Directivity in Loudspeaker Systems

1.0 Background

Constant Directivity or Controlled Directivity (CD) has become a common buzz word these days. Many loudspeaker designers and loudspeaker systems are extolling the importance of this concept, but as often happens with important concepts it is not being clearly explained nor always used properly, especially by marketing and in advertising claims. The purpose of this white-paper is to outline what directivity is and what it means in a loudspeaker system, and very importantly, how it is measured and displayed. Its importance in loudspeaker design for small rooms will be noted. The idea of an "ideal" directivity will be examined and it will be shown that CD alone is not enough, in a small room one needs a narrow directivity that is also CD.

2.0 Directivity

In its simplest definition, directivity is the characteristic of how a loudspeaker sends sounds in different directions. The figure on the top of the next page shows a typical polar diagram of what is usually referred to as directivity. It is directivity of a sorts, but a picture like this alone does not tell us very much about the distribution of the loudspeakers frequency response in angle, just its change with angle at a fixed frequency. Other frequencies can be anything, we just can't tell. What we need is a frequency response at a variety of angular positions. This data can be shown in different ways. But just looking at singular polar diagrams like this is not very useful.

There are two problems with this old standard form of displaying polar response. The first is the fact that this one is not in Log scaling (sometimes they are, sometimes not) and so it can be misleading if Log is assumed and its linear, and visa versus. At the very least they should be clearly labeled. Second, it can only show one, or a few, frequencies at a time without getting confusing, or a very broad brushed average across a wide frequency ranges must be used. Figure 1shows how a typical piston source would look in a standard polar plot as averaged in one-octave bands. Note that the axial point is normally set to be 1.0, or 0 dB, i.e. the data is normalized on-axis, thus suppressing any response variations across frequency. In this figure the colored data lines are at values of *ka* =.5 - red to *ka* = 4.0 - the orange line, with each curve doubling the value, i.e. *ka* =.5, 1.0, 2.0, 4.0. The dimensionless value *ka* is a convenient number that represents frequency - this polar diagram will apply to any circular piston when used in the dimensionless *ka* units.

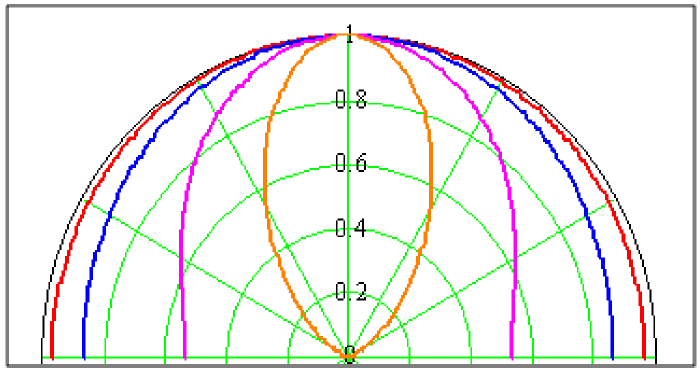


FIGURE 1. Polar directivity graph

To get the actual frequencies from Figure 1 we note that $k=2\pi f/c$ and a is the source radius. This means that $ka = 2\pi fa/c$ and if a is in centimeters then the frequency (in Hertz) is

$$f = 5460\ ka\ /\ a \quad \text{(EQ 1)}$$

Hence for a 2.0 cm radius piston, the blue line (*ka* = 1.0) would be a one octave averaged polar response centered at about 3 kHz. In this figure it should be noted that significant directivity (*ka* = 2.0) does not happen for a 2 inch piston until about 10 kHz. Below that it is pretty wide.

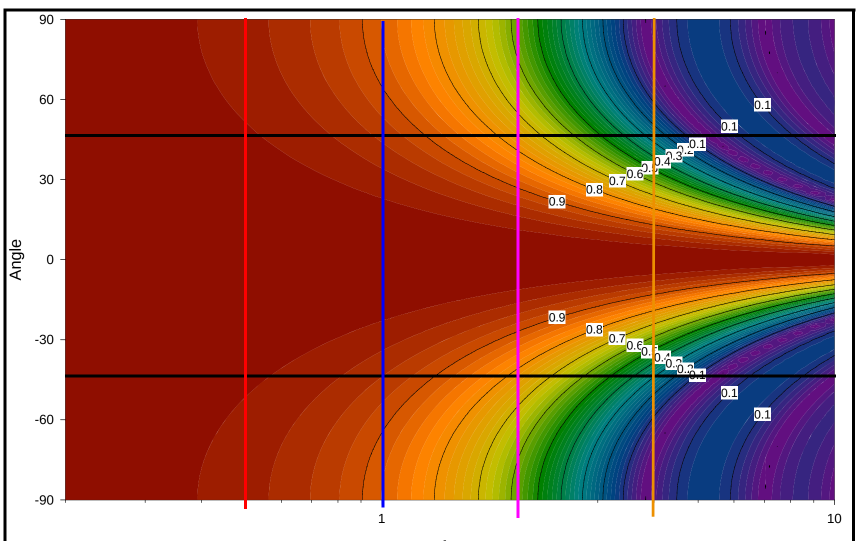


FIGURE 2. Polar map for generic piston

Figure 2 is actually the same data as shown in Figure 1, but here I have accumulated a continuous range of *ka* values on the same plot with *ka* running along the horizontal. It must be understood that Figure 1is actually contained within Figure 2 as the data along the four vertical lines (the colored vertical lines correspond to the same colors as shown in Figure 1). It is important to understand Figure 1 before moving on since the rest of this paper will be based on these type of polar maps.

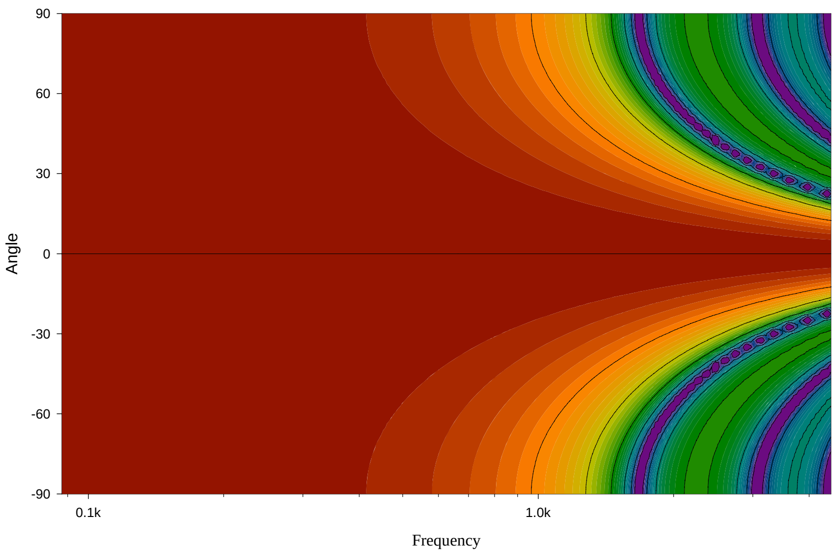


FIGURE 3. Log and frequency scaled polar map

Figure 3, shows the same data as Figure 2, but now the amplitude is scaled in dB (20 $\log_{10}$) and the *ka* value has been replace with its frequency equivalent for a piston of 12 cm radius. The dB scale is natural for any audio based data presentation because it much more closely aligns with the way we hear. Virtually all of our senses respond logarithmically. Note how the nulls are far more evident in the dB map.

