Proceedings of The Institute of Acoustics

DEVELOPMENTS IN LINEAR PHASE CROSSOVER DESIGN.

K.V.Castor-Perry.

Kemo Ltd., Beckenham. Kent.

1) What are crossovers in general?

Crossovers are circuits or networks intended to divide composite signals (such as, but not limited to, music sources) into several channels which can be separately processed and subsequently recombined. The most familiar application of these circuits is in the division of a musical signal into several frequency bands which can be then 'reproduced' as acoustic vibration by transducers optimised for these particular bands. The recombination process is one of acoustic mixing in the space around the enclosure supporting the transducers.Other applications include broad band noise reduction and data splitting.The problem of splitting a source up into separate portions of the frequency spectrum and then putting it back together again has received an immense amount of attention. Much controversy has raged over the subject in recent years, and it is the aim of this paper to clarify the problems involved in the typical case and to introduce some computational approaches to their solution in an informal way prior to the publication of more detailed research.

Throughout the text the system which will be considered is a simple two-transducer loudspeaker system. It can be shown that the complexity of the problem increases in a controlled (i.e.not exponential) manner as the number of bands to be produced increases, although initial investigation into multiband systems indicates that the amount of signal processing involved might act as a disincentive.

11) What is 'linear phase'?

A 'Linear phase' system is one whose input/output phase lag is proportional to the signal frequency over a certain range of frequencies. This might not seem like a very important property,but it can easily be shown that as a result any signal at a frequency inside this range is passed through with a time delay that does not depend on the frequency. Assuming that the amplitude response is flat, this means that a mixture of frequencies in this range, forming a composite waveform with a particular, unique shape, will pass through the system with the shape unaffected, since each of the component parts will have been 'pushed along' the time axis by the same amount. It has often therefore been a natural assumption that the systems of this kind are the best way of processing signals, since the information about the shape of complex signals is not lost.

If the same signal is put into a system where the phase shifts at the component frequencies do not lie on a straight line, the 'shape' of the output signal will be different from that of the input, and it can be argued therefore that a form of distortion has been introduced.

It is not within the ambit of this paper to comment on whether this distortion can be heard.

Until recently, however, it was not feasible for comparative listening test between non-linear phase and linear phase speakers to be made because of the absence of useable linear phase crossover networks.

111) Which networks are linear phase ?

Modern crossover design addresses a multitude of problems. The designer must acheive a brisk but well controlled transfer of control between the two drivers in a modern two way loudspeaker. The large low frequency driver can become incoherent and uncultured at high frequencies, and the large energy content in the lower mid-band will fry most high frequency drivers, with their light diaphragms and small coils. At the same time, the phase shift between the two drivers must not be too extreme, or most of the projected energy in a typical non-coincident driver system will get squirted out at odd angles from the listening axis.

The first order crossover (in its simplest form just an inductor for the1f unit and a capacitor for the hf unit) and the 'artificial' low order linear phase crossovers have been largely rejected by designers because of the large overlap region and poor final slope properties. All the higher order crossovers have an overall behaviour which is non-linear phase; low frequency signals have a certain finite time delay applied, whereas very high frequency signals tend to come out with no delay. Sometimes frequencies around the crossing-over point are delayed even more, causing very uneven time-domain behaviour.

A milestone in the field is the paper by Lipshitz and Vanderkooy[1]. They reasoned that, if the circuit driving the low frequency driver introduced a certain delay at low frequencies, then taking this low-pass signal away from a signal equal to the input but delayed by the same amount as the low-pass signal, then the low frequencies would cancel out and just the high frequencies would be left, idea not new but their paper most comprehensive. This indeed proved to be the case and all that was left for them to do was to find the particular low-pass circuits which would, when used in this way, give the best high-pass signal. Their paper is rich with examples, the one giving the most promising slope behaviour having a low-pass slope Of 18dB/oct. and a high-pass slope of 24dB/octave, this latter figure the equal of modern state-of-the-art non-linear phase designs.

There is always a catch; in this case there are two. The phase shift between the two signals produced by this elegant process is not zero, and the polar (dispersion) behaviour of our two-way loudspeaker would not be ideal. This represents a retrograde step since modern non-linear phase designs have this problem well under control. We will come back to this problem later.

The other catch is the generation of the time delay needed in the processing to obtain the high pass signal. Of course digital techniques can be used but this begs the question of fully digital crossover systems, which should be the subject of a whole paper in themselves. The construction of analogue all-pass circuits is a well-known art, but they are not trivial and the component sensitivity is high enough to cause worrying errors. The subtraction of two equal signals is an error-prone technique and it is the first original aim of this paper to point out that this subtraction can be carried out mathematically before the physical realisation, rather than on the signals themselves.

