# ELECTRONOTES 202

Newsletter of the Musical Engineering Group

1016 Hanshaw Rd. Ithaca, NY 14850

Volume 21, Number 202 August 2003

## GROUP ANNOUNCEMENTS

Contents of EN#202



## BOOK REVIEWS

I recently read The Legend of Barjo Restaurant - the Life of Josephine McAllister Stone by J. Emily Foster, Soleil Press (2001). The book is pretty much self-published (but available from Amazon.com!), and likely mainly of local interest. Possibly some of you have been fortunate enough to have eaten at the Barjo Restaurant (in Norway, Maine). It was a place where they served you plenty of great tasting, wholesome food, and never worried about giving it fancy names. Jo served her hot meals on hot plates (she heated them in the oven) and the waitresses always warned you not to touch the plate when the food was set on table, and you remembered not to try it a second time! It was a place where if you asked to take out a slice of pie for your grandfather up at the nursing home you were asked which flavor he preferred, and if you tried to keep things simple by saying he liked all of Jo's pies, you were asked his name and then told which flavor he liked best, and then they wouldn't accept any money for the slice.

Of course, I loved the book, but it was surely because I knew enough about the people and the places to relate. Great story that it is - it is hard to imagine that a general reader would get too excited about it. It occurs to me that Tracy Kidder could pull this sort of thing off and write a best-seller called Restaurant. But alas, Kidder only writes about ordinary people.

Reading the Barjo book was for me what many people would describe as a nostalgic experience: back in time, to previously visited places, and to old friends. But it is more than that. It is of course not my diary of events I myself experienced. I learned much much more than I originally knew, and pasted these things on an increasingly familiar framework.

Analog Days - The Invention and Impact of the Moog Synthesizer, by Trevor Pinch and Frank Trocco, Harvard University Press (2002), which I was reading at the very same time, was also a book which transported me in time and in place. It too was a book which reminded me of many things I did know at one time, while at the same time adding details. Most importantly, it told me many things I did not know before. Many of the people in the book I knew quite well, while others of them I never met. A lot of the stories I knew, but there were many more that I now learned, and I could fit them into a more complete picture. And it triggered fond memories of stories that I knew that were of course not in the finite-length book. During the evenings I was reading these two books, it was often a jolt to find myself on my couch in Ithaca in the year 2003.

Analog Days is first of all a scholarly work, extremely well researched, and yet told in a lively way uncharacteristic of many histories dealing with technological issues where it is often left to the reader to provide any elements that would make the work compelling. For readers of this newsletter, this was probably not necessary - they would have had to read it for the information - but the lively presentation is a bonus. By the way, it should go without saying at this point that all our readers should read this book.

I first heard of this book (that it was being written) back about 1995 from Ron Kline at Cornell. and subsequently met with and talked with the authors a couple of times. In the years that followed, I sort of wondered if it was actually going to come out. When it did, it was clear that during the years they were working on it, they had been working hard. It seems that they spared no effort to get to the right people to interview and get to original sources. They could have cut corners - but they didn't.

This book is about Bob Moog. It is his story they are telling, as the subtitle suggests. But it does not ignore the other pioneers such as Don Buchla and Alan Pearlman. We meet these people and many more, much as Bob, initially quite isolated, likely also met them as time went on. And the portrayals are honest and usually charming - no need for a hagiography when the characters were so interesting.

( - continued on page 28 )

# ACTIVE COMPENSATION OF SALLEN-KEY LOW-PASS FILTERS

-by Bernie Hutchins

## 1. BACKGROUND

Active filter designs are often trivial to do on paper (where we assume idealized elements). Realization with real components and solder may or may not be more difficult. While the use of real elements can complicate design, many "ideal designs" will be perfectly satisfactory in practice, with no modifications to the "cookbook" paper design. We are likely to have this sort of success if our filters are of low order and designed for low frequency operation. However, if the order is high (perhaps greater than eight or so) and/or if any of the poles are high-Q (perhaps greater than 20 or so) and/or if the frequency is high (perhaps greater than about 10% of the gain-bandwidth product of the op-amp used), we may have to be more careful. Care in this sense may relate first to the need for passive components with greater precision (lower tolerances): so-called "passive sensitivity." Or (very likely) it may also relate to the choice of active elements (op-amps) with higher gain-bandwidth product (GBP): so-called "active sensitivity." Because it is usually impractical to just get better and better op-amps, the preferred approach is to employ certain compensation techniques to tailor the performance.