In Figure 3 the contour labels have been left off to highlight an important aspect of polar maps when displayed as shown here. The contours change color every dB and the major black lines occur for every 6 dB drop in level. Carefully note the region around the first null at the extreme right side of the plot - the high frequencies - higher *ka* values. The continuous dark purple regions are nulls of the radiation and drop to very small values. Notice how the smooth valley breaks down into a series of spots at the end instead of sweeping lines. This is a clear indication that spatial aliasing has occurred and the resolution is no longer sufficient to resolve the sharp nulls. In numerical analyses involving FEA or BEM, these aliasing spots are often seen instead of sweeping lines as seen in Figure 3. This is because these calculations are usually done on fairly sparse far-field locations for the sake of calculation time. The field density in the plots shown in this paper are extremely high at 2.5° and about 1/200th of an octave in the simulated plots. (Its lower in the measured data sets.) When spots like this are seen in simulation data, the data can no longer be considered reliable since spatial aliasing has occurred.

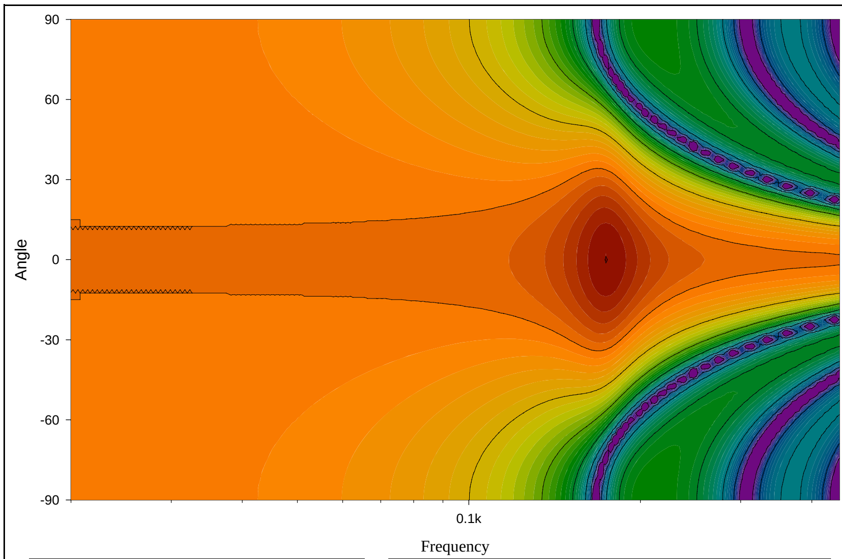


FIGURE 4. An electro-mechanical domain resonance

In the simulations done here the contours meet the edges at 90° since the model assumes an infinite baffle. but in a free-field measurements this does not occur although the response generally falls off at a very similar rate to the simulations.

As yet these figures haven't shown us much of interest other than the obvious aspect that pistons beam sound more at high frequencies - a fact that we will return to later. Figure 4 shows a resonance in either the electrical or mechanical subsystem of the loudspeaker and the effect it has on the polar map. This resonance is about 6 dB at 1.8 kHz with a Q of about 3. We can see that it is a strong perturbation on the map that lies strictly in a vertical direction, i.e. a constant frequency (iso-frequency?). This is an important point, because a resonance that is not in the acoustic domain will always show up in a polar map as a simple vertical perturbation at a constant frequency. We will see that an acoustic resonance will not have this same feature - that's good to know for reasons that we shall soon see!

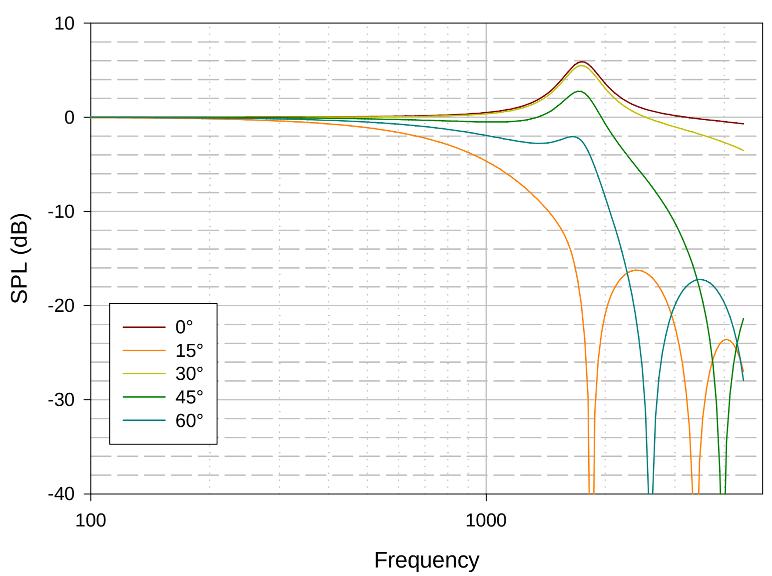


FIGURE 5. Frequency response data

Figure 5 shows a more typical view of a polar response, Note that this type of plot only shows a small fraction of the data seen in the full polar map of Figure 4.

Of far more interest is Figure 6, which shows a simulation of the acoustic aberration that is created here by the addition of a very small secondary source of coherent sound - something like a small reflection or diffraction, a cabinet edge or small object on the baffle causing a reflection, that sort of thing. Here the acoustic perturbation signal is broadband, but not very significant in level. The effect on-axis is a mere ±.5 dB ripple, but off axis the effect is much more pronounced. Note that the aberrations are not at constant frequencies (non-iso-frequency), but are more like curved sweeping lines, a clear indication that this problem is caused by something acoustic in nature - most definitely not electro-mechanical (one dimensional) in nature. Note also how the easily distinguished curves for the lobes in the previous figures have almost disappeared.

Note in Figure 4 on page 3 that if I correct the response to be perfectly flat on axis that it will automatically be corrected everywhere. On the other hand in Figure 6 there is no single correction curve to the left plot as the right "corrected" plot shows. Along the central axis everything has been corrected to be perfectly flat, but in no way, shape, or form is the problem "corrected". In fact it might even be worse. It should be obvious that electronic EQ can correct for any linear prob-

