



HIGH EMOTION AUDIO®  
HIGH EMOTION AUDIO, LLC.  
NASHVILLE, TN 37205 USA

# HIGH EMOTION AUDIO

## What is it and how do we achieve it?

*D. Michael Shields and J. Craig Oxford  
High Emotion Audio, LLC.*

We design and build loudspeakers that can be heard in the heart as well as the ears. We believe our loudspeakers satisfy in ways that no other loudspeaker can, and we are not alone in our thinking. Musicians, music lovers, reviewers and audiophiles agree and, interestingly, many designers and manufacturers who once criticized our work, are now attempting imitations of it.

So what is it, exactly, that we are doing?

We are providing room correction the natural way: We are creating systems that deliver absolutely uniform dispersion of sound independent of frequency throughout the entire operating bandwidth. Our technical design objective is flat power delivery throughout the room. Our central design objective is to deliver the full emotional experience of the music itself. Also, we are liberating the music from the technology by recognizing that hearing and feeling, not measurement, are the final arbiters.

We realized early in our research that human hearing judges things that aren't easy or even possible to measure by technical means, but there is still a need to establish objective technical competence.

Our basic approach to loudspeaker design is, in essence, the inverse of the norm. Rather than just create loudspeakers that do a good job of performing in the usual scenarios of technical measurement (and then hoping for the best in the listening room), we first applied an even loftier set of standards focusing on the cardinal attributes of human hearing.

- Frequency range
- Dynamic range
- Temporal resolution
- Spatial discrimination

Simply put, how does the human ear perceive these attributes in a natural environment; and considering that it is actually the complex interrelationships of these attributes that create and deliver the natural and realistic experience of listening to live music, what does that mean in our design of loudspeakers?

## **Frequency range**

It has been long established that the frequency range of human hearing for stationary tones is about 16 Hz to 16 kHz. In musical terms, this is about 10 octaves. The very young can sometimes hear beyond 16 kHz and there are reports of audible sensation (as opposed to vibratory sensation) as low as 8 Hz.

A useful indicator is that if you can hear the squeal of the horizontal flyback in an analog TV set, you can hear 15.75 kHz. If you are more than 30 years old and can't hear it, don't feel bad; most people your age can't either. Women tend to hear slightly higher frequencies than men. Lucky them.

There are also reports of measurable physiological responses to auditory stimuli up to about 30 kHz – even though they are not necessarily accompanied by conscious perception. This may be important (or not).

One should not confuse frequency (which is objectively measurable) with pitch (which is a percept). The unit of frequency is the Hertz (1/second). The unit of pitch is the mel. Perceived pitch involves the harmonic structure and duration of a sound where frequency is simply a measure of the fundamental cyclicity. In order to achieve clarity of pitch, not only is frequency response important, but also distortion and temporal response.

All PipeDreams® loudspeakers are three-way systems that operate over the range from slightly below 20 Hz to slightly above 20 kHz. In general, our crossover networks are optimized for flat power in consideration of the wide dispersion of the PipeDreams configuration.

Our smaller monitor systems have similar extension of the high end. In the bass, the two smallest monitors (S5 and S7) are sufficient for many types of music if the room is small and speaker placement is optimized, but S5s are really intended to be used with subwoofers. Using a crossover frequency around 85 Hz, they will integrate seamlessly with most subwoofers. The S7 delivers bass extension to below 35 Hz while still being capable of

substantial output.

Achieving adequate and smooth frequency range in a loudspeaker to a fixed axial position of the listener, is relatively easy. Achieving it over a broad range of listening positions is not. Achieving it without compromise to the qualities affecting the other important attributes of hearing is more difficult still. This is the area where we excel.