In large part [1,2], we have confined our efforts in dealing with active sensitivity to two paths. In the first path, we recalculate the transfer functions using real rather than ideal op-amp models. In the ideal op-amp, we assume that the differential input voltage is zero. In the real op-amp, the output voltage of the op-amp is related to the differential input voltage ($V_+ - V_-$) as:

$$V_{out} = (G/s)(V_+ - V_-) \qquad (1)$$

where G is the GBP of the op-amp in radians/sec, while the "s" in the denominator indicates that the usual single-pole roll-off (integrator) model for the op-amp's internal compensation is being used.

This adds a pole to the network for each op-amp it contains. For example, there are many second-order networks that use a single op-amp, and when we use a real op-amp, we now find (typically) two poles at the near-nominal positions and a third pole which is real and often relatively far away out on the negative real axis. Our concern is than usually with the two near-nominal poles, which we find have moved to some degree, and we wish to know if the actual performance will be nonetheless satisfactory.

In cases where it is not satisfactory, we can often use an "overdesign" technique. This is a classic ploy of determining the direction and amount by which the performance degrades. We then redesign the filter in the opposite direction by a corresponding amount. The hope is that it may then "fall back" to a position that is just about right. With iteration, this can be very effective.

The second path is one of constructing our filters with improved building blocks [2]. Most commonly we have used this approach when dealing with filters based on integrators and summers (the "state-variable filter" and certain signal-flow-graph realizations). In an initial attempt to realize these structures, we quite naturally choose integrator and summer circuits that are perfect if we consider the op-amps to be ideal. We know that when we use real op-amps, we pick up additional poles, and because the circuits involve several or many op-amps, serious performance degradation can result (again, at high orders, frequencies, and Q's). Our attempt to deal with these real op-amp problems is to use sub-circuits that are "improved" integrators and summers. This we have approached in terms of "passive compensation" (of active sensitivity), adding small "trimming" resistors or capacitors; or by "active compensation" (of active sensitivity): adding even more op-amps - but in a way such that undesirable effects cancel at least in part. Thus we have used both passive and active methods of dealing with the active sensitivity problem, which we have applied to the individual building blocks of the circuits.

This second path has been employed to construct successful state-variable filters [1,2]; not infrequently using more trial-and-error than theory. We have also made efforts to look at the state-variable filter by real op-amp analysis [3,4], with the idea that we could then always apply an overdesign method (at least for fixed-frequency cases). The analysis further shows [4] that when active compensation of the integrators and summers is used, the positions of the desired poles moved much less (relative to no compensation) and the residual poles and zeros took on a particularly advantageous arrangement - said arrangement apparently being inherent in the general method [5,6].

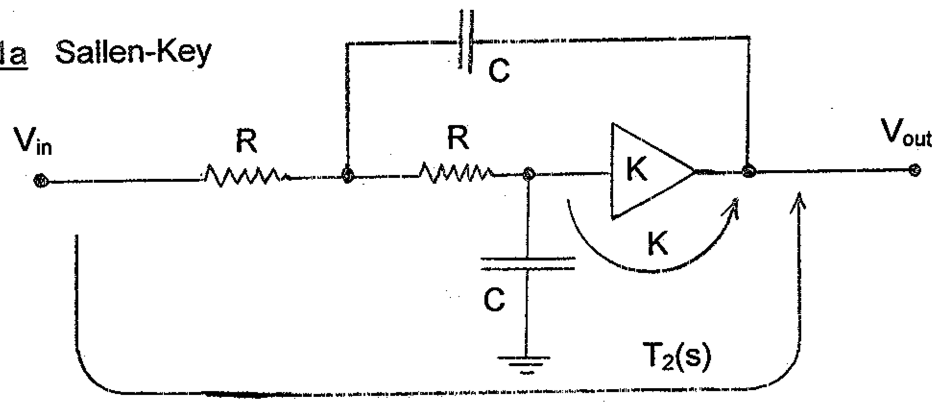
## 2. SALLEN-KEY WITH A REAL OP-AMP

One thing which we have not yet looked at is the idea of using active compensation of single op-amp filters (as stated, we have relied on overdesign). In fact, we can try this most easily with the Sallen-Key arrangement, since this circuit is well studied and uses a finite gain voltage amplifier, and we know how this amplifier behaves with a real op-amp [7], and how to actively compensate it [8].