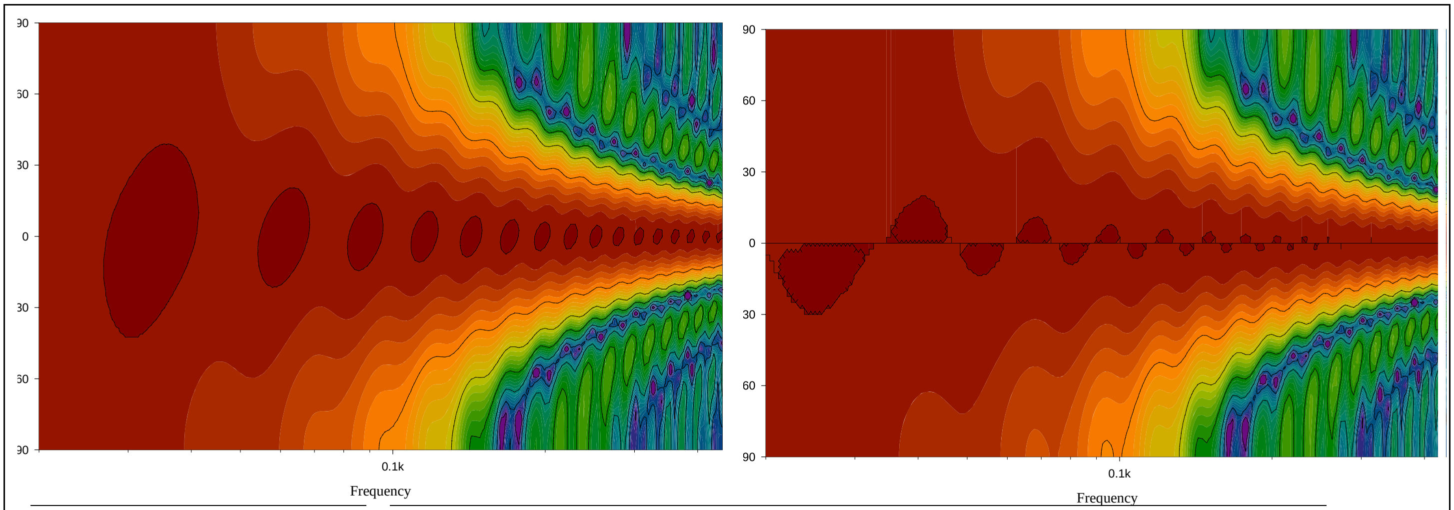


FIGURE 6. Polar map of a typical acoustic problem

lem found in a loudspeaker prior to the acoustic domain, but it does not correct for problems within the acoustic domain itself, except along a single direction, i.e. a single point. This is independent of whether or not this problem comes from the speaker or the room itself. The correction is valid only at that single point in space and nowhere else. This is a key point.

Let's now look at what a more typical acoustic problem might look like, i.e. a band-limited diffraction, sort of like an acoustic resonance. This is shown in the Figure 7 on page 5. While the effect is somewhat localized in frequency, no electronic correction can possibly fix this issue globally. An attempt to do this is shown in. In what sense the lower curve could ever be considered an improvement on top curve is not at all evident.

Hopefully by now the reader has achieved a better feeling for polar maps and how they show a great deal more information than any of the more standard types of performance plots. Details that appear to be quite important in a global sense (as opposed to the standard single axis sense) are readily seen in the polar map, but not evident at all in a single axial curve and not even that apparent in a multiple angle set of data like Figure 5. Assuming that the reader is now comfortable with this new way of showing polar data - I want to take a step backwards and talk about what might be considered as an *ideal polar map*, what we would want to see if possible.

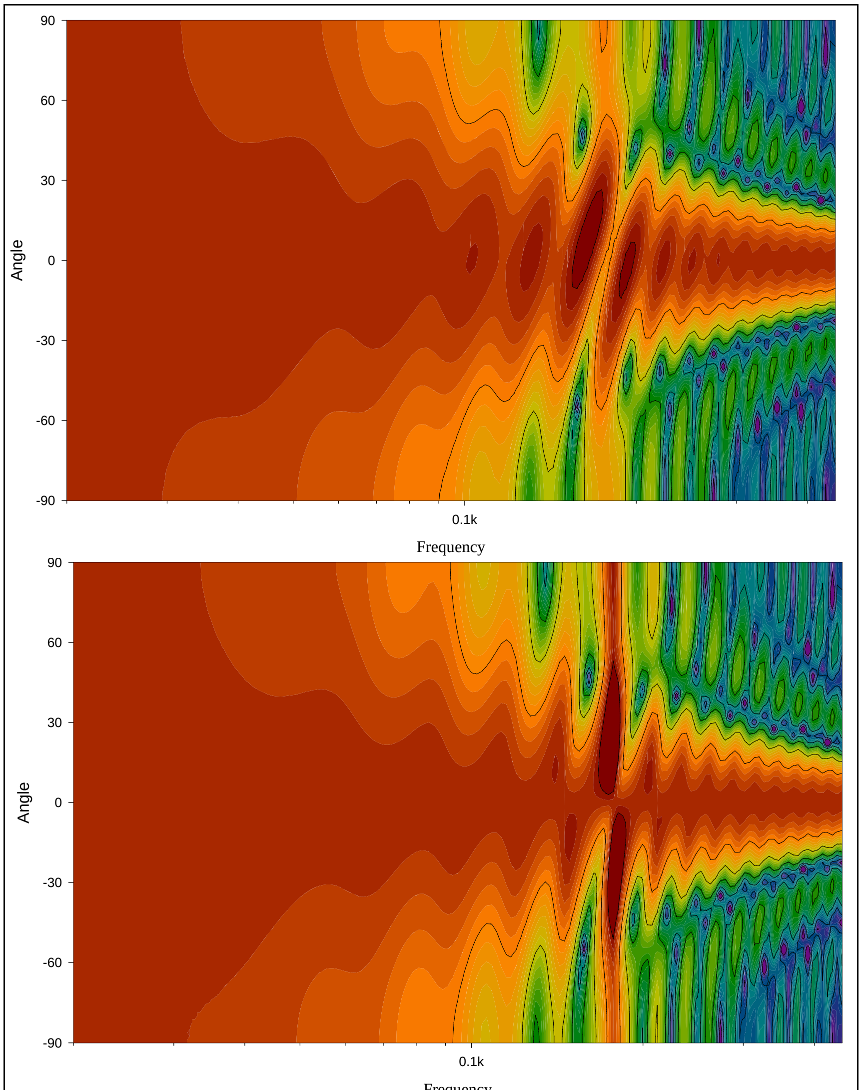


FIGURE 7. An acoustic aberration and its correction

It is well understood that the ideal frequency response is basically a flat curve, although a lot of evidence suggest a small drop in the respond at HFs (mostly room averaged power response) in smaller rooms. We all know that this can only be achieved within some limits, but the discussion above shows that there is a lot more to the story than just a single directional response. Some very important (indeed, critical) aspects of a loudspeakers sound quality is not at all apparent from an axial curve or any single direction curve and even several polar directions don't yield as much information as would generally like to see. I've shown how to look at a far more complete set of data can be shown by using a topographical plot called a polar map. And we have seen what kinds of data this type of data presentation shows, now its time to figure out what the ideal might be and what kinds of physical limitations to this ideal that we might expect in the real world.

3.0 The Idea of an Ideal Polar Map

With the assumed familiarization with polar maps developed in the previous section and the kinds of things that can be seen in them, the rest will be easy. First I want to talk about what a playback system in a real room should do. This will by necessity be rather brief since this is a vast subject. Basically, I'm going state my beliefs about what the ideal polar pattern is and why, but I won't give an elaborate proof for this position, although I believe that I could. However, for brevity (and the sanity of the reader), I will simply state my position as my opinion and leave it there. In any case, it should be appreciated that my opinions here are based on a lot of training, experience and data.

Many say that an omni-directional loudspeaker response is the ideal, but I do not accept this for a small room. That's because the Very Early Reflections (VER) increase dramatically with this type of source because of its very wide polar pattern. Because of the omni-directional response, this type of source cannot be situated such that it does not excite every possible VER, and there simply is no way to fully absorb these VER without creating a room that is so dead as to be acoustically lifeless. As the directivity gets narrower and narrower however, it should be intuitively obvious that I can orient the speaker in such a way as to avoid the nearer boundaries and thus minimize the VER to a much greater degree. In fact, it can be shown that the higher the directivity, the greater the average Reflection Free Time (RFT) immediately following the direct sound. Ideally this would be about 20 ms, but that is never going to happen in a small room. Even 10 ms requires some significant room modifications to achieve in typical sized listening rooms. However, it is quite true that the more that can be done to extend this RFT the better the image will be, with diminishing returns beyond about 10 ms.

