eng. Soc.* Vol.42, pp.651-669 (1994 September).

[74] S. Bech, "Perception of Timbre of Reproduced Sound in Small Rooms: Influence of Room and Loudspeaker Position," *J. Aud. Eng. Soc.* Vol.42, pp.999-1007 (1994 December).

[75] F.E. Toole, "Loudspeaker Measurements and their Relationship to Listener Preferences: Part 2," *J. Aud. Eng. Soc.*

Vol.34, pp.327 (1986 May).

[76] M. Colloms, *High Performance Loudspeakers*, Fourth Edition, pp.317-319 (John Wiley & Sons, 1991).

[77] V. Dickason, *Loudspeaker Recipes Book 1*, pp.viii-x (Audio Amateur Press, 1994).

[78] F.E. Toole, "Loudspeaker Measurements and their Relationship to Listener Preferences: Part 1," *J. Aud. Eng. Soc.* Vol.34, pp.227-235 (1986 April).

[79] F.E. Toole, "Loudspeaker Measurements and their Relationship to Listener Preferences: Part 2," *J. Aud. Eng. Soc.* Vol.34, pp.323-348 (1986 May).

[80] J. Atkinson, "JA on Loudspeakers," *Stereophile* Vol.14 No.10, pp.205-212 (1991 October).