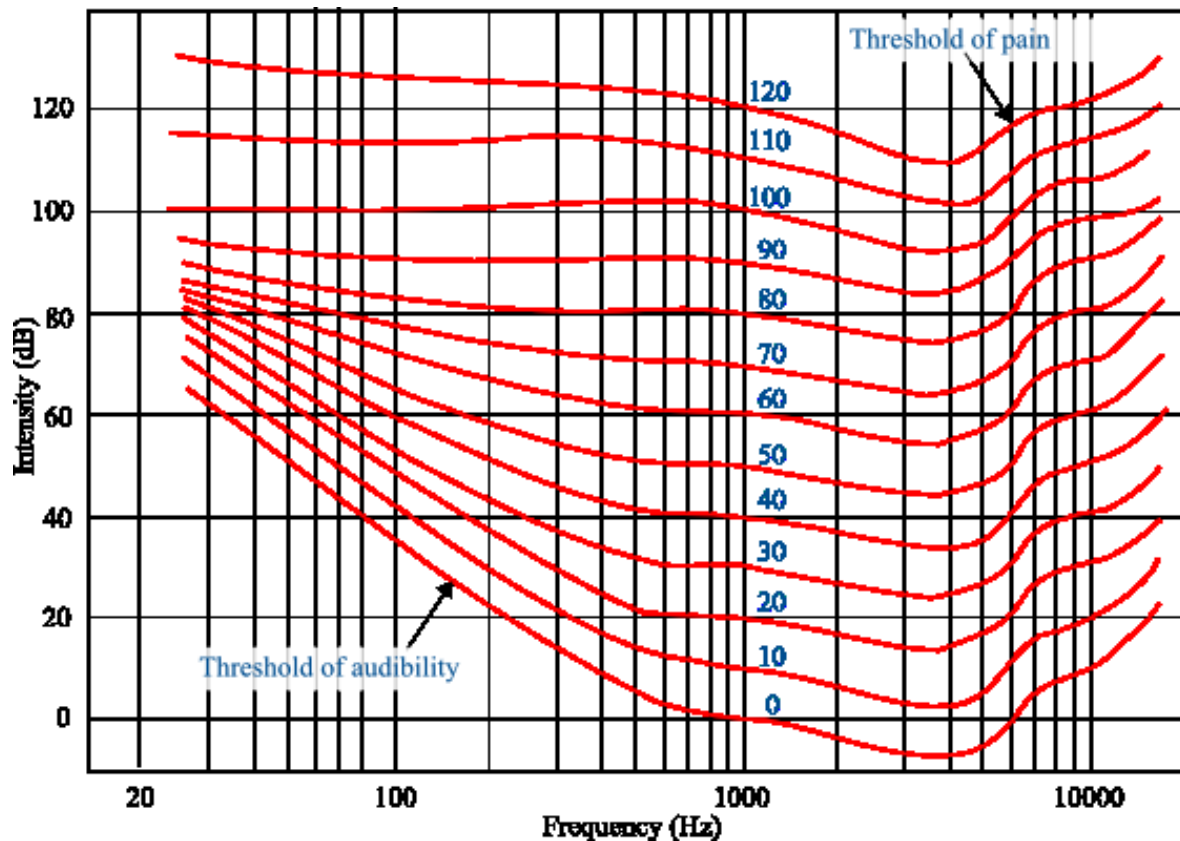
### **Dynamic range**

The dynamic range of hearing is the range from the softest perceptible sound to the loudest tolerable sound. In the middle frequencies (around 1 kHz) it is about 120 decibels (dB). It must be remembered that dB is a logarithmic measure of ratio. In terms of acoustic power, 120dB means that the loudest tolerable sound is about one trillion times ( $10^{12}$ ) greater than the softest perceptible sound. A symphony orchestra has an actual dynamic range of about 70dB, or a power ratio of 10 million to one ( $10^7$ ).

At low and (to a smaller extent) high frequencies the dynamic range of hearing is much less. This is because the bottom comes up – i.e., the minimum perceptible sound at 50 Hz will be measurably much higher in level than at 1 kHz.

The most well known examination of this subject was conducted by Fletcher and Munson of Bell Laboratories in the early days of audio technology. They collected a statistically significant body of data by measuring the reaction to pure tones of thousands of people. An important result of this work was the so-called “Fletcher-Munson” curves, more properly called the Contours of Equal Loudness. This work has been repeated with variations by other researchers in other countries with substantially similar results.

The contours of equal loudness are shown on the next page:



The “crowding” of the curves at low frequencies means that care must be taken in sound reproduction to avoid deficient low frequencies. Because small decreases in the intensity at low frequencies produce large decreases in perceived loudness.

Reproduction of the sound of a symphony orchestra at levels comparable to a live performance requires a sound pressure level (SPL) up to about 105 dB SPL at the listening position. Much popular music in certain genres requires higher SPLs. Instantaneous peak SPL may be up to 10dB greater than this.

As with frequency vs. pitch, the contours of equal loudness show that there is a variable relationship between intensity, which is objectively measurable, and loudness, which is a percept.

The essential sonic attributes of music are melody, harmony, rhythm, dynamic contrast and absolute loudness. Dynamic contrast and absolute loudness are tightly related. It is not enough to simply be able to make a loud noise with a loudspeaker system. The accurate rendering of dynamic contrast requires the accurate preservation of different levels of absolute SPL.