It is also true that VERs and a short RFT do add to the perception of *spaciousness*, which is considered to be, by many, a good thing, but just as certainly it is know that they detract from the ability to image the content of the source because of the confounding influence of the VER. Toole is a strong proponent of a large amount of VER because of its increase in the spaciousness effect. He appears to discount the negative aspects of this on imaging however. (Dr. Toole, does not make many statements about “image”, perhaps being concerned over its loose definition.) It seems to me that *if I can create “spaciousness” without increasing the VER, then I can achieve the best of both worlds*. This can in fact be done by making the room fairly reverberant, particularly behind the listener, which will improve the feeling of spaciousness through the multitude of lateral and rear reflections that will occur. However, the use of this technique with wide directivity speakers is not going to yield a very good image due to the VERs from the nearer walls, and the speakers should have a fairly constant frequency response in all directions, i.e. the power response, otherwise the sound quality will be colored. If the speakers do have a narrow directivity then the frontal VERs have been lowered, hence improving the image, and yet I can still retain the feeling of spaciousness because the room itself is fairly reverberant. In fact, if most of the reflections are coming from the sides and rear, as opposed to the front, then the quality of the spaciousness is know to improve. Hence a narrow directivity lowers the VERs and extends the RFT yielding good imaging, while the rooms high reverberation yields good spaciousness. It is most curious that this is quite often the opposite of what is done in many rooms.

Another factor in favor of the high directional response is the possibility of increasing the so-called sweet-spot of the loudspeakers through judicious use of the directivity fall-off with angle. This can be a little hard to follow, but I'll give it a try. It is easiest to start with a discussion about why there is a sweet-spot. In normal piston loudspeakers, the directivity is getting narrower at the higher frequencies and so it is usual to design the speaker such that the listener is on the axis of the speaker. This has been the sole reason why the “axial response” has carried so much weight, because it has been assumed that the listener is always on-axis. But what then happens if the listener moves off axis. In general the response from the farthest loudspeaker will fall in level and there will be an increase in the time delay relative to the nearer speaker. This “double whammy” of a lower level and a greater delay to the farther speaker, will immediately pull the image to the nearer loudspeaker. In most piston source loudspeakers the image collapses very quickly as the listener moves away from dead center.

Now, if somehow we could cause the nearer speakers level to fall while the farther speaker's level actually increased then we could partially offset the subjective effect of the image collapsing to the nearer speaker. (The farther speakers time delay will increase, but its level will increase as well.) Nothing can be done to correct for the time delay differences, but if we can make the level differences great enough then we might be able to achieve the effect that we want. Keep this highly desirable feature in mind as we look at directivity requirements on the polar maps.

The next thing that we need to talk about is Low Frequencies (LF) in small rooms. This too is a massive subject that I will only touch on here. Basically from the perspective of the main speakers, the LF directivity is not very important. This is because the nature of the sound field in the modally dominated frequency region that exists at LF in a small room is such that the directivity does not even enter into the picture. With the use of multiple subs the mains directivity at LF is not an issue at all. Certainly above some frequency it starts to become an issue, but clearly this is not 200 Hz, that's far too low, and just as clearly by 1 kHz we need to have the directivity well under control since, by this frequency, the ear is beginning to get quite sensitive to reflections and timing aberrations. Hence, somewhere between 200 Hz and 1 kHz the directivity needs to start to become narrower - down to the design angle - and then remain at that angle for the remainder of the audio bandwidth.

It might be said that lower is better as regards the point at which the directivity begins to narrow, but this comes at some cost in terms of the loudspeakers size and its actual cost to produce. The sizes and costs will approximately quadruple for every octave that one wants to push the directivity control down in frequency. A reasonable sized and priced system can achieve about 800 Hz and a large and expensive system might get down to about 500 Hz. But going below this gets extremely big and expensive as might well be anticipated with a quadrupling of these factors just going to 250 Hz. For the purposes of this paper I will assume 500 Hz as a practical limit.

With these requirements in mind we can see the first step in the progression of our ideal is shown in Figure 8. This figure shows what the polar map would look like for a system that had constant directivity - the horizontal nature of the strips - out to 45°, which then drop to zero. This is neither an achievable or desirable characteristic, but it is shown here just to give a feel for what it would look like. Physically it is impossible for the response to immediately drop to zero with angle. The higher the frequency the faster this drop can theoretically happen, but it is never immediate.

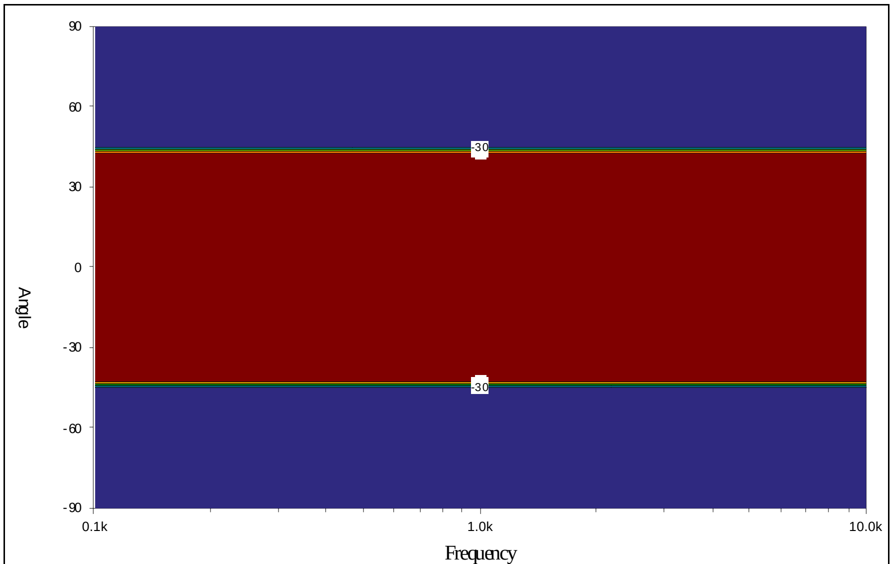


FIGURE 8. Perfect CD - unobtainable, but perfect

Taken literally the term “Constant Directivity” would mean exactly what is shown in this figure - a physical impossibility. That is why it is important to understand that there is no universal definition for what constitutes CD. This creates some problems, especially in the marketplace where marketing people hear terms like this and like the way they sound. When they learn that there actually isn’t a solid definition of it, they go wild. I have seen the qualification “CD” used for speakers that are anything but. The only way out of this dilemma is to actually show the polar response of the claimed CD loudspeaker and let the reader judge for themselves - “*Good luck with that*”.