### 2a. Ideal Sallen-Key

Fig. 1a shows the usual network for a Sallen-Key low-pass filter. Here the triangle indicated by K is understood to be an ideal voltage amplifier with a gain of K. Fig. 1b

Fig. 1a Sallen-Key



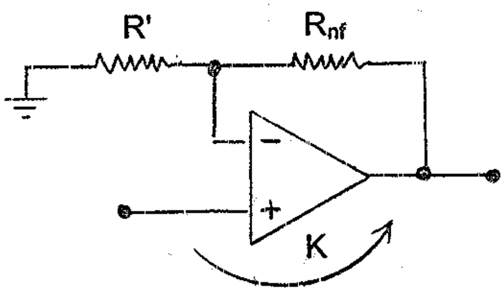




Fig. 1b Non-Inverting Amp.

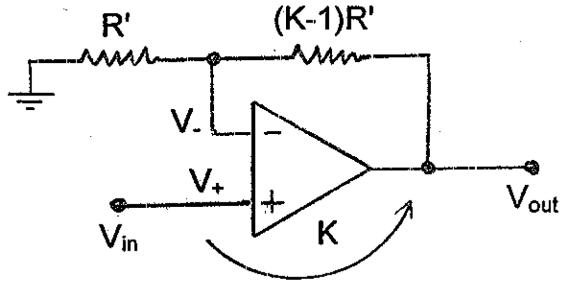



Fig. 1c Simplified Form

shows the way the amplifier is usually realized with an op-amp. If we assume an ideal op-amp, we find an ideal voltage amplifier with gain $1+R_{nf}/R'$. Note that only the ratio of resistor values matters, so it is convenient to represent them as in Fig. 1c. Network analysis easily leads us to the transfer function of the overall filter:

$$T_2(s) = K/(R^2C^2) / [\, s^2 + (3-K)s/RC + 1/R^2C^2 \,] \qquad (2)$$

and this is likely very familiar [1]. At this point we note that we are thinking of K as a constant (independent of frequency). But it is perfectly possible to have K(s) and simply plug the expression for K(s) into equation (2) in place of K. Shortly, we will do just that. We will use K(s) for the real op-amp case and for the active compensated case. In reality, this plug-in simplifies the algebra only some - we do not have to redo the network analysis for the rest of the network, but we still have some algebra to chug on after the substitution. For notational purposes, we have indicated equation (2) as being $T_2(s)$ where the 2 corresponds to two poles (the two nominal poles).

## 2b. Non-Ideal Op-Amp

In the case of a real op-amp, we need to bring in equation (1). Using Fig. 1c we note that the (+) input to the op-amp, $V_+$, is just $V_{in}$ The (-) input to the op-amp, $V_-$, is determined by the passive voltage divider:

$$V_- = V_{out} [ R' / (KR'-R'+R')] = V_{out}/K \qquad (3)$$

We note that when we have an ideal op-amp such that $V_+ = V_-$ we arrive at $V_{out}/V_{in}=K$ as we have mentioned. However plugging into equation (1) we arrive at:

$$K_1(s) = V_{out}/V_{in} = KG / [Ks + G ] \qquad (4)$$

which indicates that the voltage amplifier is not just a gain K but rather a frequency dependent gain $K_1(s)$ which is a first-order low-pass. (Of course $K_1(s) \rightarrow K$ as $s \rightarrow 0$ or as $G \rightarrow \infty$.) The pole of $K_1(s)$ is at -G/K; hence the low-pass cutoff is at G/K. Since $G/2\pi$ is something like 1 MHz to 5 MHz, this low-pass amplifier would have a cutoff well above the audio range. But what does it do to the Sallen-Key filter?