In most loudspeaker systems, (especially ones which look big but are really just a mini-monitor co-packaged with a subwoofer), the problem is one of **dynamic compression**: as electrical energy fed to the loudspeaker increases, there is not a corresponding increase in the acoustical energy that comes out.. This occurs mainly for two reasons. First, the driving force in a speaker is never completely linear with displacement. As the cones are required to move farther to reproduce a louder sound two things happen: 1) the elastic elements of the suspension start to “tighten up” and reach their limits 2) the driving force starts to diminish and may become assymetrical. The result of these problems is an instantaneous error which simultaneously distorts the waveform and reduces the output. This type of dynamic compression is called **displacement compression**.

Second, high input power to the driver(s) causes the voice coil(s) to get hot. Since the voice coils are made with either copper or aluminum wire, both of which have a high temperature coefficient of resistance, their resistance increases when they heat up. Almost without exception, power amplifiers are constant-voltage sources. This means that for any given input voltage, the amplifier’s output voltage will be a certain value regardless of the resistance at the output (the loudspeaker). As the voice coil gets hotter and its resistance rises, the power it can absorb falls – yet the amplifier output voltage remains the same. Another type of dynamic compression, this is called **thermal compression**. It occurs at a slower rate than displacement compression and does not result in much waveform distortion, but rather in a slow decrease in the sensitivity of the loudspeaker. Thermal compression is a greater problem with loudspeakers of low sensitivity because they require more power in the first place.

Both of these mechanisms of dynamic compression are damaging to the perception of dynamic range because they have no counterpart in natural sounds. Other types of distortion mimic natural acoustic phenomena. Harmonic and intermodulation distortions are not unlike natural modifications of sounds that can occur. This is important because when our hearing is presented with types of distortion which do not occur in natural sounds, those distortions are likely to be more offensive or at least more noticeable.

There are a few ways to reduce or eliminate dynamic compression. Since it is due mainly to large diaphragm movement or high input power or both, the obvious solution would be to improve the sensitivity and reduce the diaphragm motion. One age-old way of accom-

plishing this is to use a **horn**. But even though the horn can solve the dynamic range problem, it does so by introducing numerous other difficulties – directivity, ragged frequency response, total lack of temporal resolution, etc. – making it a poor choice.

Another way to eliminate dynamic compression is to make the diaphragm larger. This reduces the excursion of the driver in proportion to the increase in its physical area. This approach is found in big full-range electrostatic loudspeakers and in planar magnetic loudspeakers. Like the horn, this solution is accompanied by increased directivity. Also, and this is very important, both electrostatic and planar magnetic loudspeakers operate with stationary structures directly in front of (and behind) the diaphragm. All claims to the contrary notwithstanding, these structures are not acoustically transparent. In some cases they produce cavity resonances which are so severe they have to be trapped in the crossover network. Also, these types of large area loudspeakers are almost invariably **dipoles**, i.e. they radiate in opposite polarity from the back. Dipoles are hard to place because the placement which achieves passable bass response is almost never the same as the one which gives good image rendition.

Then there is “our way” – a straightforward solution to dynamic range that simultaneously addresses diaphragm motion, directivity, sensitivity and power handling:

*Extend the radiator in one dimension only, i.e. a **line source***

A true line source is not readily achieved. There are some attempts with “ribbon” loudspeakers and narrow planar magnetic transducers, but these are either fragile or plagued with the problem of fixed obstacles on either side of the diaphragm.

A sufficient approximation to a “true” line source is accomplished with a multitude of identical dynamic drivers that are small enough and closely spaced enough to coalesce into an apparently continuous line. Further, this allows us to take advantage of the fact that this is the most mature and highly developed loudspeaker transducer technology. There are no obstacles in between the diaphragm and the listener, and very high levels of driving force are possible.

The use of many drivers means no single driver ever receives very much power, so thermal compression is eliminated. The same is true of diaphragm motion, so displacement compression is similarly eliminated. Because the individual drivers are virtually always operating in the small-signal regime, all the non-linearity and distortion are held to very low levels.

We see from various manufacturers heroic design techniques applied to individual drivers. This is usually because such drivers are being asked to do too much. Our line array directly circumvents this problem avoiding the need for tortured driver design. An important result of this is our use of soft dome tweeters which, while not ideal at extremely high drive levels, are surpassingly transparent when not overdriven. In a PipeDreams array, they cannot be practically overdriven.