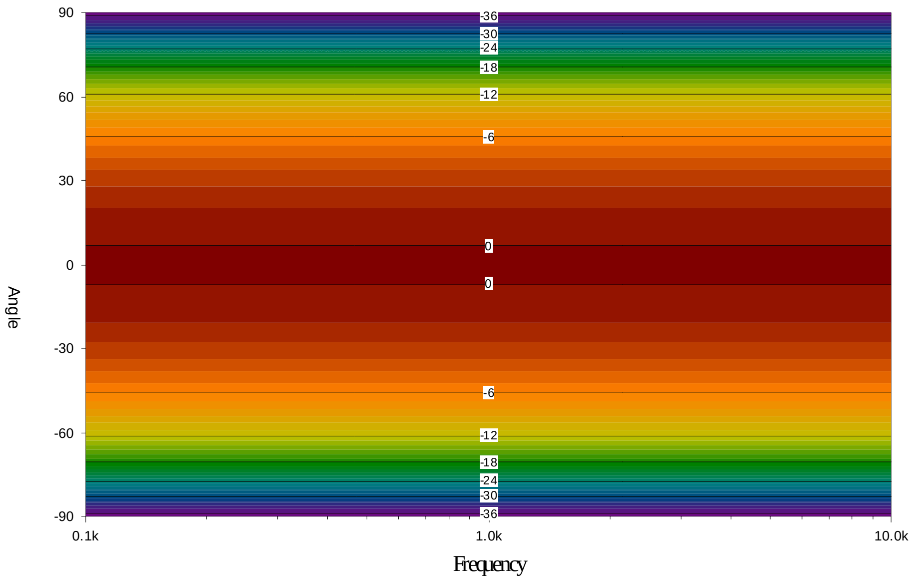
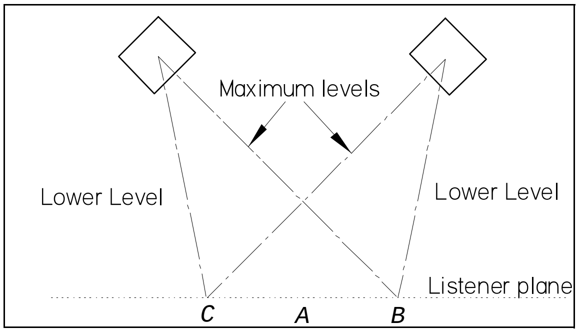


FIGURE 9. A more feasible CD polar map

A more feasible and desirable directivity pattern is shown in Figure 9. This figure shows a nearly constant directivity with frequency - it varies at low frequencies, but remember that the left side of the plot, the low frequencies, are not as important as the right side. This figure has the response down by 6 dB at 45°, and unlike Figure 8, the response falls slowly with angle as required by the physics.



Recall now our discussion of image shift in a stereo situation where we hypothesized that if the farther speaker could get louder as one moved laterally while the closer speakers level decreased, that we might be able to offset the time delay differences and maintain a fairly stable image with listening position. A glance at the drawing on the next page will show that if the central listening position is at about 22° off-axis of the main left and right speakers, i.e. toe-in, then the situation that we are looking for has been achieved. Namely, at position A, the sound from the two speakers is equal and the image is as required. At point B the closer speaker's level falls while the speaker farther away has an increase in its level. This will tend to offset the normal situation found where the image collapses to the nearer speaker whenever the listening position is not dead center. What is not so obvious is that this technique *requires* that the frequency response at A, B and C remain constant. In other words, the directional characteristics must fall with increasing angle, but it's also critical that the response remain uniform, i.e. constant with a direction change about the loudspeaker. Only a CD speaker can perform this trick correctly and as we will see piston based loudspeakers are never CD.

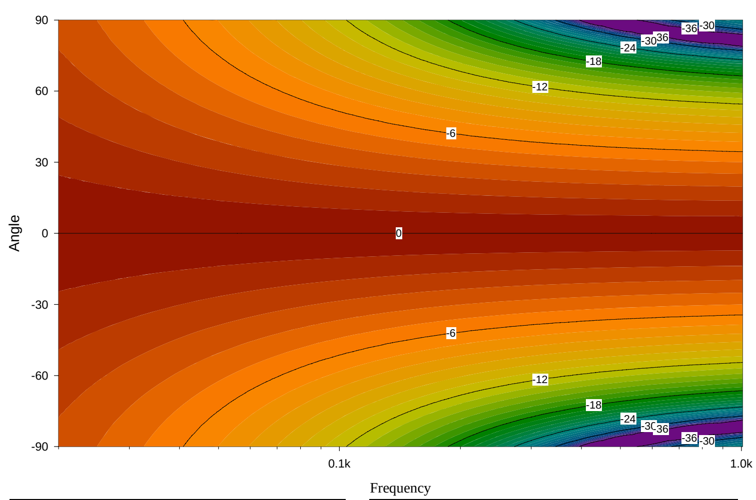


FIGURE 10. An achievable “optimum” polar map.

Finally let's consider the practicalities of a real loudspeaker, namely that at very LFs the polar response is going to widen. How much depends on the design (OB, IB, closed box, etc.) but in all cases it will get wider. Figure 10 shows a practical loudspeaker that theoretically could be achieved, albeit it would not be easy. For comparison Figure 11shows the more typical frequency response plot - various, but limited, polar angles. Note that Figure 11only shows a fraction of the data shown in Figure 10.

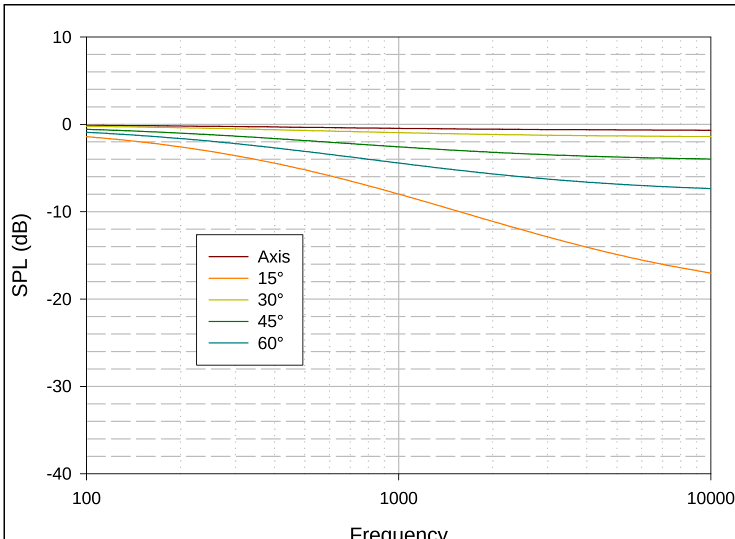


FIGURE 11.