## 2c. Non-Ideal Op-Amp in Sallen-Key

Here we borrow equation (2) and plug in equation (4):

$$T_3(s) = K_1(s)/R^2C^2 \ / \ [ s^2 + (3 - K_1(s))s/RC + 1/R^2C^2 ]$$

$$= (G/R^2C^2) / [ s^3 + (G/K + 3/RC)s^2 + (3G/KRC - G/RC + 1/R^2C^2)s + G/KR^2C^2 ] \qquad (5)$$

The result is of course third-order. There are three poles. This is not a new result and we have used this to show the active sensitivity of the real Sallen-Key [9]. None the less, $T_3(s)$ will be a fundamental comparison case for the compensation method that follows

# 3. APPLYING ACTIVE COMPENSATION TO SALLEN-KEY

## 3a. Fixing the Amplifier

How do we fix the amplifier? Well, this we have looked at [8], and the method is to put a second op-amp in the feedback loop of the first as in Fig. 2a. This is a tricky-looking circuit, but only because we are being efficient. All that we really have is an op-amp stage (upper op-amp) with an attenuator 1/K followed by a gain of K (net: unity) in the feedback loop of the lower op-amp. This was the "magic trick" we have seen [2] -

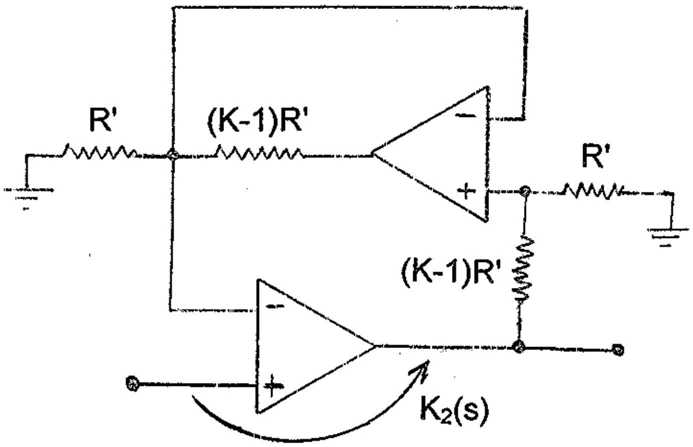



Fig. 2a Active Compensated Non-Inverting Amp.

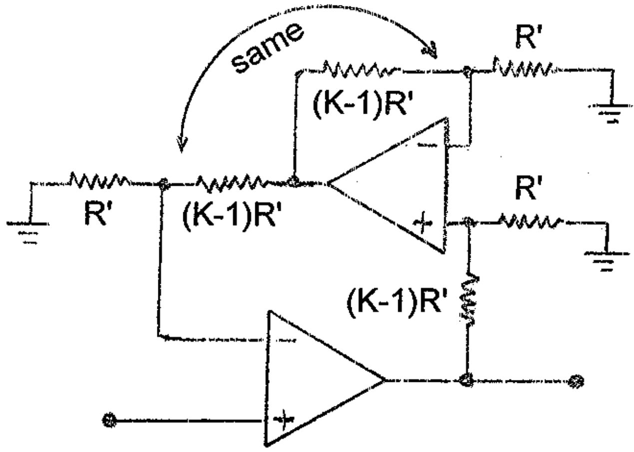



Fig. 2b Equivalent But Less Efficient Form of 2a

make the two op-amps have the same "noise gain" (inverse of feedback ratio from the op-amp output to the inverting input). This is more clearly seen in Fig. 2b. Fig. 2a is just sharing the same voltage-divider, and is preferred because it uses two fewer components, and there is no issue of having to match the divider ratios.

Analysis of Fig. 2a for the real op-amp case is just a matter of applying equation (1) to both op-amps. The result is:

$$K_2(s) = V_{out}/V_{in} = (G/K)(G+sK) / [\, s^2 + Gs/K + G^2/K^2] \qquad (6)$$

This hardly looks like an improvement. $K_1(s)$ was frequency dependent, but only first-order. $K_2(s)$ is second order. In fact, $K_2(s)$ has a zero at $s = -G/K$ and poles at:

$$p_{1,2} = -G/2K \pm (G/2K)\sqrt{3}\ j \qquad (7)$$

Again, this pole/zero array (Fig. 3) is the "magic result" we have seen where the phase due to the zero exactly cancels the phase due to the two poles immediately around zero frequency (with significant cancellation for a range of useful frequencies as we go above zero).