The direct solution of the dynamic range problem confers ***unique listening attributes*** which must really be heard to be appreciated. First, and most unusual, the “surface loudness” is low. This is noticed as you approach the speaker and it doesn’t get a lot louder. Rather, there is an impression of listening “through” the speaker. Conversely as you move away, the sound doesn’t become “small.” But the most rewarding attributes are the qualities of

- effortlessness
- power, and
- scale

in the most dynamic kinds of music.

So, to refer back to the dynamic range of hearing discussed earlier, the PipeDreams Loudspeaker systems properly applied are capable of reproducing, in a normal listening environment, the entire dynamic range of all music. Continuous SPL of over 110 dB is possible with low distortion, high reliability and reasonable amplifier power, even in very large domestic rooms.

The vertical line source, because it remains small in the horizontal dimension, retains wide dispersion in the horizontal plane. The long vertical dimension causes the dispersion in the vertical plane to be limited to the height of the line. As long as the line is tall enough to encompass all listening heights, this is not a limitation.

### **Temporal resolution**

This is a very inadequately considered aspect of human hearing, largely because most speaker companies don’t want to deal with the expense of addressing it. Much auditory research concerns itself with pure, continuous tones. This leads to the frequency range of hearing being described as extending to about 16-20 kHz as discussed above. Yet we are able to hear

transient sounds whose time-domain representation indicates (via the Fourier transform) a frequency content that extends beyond “audibility.”

Very short transients, on the order of 10 microseconds (usec, or ten millionths of a second), are less audible than continuous tones of the same amplitude, but they are still audible. There is anecdotal evidence that editing a 44.1 ksp/s digital recording by randomly removing single samples causes some listeners to object to the result. A single sample represents a missing time of only 22.7 usec. There have been several studies directed to determining whether listeners can perceive differences (or improvements) from bandwidth extensions beyond 20 kHz in recording/reproducing systems. The general finding is ambiguous.

There is, however, far more to temporal response than bandwidth. For sounds composed of many frequencies simultaneously (virtually all natural sounds), it is important to preserve the simultaneity of the various components. They must all emerge from the recording/reproducing system at the same relative times that they went in. Failure to do so is referred to as **group delay distortion**. The ear is fairly sensitive to group delay distortion but the exact thresholds at various frequencies have not been determined. There is a criterion established by Blauert & Laws, but this relates to the impairment of speech. The perceptual thresholds may be quite different for other types of signals.

The meta-linguistic descriptor of loudspeakers as being “fast” or “slow” appears to be related to group delay distortions (assuming adequate bandwidth is available).

Temporal resolution in loudspeaker terms is the ability to render the fine structure of sounds. This is commonly called transient response. In the time domain, it is a measure of how quickly the loudspeaker can respond to changes in the input signal, not only with respect to the onset of a sound but also its cessation. In the frequency domain, it is a measure of bandwidth and group delay distortion as discussed here. The frequency and time domains are not exclusive to each other but are rather alternative views of the same thing. The Fourier transform allows the view in one domain to be converted to the other.

Accurate response to a rapidly changing signal requires the diaphragm to be accelerated (and decelerated) rapidly. Since acceleration equals force divided by mass (Newton’s Second Law), it follows that for a given moving mass, acceleration will be proportional to available force. The mass is defined by the weight of the moving parts (voice coil, diaphragm,

suspension) as well as the air load on the diaphragm. The force is provided by the motor (magnetic circuit and voice coil) due to current from the amplifier passing through it.

Using many drivers in an array does not alter the ratio of force to mass. Therefore, a line array does NOT present an opportunity to use low-performance drivers. What happens is this: for a given acoustic output, the displacement per driver is reduced in proportion to the number of drivers (forgetting for a moment that there is also an improvement in radiation resistance). This reduces diaphragm velocity by the same amount. Since kinetic energy is proportional to the square of velocity, the kinetic energy that must be imparted to each diaphragm and then removed from it is tremendously reduced (by the square of the number of drivers). The ability of the drivers to respond to transient inputs is correspondingly improved.

Loudspeaker drivers are minimum phase devices. This means that the phase response is directly predictable from the amplitude response. If the amplitude response is smooth and flat, so will be the phase response which is equivalent to low group delay distortion as discussed earlier. It then remains to design the crossover so that the transitions from woofer to midrange and midrange to tweeter also exhibit low group delay variations.

The combination of drivers with good force/mass ratios, proper crossover design and large arrays leads to subjectively amazing qualities of

- clarity, and
- speed

that are encountered to this degree in no other loudspeakers.