If it is supposed that the signal to drive the high-frequency unit is to be obtained by subtracting the output from one realiseable filtering circuit (our previously chosen low-pass filter) from that of another realiseable circuit (an all-pass filter, with a linear phase response over as much of the frequency range as we decide necessary) then it can be seen that this resultant signal is related to the input signal by an algebraic transfer function. It would seem sensible, therefore, to attempt to realise this transfer function directly, thus avoiding the sensitivity problems.

This was indeed attempted, and a condensed version of the working and results is shown in the figure. The method adopted was as follows.

1) Choose a ball-park figure for the crossover (-6dB) point.
2) Calculate the delay that the low-pass filter of choice (for instance, T3 from L&V) will have for this cut-off.
3) Choose an all-pass transfer function which will give this delay at all frequencies up to a chosen maximum near the audible extreme. A Bessel alignment is the conventional choice; here, a 6th order .5 degree equiripple phase function was used, as it trades in absolute linearity of phase for some extra bandwidth.
4) Normalise the all-pass function to match the low-pass function. This eases the sums somewhat.
5) Form the numerator and the denominator of the difference between these two functions. The denominator will just be the product of the two denominators and needs no further work, but
6) Expand and order the numerator in powers of s

At this stage, it should be obvious that the coefficients of low powers of s are very small indeed. Using a perfect all-pass filter and perfect arithmetic the coefficients of s 0 to s 3 should be zero, reflecting the fact that the filter we are aiming at has a slope of 24dB/octave as the frequency tends to zero. Now in this working they were not quite zero, as the all-pass was not perfect. However, we drop them anyway - the error is very small, and anyway that was the desired result.

7) Split off all free powers of s (here it's s 4)

DEVELOPMENTS IN LINEAR PHASE CROSSOVER DESIGN.

Figure 1

Example using T3 alignment from V & L with fc=2438Hz.

Normalised lowpass:

$$\frac{1}{(s+0.8787)(s^2+0.7726s+1.138)}$$

Suitable allpass up to audible limit:

$$\frac{(s^2-4.2393s+6.4438)(s^2-4.0374s+21.861)(s^2-3.1736s+48.046)}{(s^2+4.2393s+6.4438)(s^2+4.0374s+21.861)(s^2+3.1736s+48.046)}$$

referred to as $Hd(s) = \frac{Pd(s)}{Qd(s)}$ and $Hl(s) = \frac{1}{Ql(s)}$

We want to find $Hh(s) = \frac{Pd(s)Ql(s) - Qd(s)}{Qd(s)Ql(s)}$

note that $Pd(s) = E(s) - O(s)$ and $Qd(s) = E(s) + O(s)$, so that the the numerator is $E(s)\{Ql(s) - 1\} - O(s)\{Ql(s) + 1\}$, and subtraction of the two sides gives:

Power of s	LHS	RHS	result	comments
9	1	1	0	numerator is n=8
8	1.6515	11.4503	-9.7988	
7	121.55	18.91	102.64	
6	197.73	681.3	-483.56	
5	2917.4	1113.7	1803.6	
4	4458.8	7349.7	-2890.9	
3	11673	11477	196	ideally 0
2	11178	11173	5	ideally 0
1	12297	12299	-2	ideally 0
0	0	0	0	by design

drop last four powers of s and find the roots, giving numerator:

$$s^4(s-2.5949)(s^2-3.9908s+23.2915)(s^2-3.213s+47.832)$$

the complex roots are near roots in the original Qd(s) so shift them to these positions, rearrange, and get Hh(s)=

$$\frac{(s-2.5949)(s^2\qquad\qquad\qquad)(s^2\qquad\qquad\qquad)}{(s+0.8787)(s^2+0.7726s+1.138)(s^2+4.2393s+6.4438)}$$

$$\times\ \frac{(s^2-4.0374s+21.861)(s^2-3.1736s+48.046)}{(s^2+4.0374s+21.861)(s^2+3.1736s+48.046)}$$

which is -24dB per octave HP as desired.

8) Stuff the remaining polynomial into a root-finding routine. These are available on any small micro; there is a good one on the tape that com comes with the Hewlett-Packard HP-85.
9) Form up the quadratic factors from the conjugate root pairs
10) Write down the whole thing in one place.

In this work example, it can be seen that some of the derived roots on the numerator are rather close to the original roots of the all-pass filter. A proof of the maximum bound on the shift of any pole is beyond the scope of this paper, but in this case it was felt valid to re-shift the poles to form individually realiseable all-pass sections. The final result of all this approximation is shown in the figure; the response is rather good, although the wobble in the power response gives the unexceptional interdriver phase away by comparison.

This derived circuit can be built and aligned separately from the low-frequency circuit, and although the response can change with inexact components there is no danger of the response at low frequencies being perturbed; this is guaranteed by design.

1V) Further developments on this technique.