### 3b. <u>The Active Compensated Amplifier Inside the Sallen Key</u>

Finally we get to something new (Fig. 4b). Just as we plugged $K_1(s)$ into equation (2) we can plug equation (6) for $K_2(s)$ into equation (2). We arrive at:

$$T_4(s) = \frac{(G/K)(G+Ks)/R^2C^2}{s^4 + (G/K+3/RC)s^3 + (G^2/K^2 + 3G/KRC - G/RC + 1/R^2C^2)\, s^2 + (3G^2/K^2RC - G^2/KRC + G/KR^2C^2)s + G^2/K^2R^2C^2} \qquad (8)$$

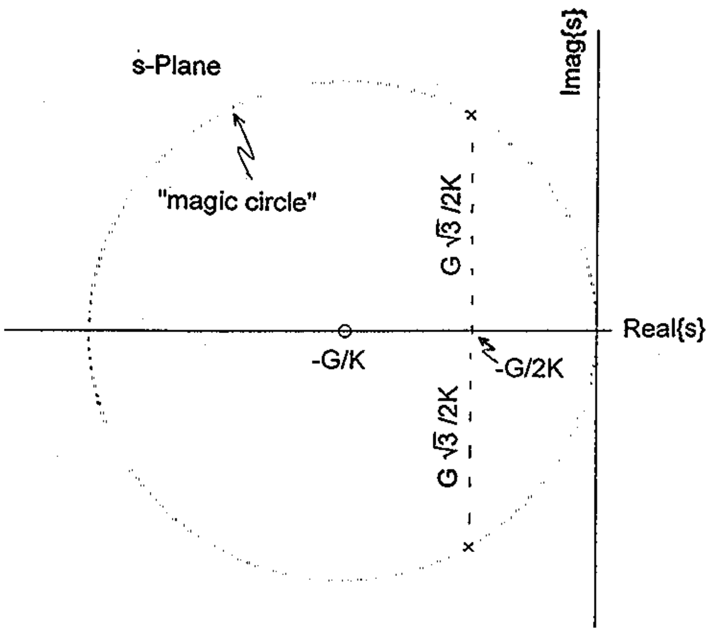



Fig. 3 The active compensated non-inverting amplifier has the "magic" array of two poles and one zero. The poles are on a circle that passes through s=0 and is centered on the zero.

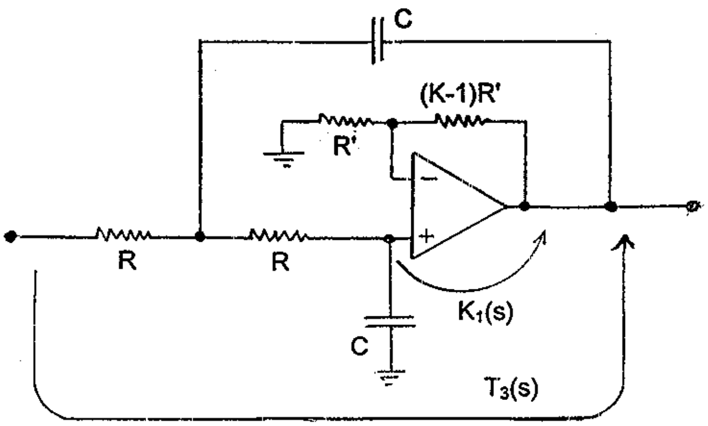



Fig. 4a Sallen-Key with real op-amp. This is just the original circuit except we assume a real op-amp is used. It has three poles - two due to the capacitors, and one due to the op-amp.