### **Spatial discrimination**

This is the human ability to locate the origin of sounds. It is a very complicated mechanism that is becoming fairly well characterized if not yet well understood. In general, the differentiation of direction in azimuth (the horizontal plane) is related to inter-aural time differences below about 700 Hz and to inter-aural intensity differences above about 2 kHz. The region between seems to use both mechanisms. Differentiation in elevation is due to changes in the frequency response at the ear canal from constructive or destructive reflections from the outer ear structure, the pinna. Propagation of sounds around the head for azimuths other than zero degrees (straight ahead) is different for the two ears. These differences consti-

tute the head related transfer functions (HRTF) which can be exploited in various audio processes to create virtual locations of sounds, but this technique only works reliably in conditions of very low reflection.

A great deal has been learned about spatial hearing by studying listener responses in anechoic conditions, i.e. an acoustically non-reflective environment. Yet we live all day, every day, in environments that are acoustically uncontrolled and almost always fairly reflective. Our spatial discrimination operates quite well in these situations. Indeed, it is a requirement for survival. Presumably our ear-brain system understands how to factor the local environment so it does not interfere with understanding what we hear.

In loudspeaker terms, spatial discrimination is usually referred to as *imaging*. The general sense in which the word is used, suggests the ability to locate sounds, but there is more to it than that. Lateralization (across the soundstage) is just one dimension of imaging. Localization (depth) is another. Yet another is the sense of *space* surrounding each instrument or voice creating the music and the related sense of *scale*.

Not all these aspects of musical presentation are equally well encoded in all recordings, so discussion of these attributes presupposes a well made recording.

The attributes required for temporal resolution discussed above are also important keys to preservation of the spatial attributes. But now we come to one of the most important and difficult matters in loudspeaker system design: the interaction between the loudspeakers and the room.

Most loudspeakers are optimized for flat on-axis response along a particular line perpendicular to the front of the unit. In general, some attention is paid to the response up to 30° off-axis horizontally and maybe 10° vertically. Beyond that very little attention is paid. The assumption is that the far off-axis response doesn't matter. This is **incorrect** for at least two reasons: 1) the far off-axis sound is a strong contributor to the reflected soundfield in the room (correctly called reverberation only in very large rooms) and 2) the far off-axis frequency response of the loudspeaker is likely to be very irregular. The irregularity of the off-axis response is usually due to diffraction around the edges of the box. The diffraction angle is wavelength dependent and the resulting constructive and destructive interferences produce very irregular response. In the case of horns, diffractions near the mouth have the same effect.

Dipole loudspeakers are relatively free of this particular problem

Our hearing allows us to have excellent spatial discrimination even in the presence of reflections of the sound we are trying to discriminate. So presumably, our ear-brain system is able to use the reflected information, or possibly disregard it. But if the reflections are initiated by a severely deformed version of the direct sound then our ability to factor out the reflections seems to become impaired. This is what happens when bad off-axis response excites reflections.

The usual proposal to remedy this situation is to eliminate the reflections. This is accomplished by deadening the room with absorbent material on all its surfaces. If the loudspeakers required this room treatment in the first place, then for the same reason they have poor off-axis response and they will also have a narrow frontal coverage angle. So the result is a narrow listening area in an unpleasant room. This is such a solipsistic solution you might as well buy a good pair of headphones and listen all alone and forget about the room. This situation is the result of treating the symptom rather than the disease. The disease in this case is conceptually defective loudspeakers.

Listening to music is a fundamentally social activity that is most pleasantly conducted in a comfortably furnished room with many appropriate uses. To this end we are careful to make all our loudspeaker designs interact with the room in a smooth and predictable way. The important results of this approach are:

- Broad useful listening area (big sweet spot)
- Easy placement of the speakers
- Superior imaging
- Aesthetic preservation of the room

An important consequence of this emphasis is that the acoustic *of the recording* is clearly intelligible because the local acoustic of the listening space is not suppressed but rather is correctly "illuminated."

Another aspect of the rendering of space is the sense of the size of the *space*. Where a very large space, as in a cathedral, is to be rendered there is a requirement to preserve the extreme low frequencies. The presence of stereophonic deep bass (no mono subwoofers allowed) contributes strongly to the sense of "vastness." All our subwoofer systems maintain

uniform response to 20 Hz with significantly high output.