The techniques described is a synthesis procedure; apart from a few simplifying approximations, it produces the required filter circuit by a single application of a closed calculatory process. We must now ask how closely we can approach the ideal result using functions which are not apparently suited to our purpose. We do this by specifying how inaccurate the result is allowed to be, and then using a microcomputer to adjust circuit parameters iteratively until we reach the'best' solution. (We must have a decision procedure to decide upon which solution is 'best'). For instance, it can be seen that the shape of the high-pass section when the response is, say, 40dB down (i.e. 1/100th of the input value) is not going to have much to say in the overall response. This opens the possibility of modifying some of the transfer functions so that, for instance, the slope of one or both curves can be enhanced, or the phase match between high-pass and low-pass sections can be improved - naturally this only needs to be the case at frequencies where both transducers are contributing to the overall result.

It is convenient to select a suitable function for the low-pass circuit first, and it will be seen that the choice of this function can determine completely the acheivable phase match beteen the two sections. In principle, though, any low-pass section can serve as the initial reference point; it may be that a new high-pass section is being designed to go with an existing low-pass circuit that should not be altered. We can then make two or three stipulations:

1) The amplitude response of the summed signals should be flat to within an allowed error band;
2) The phase response of the summed signals should be linear with frequency

to within an allowed error; and perhaps
3) The phase shift between the outputs of the high-pass and low-pass sections should not exceed a certain small value over the range of frequencies where both outputs contribute significantly to the total.

The appropriate values for part 3 can be derived from geometrical considerations; see the concept of 'lobing error' in V&L. The first two criteria can alternatively be viewed as saying that the difference between the summed signals and a purely delayed version of the input signal should be as small as possible; it is this approach which is the most elegant to realise computationally.

The concept of optimisation involves forming an error function, which depends on frequency (here) and any circuit values or function coefficients which are allowed to vary, and using one of several algorithms to systematically reduce the 'value' of the error finction (often this is the sum of the squares of all the errors at the frequencies for which the calculations are done) until it cannot be reduced any further. This then represents the best that can be acheived with the elected choice of components,and can be used if the actual error resulting from its use lies inside the pre-defined boundaries. If not, a more elaborate function is used in the initial stages, chosen either by experience or by using techniques such as the one described above, until all the requirements are adequately fulfilled.

Computer programs written to develop such filters can easily be tested by using as a test input a filter design such as the one carried out longhand in this paper. If the program is functioning correctly, it should converge to a solution for the high-pass function which is very similar to the one derived here (although the all-pass portions can often be rather ill-conditioned and may turn out rather differently; this is due to the relaxation of requirements from perfection, and the particular way in which most of the recommendable optimisation algorithms form the error value)

The third criterion mentioned above was to keep the phase difference between high-pass and low-pass signals to a minimum. This can be done in two ways. The first is to perform a separate optimisation on the phase difference running concurrently with the first. The total error is regarded as a combination of the two separate errors with a weighting factor decided by experience. In these circumstances the program must be allowed to alter the parameters of the low-pass side as well as the high-pass, or at least to multiply in an additional function where existing circuitry must remain. It can be shown that the two goals of linear overall phase and near-zero phase difference can only be met if the low-pass section has a phase delay (this is the phase shift divided by the frequency) which is nearly constant for all frquencies up to the cross-over point with the high-pass. Of course a certain leeway is possible as the optimisation can take up some 'slack'. However, the alignments in V&L which are best for high-pass slope do not meet this condition. The only suitable responses in V&L are the higher order

DEVELOPMENTS IN LINEAR PHASE CROSSOVER DESIGN.

Bessel ones, with exceptional phase match. The Bessel filters, and certain other standard types of filter with similar response, are good places to start from with a computer optimisation process such as this; although the V&L synthesis only acheives 12dB/octave for the derived high-pass, there is the attractive possibility of forcing this higher if we accept a moderate departure from perfection.

Another approach to the generation of a suitable low-pass circuit is to linearise the phase of the low-pass first with some all-pass circuitry, and then to perform either the V&L synthesis or a computer optimisation.

V) What about drive units ?

Implicit in the entire discussion has been the use of ideal drive units. However, there are several techniques available to adapt the crossovers so designed to real drive units. The most comprehensive way, indeed the only feasible way if the system is to be designed with a passive crossover, is to use an optimisation program to adjust the overall response of topologically suitable networks driving the drive units themselves and incorporating the acoustic responses of the drive units, until the best match is obtained with the theoretical figures. Of course, this analysis may be performed simultaneously with the first optimisation.

The second way is to derive a transducer function which represents the particular features (resonance, directionality) which need to be removed, and divide that out of the theoretical transfer function. This was done for the long-hand synthesis in this paper for the resonance of the hf unit.

Inter-unit time delays due to construction may also be coped with by adjusting the enevitable all-pass sections which will be in the high-pass sections until the time delay is compensated.

In conclusion, the aim of this paper was to give a necessarily brief and rather rushed introduction to modern techniques in linear-phase crossover design. Some more detailed theoretical work, and the practical results from it, will be the subject of a future paper.

Reference: A family of Linear - Phase Crossover networks of high slope derived by time delay. S.P.Lipshitz & J.Vanderkooy. JAES V31 No 1/2 1983.