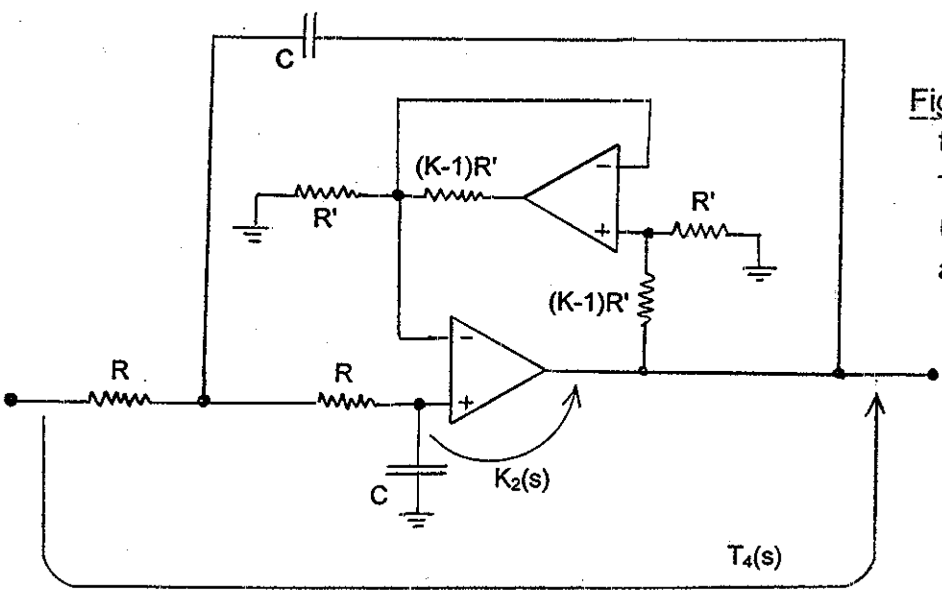



Fig. 4b Sallen-Key with two real op-amp. This network ends up with four poles and one zero.

This, as expected, has four poles, and one zero. Our interest in solving for the poles and zeros is to see where the extra two poles end up. It will turn out that we again get the "magic result." This will lead to the understanding of why when one op-amp messes thing up, a second op-amp may lead to significant correction.

## 4. COMPENSATION - FREQUENCY RESPONSE VIEWPOINT

At this point, it is useful to calculate a few examples to see how a single op-amp can degrade a filter's response away from nominal and how the addition of a second compensating op-amp can restore a useful response. The first example we will consider is an 8th-order Butterworth lowpass, assuming a normalized GBP of $G_n$=50. [ For example, if we had 4.5 MHz GBP op-amp, we would be designing for a low-pass cutoff of 90 kHz.] Note that this 8th-order network is composed of four 2nd-order sections in series.