**Scale** is an important consideration as well. The Pipedreams are very tall loudspeakers, yet the sense of how big the performer is, is remarkably correct. The singer's mouth is not perceived to be six feet tall or wide. This seems to be a counterintuitive result, but is actually due to the smooth interaction between the speakers and the room. The ear-brain system is not being deceived by spectrally distorted reflections.

### **Smaller solutions**

Most of the foregoing discussion has made the case for the line-array solution to the challenges presented by the way we hear. While a Pipedreams™ system is the most comprehensive solution we know of, solutions that are smaller in size and lower in cost are possible (and indeed, available) if we understand which compromises exact the smallest penalty.

Economics aside, the main reason for pursuing a smaller solution is to use it in a smaller room. This leads directly to a lower requirement for acoustic power output which facilitates a smaller radiating area. Hence the possibility of fewer drivers.

We especially want to preserve the directional (or should we say, "non-directional") attributes which allow smooth interaction with the room and thus preserve the qualities of imaging and timbral correctness associated with proper dispersion.

We also want to preserve the temporal resolution so we will have to use drivers that permit it. The sense of space and scale may have to be sacrificed somewhat, but if we enable (or require) the use of subwoofers, an important element of it can be retained.

To that end, we have developed small systems – one that is intended to be used with subwoofers (the S5s) and one that is a full-range stand-alone system with bass extension to 35 Hz (the S7s).

Actually, the S5s can be quite satisfying without subwoofers (depending on your taste for deep or powerful bass) when positioned in the corners of a small listening room.

These systems use our proprietary VAT (Variable Aperture Transducer™). This is a very unusual transducer in which a freely propagating, bending wave is imparted to a curved, flexible membrane by the motor. It uses the inherent properties of the material rather than fighting them. The propagation of the wave causes the radiation to be inherently diffuse. The

result of this is an extraordinarily wide dispersion with no specular “beam.” Further, because the area of the membrane is quite large, the surface loudness is low as in the PipeDreams line array.

The unique attribute of our VAT is its musicality. Virtually without exception, everyone who hears it comments on this. Also, the rendering of scale is astonishingly good. In a side-by-side demonstration with a much larger PipeDreams line-array system, and using the same subwoofers, listeners frequently mistake the small monitors for the PipeDreams.

We think this strongly demonstrates that by keeping the important objectives in mind, it is possible to devise completely different solutions for the same set of problems.

### **Summary & Conclusions**

We identified the principal attributes of human hearing to establish a context for the technical considerations in designing our loudspeaker systems:

We presented our approaches to addressing these requirements and why they are *superior* to others.

The validity of our position is established by the ability of our speaker systems to connect you to the music and deliver the emotional impact of the original performance. In terms of our line-array speakers, our position is further validated by the now widespread imitation of our products. While we recognize the flattery, remember that we were the first to use line-array so musically and effectively – and we are now well ahead of where we began. We use the **best** components, we have the longest experience, we know why it works and we deliver the most convincing result.

Make no mistake, these are not idle claims. Your experience listening to our products will prove that.

####

## Further reading

For those who want to know more, here's a short list of especially pertinent technical literature:

Beranek, Leo L., Acoustics, 1954 McGraw-Hill

Olson, Harry F., Elements of Acoustical Engineering, 1940, D. Van Nostrand Co.

Shorter, D.E.L., A survey of performance criteria and Design considerations for high-quality monitoring loudspeakers, Loudspeakers, Vol.1, An anthology, 1978, Audio Engineering Society Inc.

Toole, F., Loudspeaker measurements and their relationship to listener preferences, Part 1 and Part 2, Journal of the Audio Engineering Society, Vol.34, 1986

Queen, Daniel, The effect of loudspeaker radiation patterns on stereo imaging and clarity, Journal of the Audio Engineering Society, Vol.27, 1979, pp 368-379.

Hartmann, W., Localization of sound in rooms, Journal of the Acoustical Society of America, Vol. 7, pp 1380-1391

Lipschitz, S. and Vanderkooy, J., Power response of loudspeakers with non-coincident drivers - the influence of crossover design, Loudspeakers, Vol. 3, an Anthology, 1995, Audio Engineering Society Inc.

Blauert, J. and Laws, P., Group delay distortions in electrical systems, Journal of the Acoustical Society of America, Vol. 63, 1978 May, pp 1478-1483

Fletcher, H., Steinberg, J.C., Wentz, E.C., Scriven, E.O. et al, Auditory Perspective, Symposium on wire transmission of symphonic music and its reproduction in auditory perspective, Electrical Engineering, 1934 January.