“Optimum” Frequency response curves

4.0 Examples and discussion

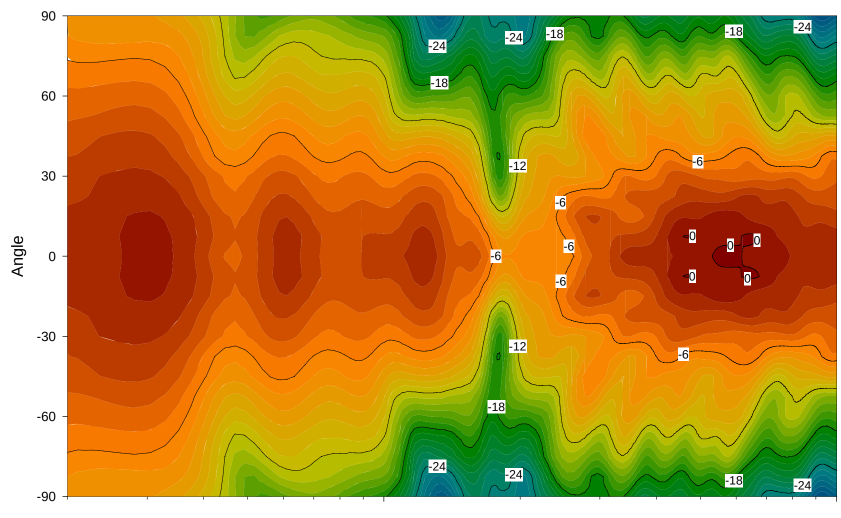
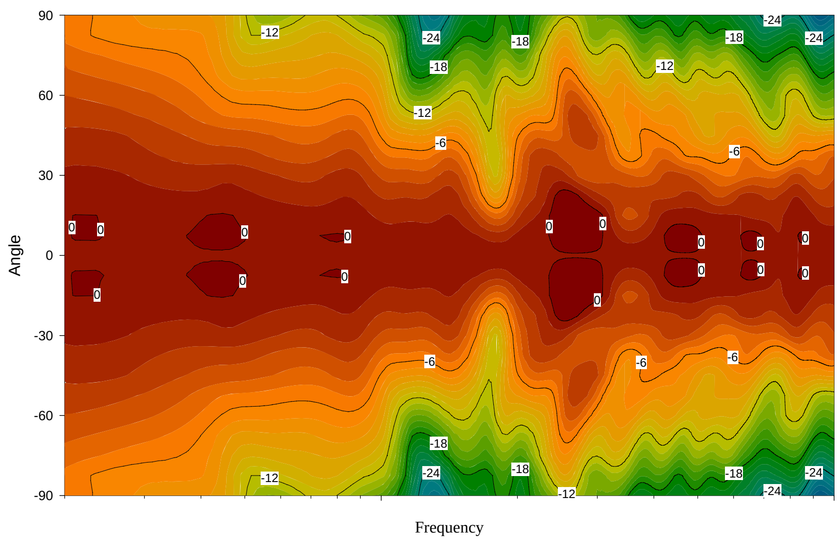


FIGURE 12. A typical 2 way speaker



Its now appropriate to look at some actual data. The examples that I will show in this section may or may not be typical, however they are all commercial products unmodified by me. In many cases their problems have obvious solutions and I will try and point those out, but in many cases no solutions are possible and the problems are cannot be treated because of the given the designs architecture. I will also point this out. I will show several examples of my own designs to show how they solve the problems, as I see them.

Figure 12 above shows a typical 2-way loudspeaker with a poorly designed crossover. The crossover is obvious at about 2 kHz where the directivity is failing. This speaker also has resonances at about 200 Hz, 500 Hz, 1.2 kHz, and a very serious one at about 6 kHz. If you look carefully at the 6 kHz resonances you will see peak lines dropping down in frequency for greater angles way from the axis. This is a clear sign of some kind of acoustic problem of diffraction, etc. Its difficult to actually see the polar responses of the individual drivers in Figure 12 because there are so many resonances in the system. In order to look at the acoustics better, and to see what the design is fully capable of, I have normalized the axial response to be flat just as a DSP system would do if placed at this point. This plot is also shown in Figure 12 above. This figure is quite interesting because now the real problems become quite evident. The woofer is behaving normally with a narrowing polar pattern and not too much in the way of acoustical problems, but then we come to the crossover. Clearly the design needs a new crossover since even EQ is not going to correct for its failings. The tweeters piston like polar response is now quite evident being truncated at the lower frequencies by the crossover. But the tweeter has some serious acoustic issue and no amount of EQ is going to help. I might keep the woofer, but the tweeter has to go and I would certainly be looking for the acoustic diffractions or whatever that has so seriously messed up the tweeters response. The system exhibits, in its normalized form, the classic and incorrect able polar response of wide at LFs, narrowing to the crossover, then wide again, then narrow again. This will be seen in each and every loudspeaker which uses piston radiators as its sources.

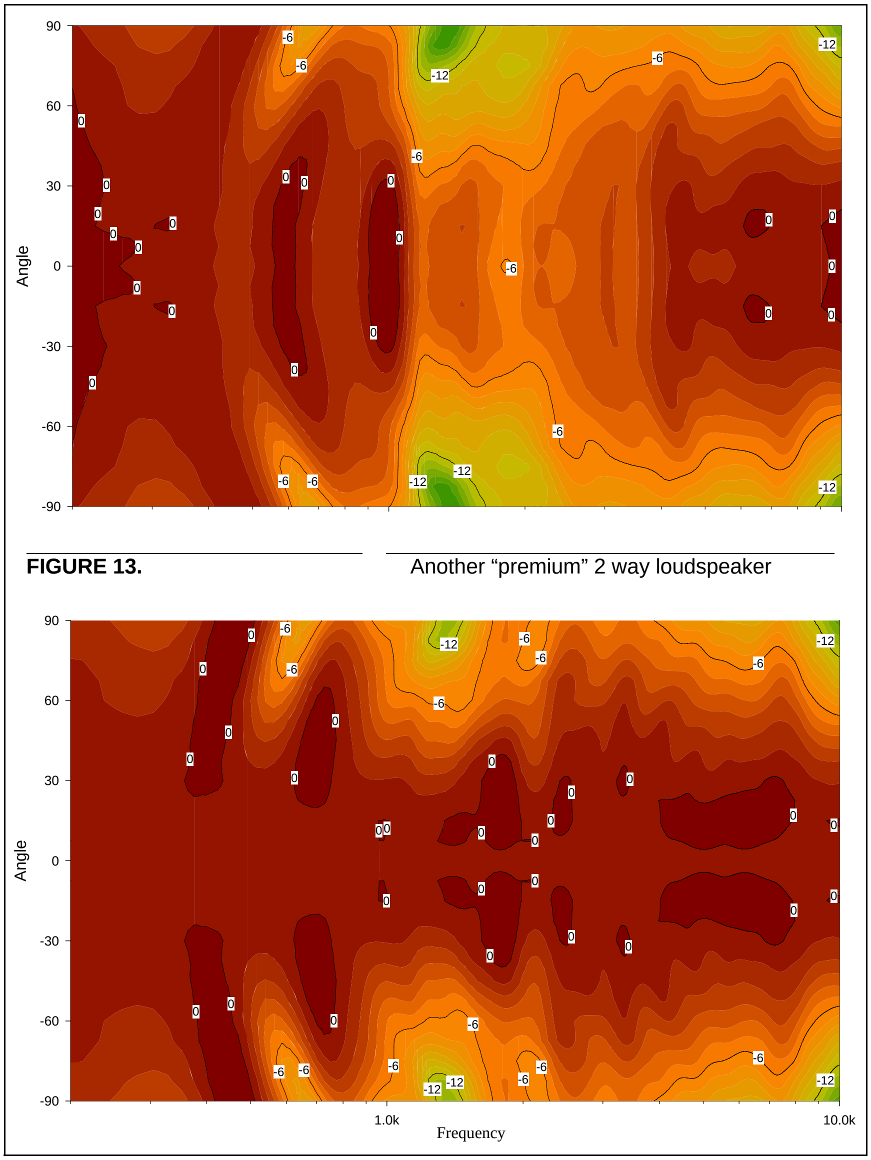


FIGURE 13. Another "premium" 2 way loudspeaker

The next example is shown in Figure 13. Not uncommon is the narrowing at the crossover, but this speaker exhibits an extremely wide polar pattern. Basically it exhibits no directivity at all. A resonance is evident 1 kHz and some acoustical aberrations at 600 Hz and above 4 kHz. A normalized plot of this response is shown in the bottom half of Figure 13. The resonance at 1 kHz is easily eliminated, but not the acoustic problems at 600 Hz and most of the tweeters range of operation. The crossover in this design is also quite poor. The classic wide-narrow-wide-narrow response is also evident. In a nutshell, while different in the details the first two examples have virtually identical major problems. An examination of the marketplace will show that virtually all of the typical two way design suffer from a nearly identical set of problem. They all sound slightly different owing to the small differences in the details, but they all have the same basic and quite major flaws in both the specific designs as well as the entire concept of the two piston two way system.