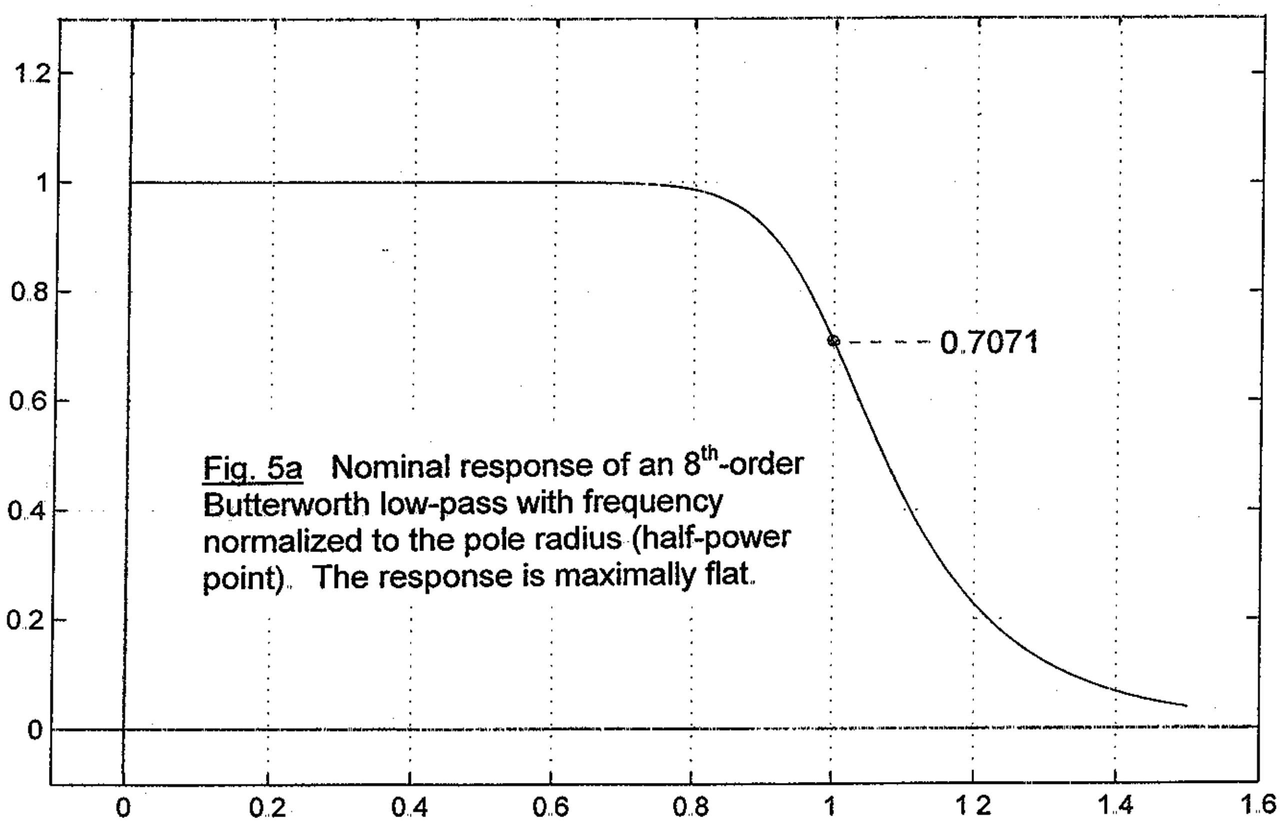



Fig. 5a Nominal response of an 8th-order Butterworth low-pass with frequency normalized to the pole radius (half-power point). The response is maximally flat.