It is often believed that using coaxial mounted drivers solves the crossover acoustics problem. But this is not the case. It solves some problems, but it creates others. Figure 14 on page 12 shows the polar map along the horizontal direction for a co-axial 2-way loudspeaker system with a woofer and a horn. Figure 15 on page 12 shows the same driver along the vertical direction. The symmetry of the responses is desirable, but this speaker is in serious need of some equalization, so in Figure 16 on page 12 I have shown the normalized - perfect EQ'd response. Note that the resonance at about 600 Hz can basically be eliminated, but the one at 2.5 kHz appears to be acoustic since EQ does not make it go away, it simply pushes the peaks response around a bit. This, again is a classic example of an acoustic resonance and how EQ simple does not work on this kind of problem. You must fix acoustic problems acoustically - there is no other choice. Were it not for the acoustic resonances in the horn, Example 3 would have a good polar response. The crossover problems have been improved by the coaxial arrangement, but other problems have crept in while doing this.

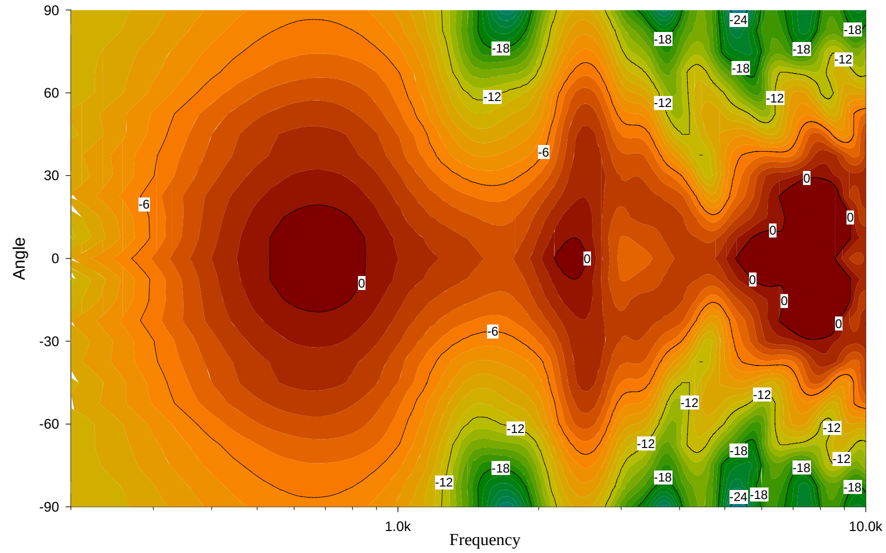


FIGURE 14. Example 3 2-way horn coax - vertical

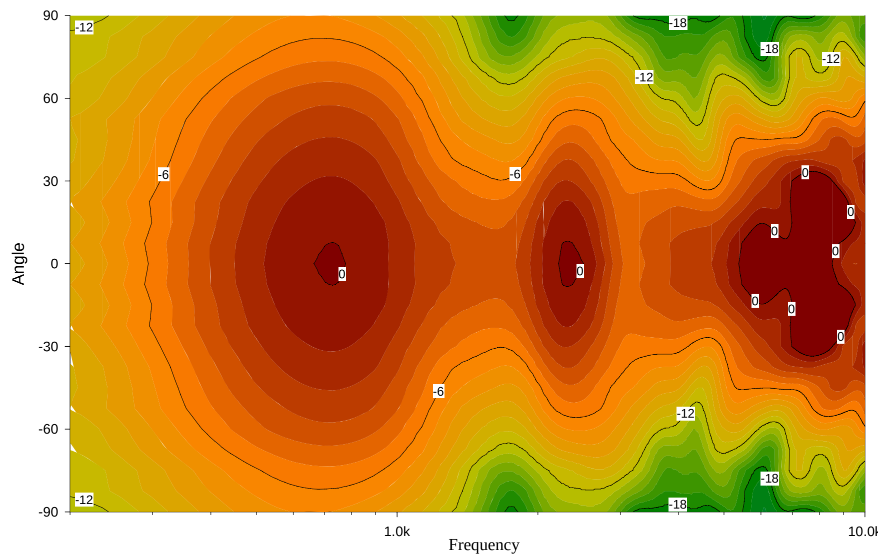


FIGURE 15. Example 3 horizontal

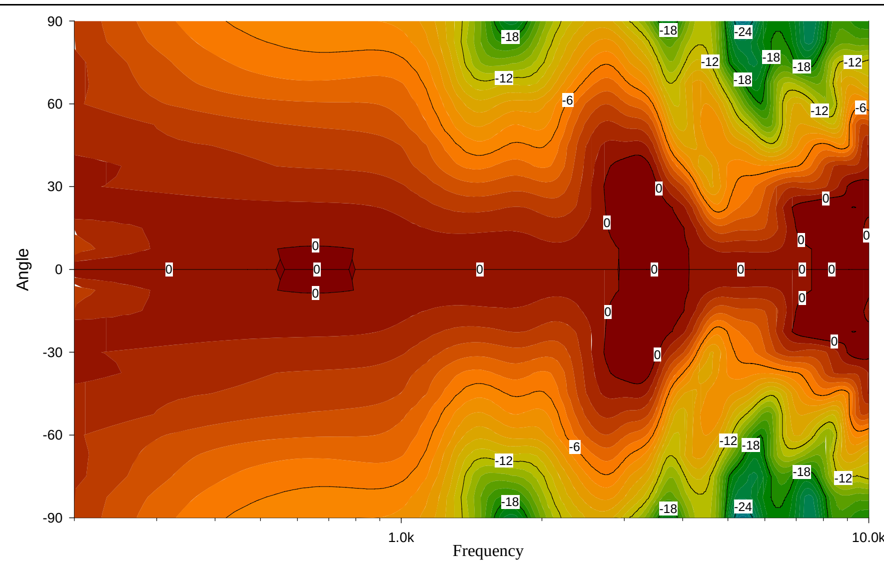


FIGURE 16. Example 3 normalized

The question then becomes could a waveguide and a piston source be made to work properly if they were not coaxially located.In Figure 18 I have shown the polar map for the Geddes Summa. This speaker is widely praised for its neutral and natural sound. A comparison between Figure 18 and the previous examples show why this would be obvious from the polar map. When toed-in as described on Page 12, one sits at the 22.5 degree axis where the response is optimized to be flat (its about +- 1 dB), the sound image that results is considered to be exceptional and is completely stable with seating location.The Summa can be seen to be very nearly the ideal in polar response with only a few small exceptions. The woofer has two resonances (since I know this speaker well, I know that they are mechanical and hence correctable). The waveguide

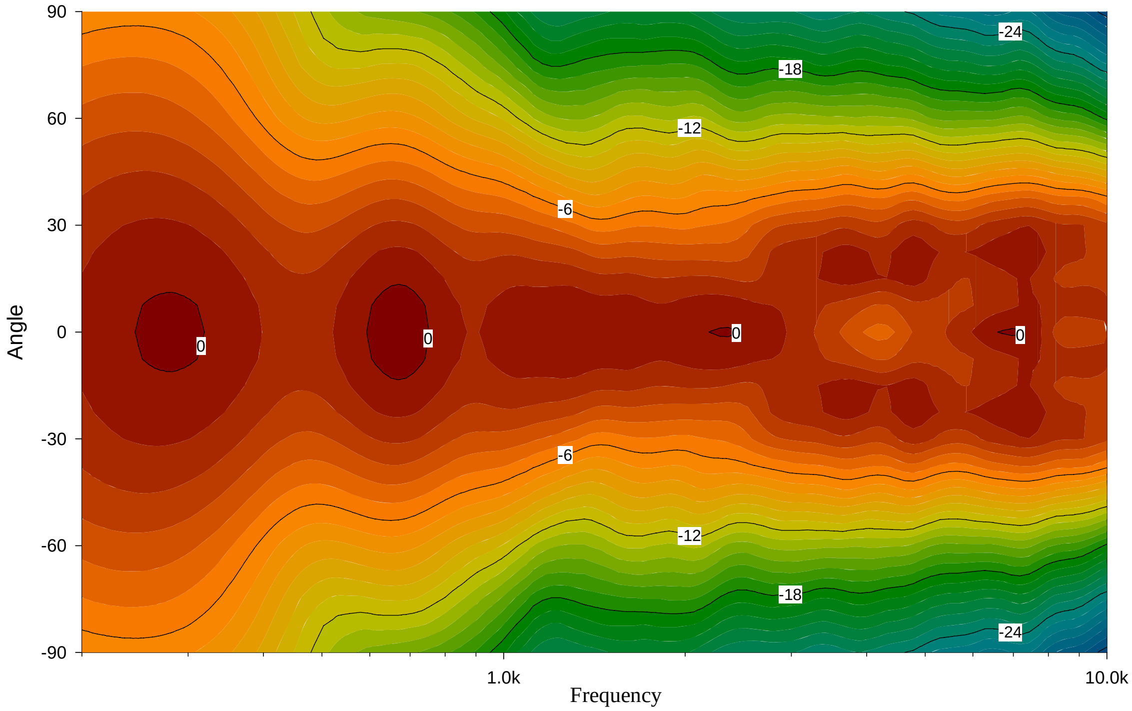


FIGURE 18. Geddes Summa loudspeaker system

exhibits a small dimple directly on axis at about 4-5 kHz. This is dues to diffraction from the mouth of the waveguide and is thus acoustic in nature and I would not expect it to be correctable. There is a slight narrowing of the directional pattern at around 1-2 kHz, but nothing severe. Normalizing the axial response, as I have done with the previous examples, results in Figure 18 on page 14, which confirms my suspicions were correct (or actually I knew they were correct when I stated them). The axial hole, being acoustic in nature, is not correctable The woofer resonances have been perfectly corrected. Basically for the waveguide no electrical EQ can provide any improvements on the device, however, the woofer could have the two small resonances corrected if deemed necessary. (This is being done in the current models.)

The take-away information from this last figure is that once one has a well designed loudspeaker, there is no need for electrical EQ and in fact it is likely to make things worse. ***Acoustic problems can only be solved acoustically, and once that is done, nothing more is possible.***

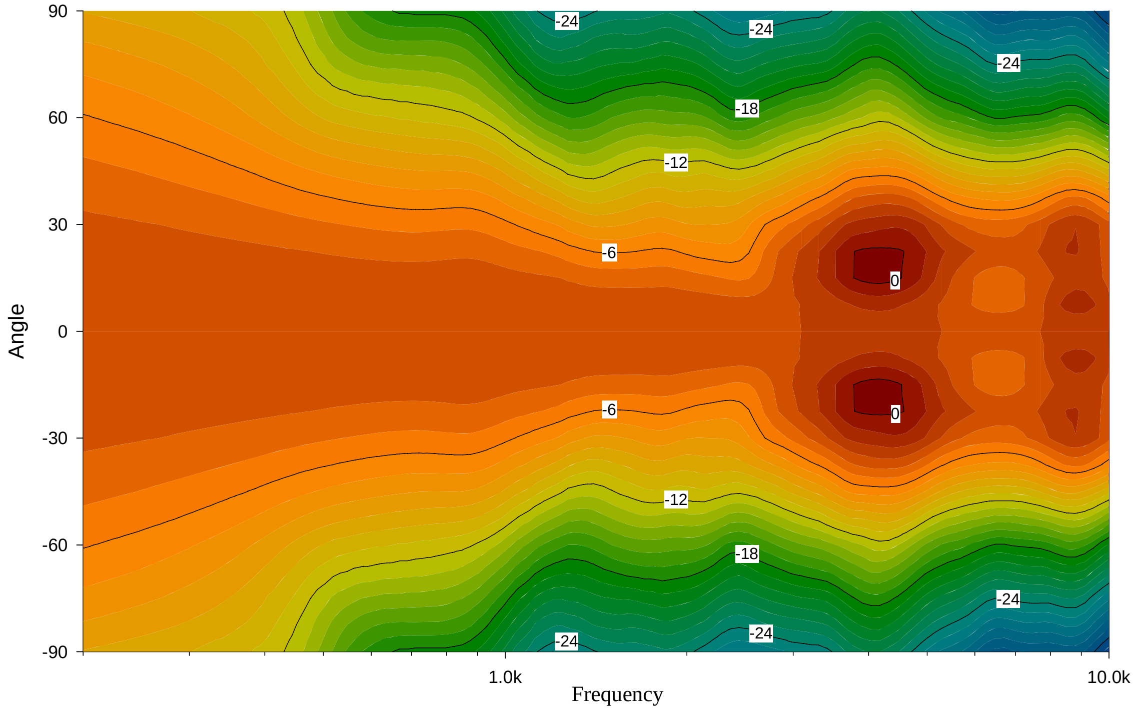


FIGURE 18. The normalized Summa

5.0 Conclusion

While the examples shown here for competitors speakers are not ideal and one could argue that they were all poor designs, I do submit that they are more typical than unusually bad examples. Could they be improved? Of course. Could any of them ever equal the Summa? No, that is simply not possible. Each one of them has built-in design flaws that cannot be circumvented without a complete redesign of the system. Basically no loudspeaker which does not use a waveguide is going to be able to match the Summas performance. Piston source loudspeakers simply cannot do what a waveguide does. But all waveguides are not equal either. The Summa has a particularly high directivity, which is desirable in a small room (another topic altogether). There are also other aspects of a design that were not touched on here that play unimportant role in sonic perception, most notably dynamics. But the fact remains that without getting the polar aspects right, the system will always be sub-optimal.

So do other speakers on the market compare to the Summa? I don’t know, no one else publishes data like this - or very few at least. And when they do, the resolution is usually not that good, making true comparisons difficult. The data shown here shows all the warts of a loudspeaker as all of the examples, including the Summa, show. The data also indicates that it is reasonable to look at data, if it is sufficient, and decide what will sound good and what won’t. This is a most controversial point, but will come to a resolution soon enough.