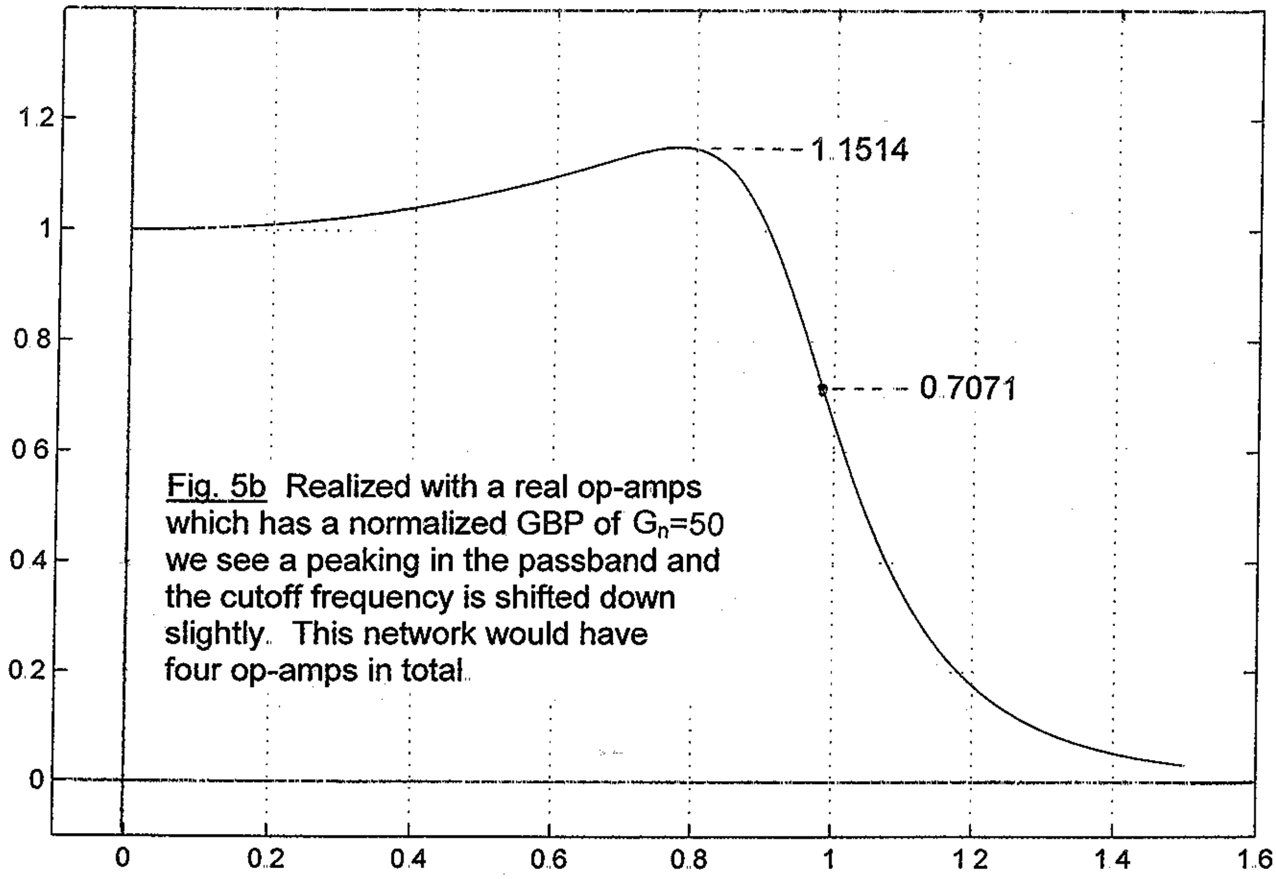



Fig. 5b Realized with a real op-amps which has a normalized GBP of $G_n$=50 we see a peaking in the passband and the cutoff frequency is shifted down slightly. This network would have four op-amps in total.

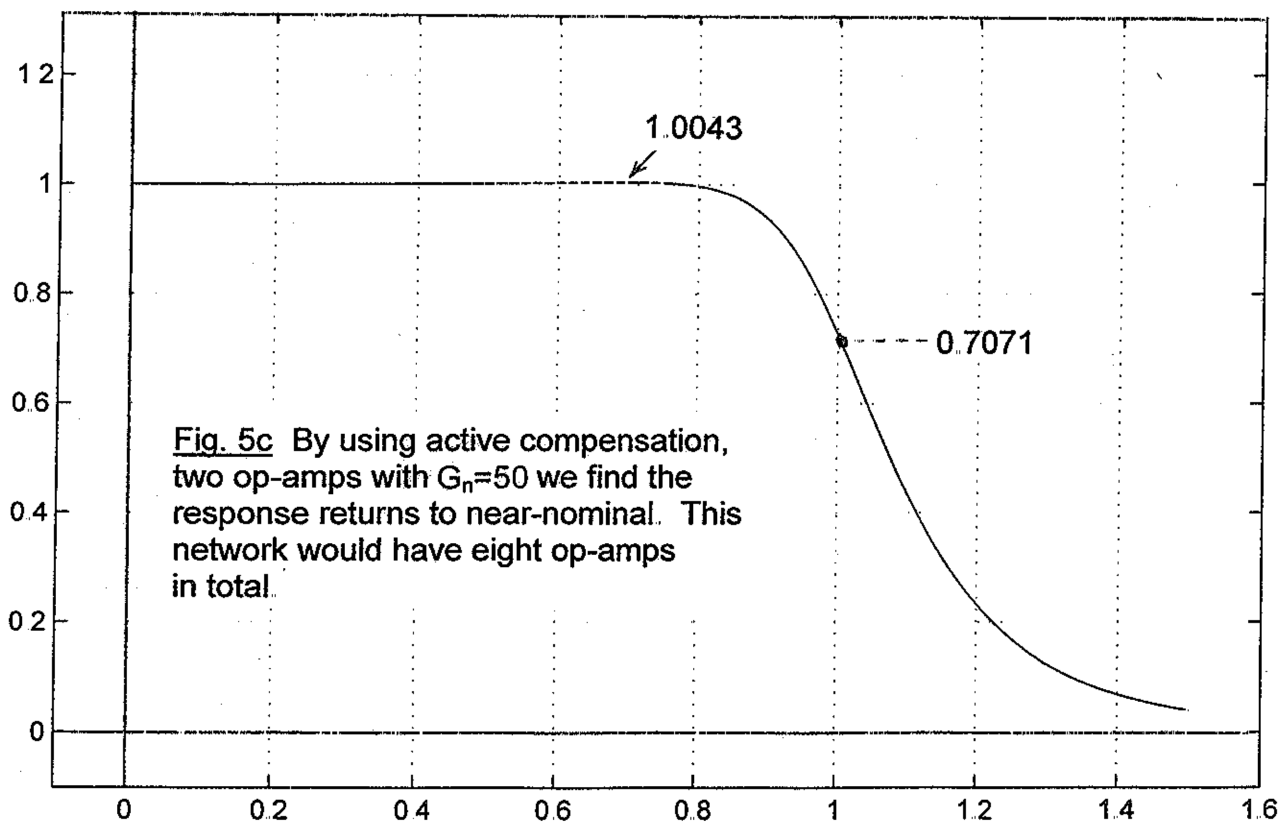
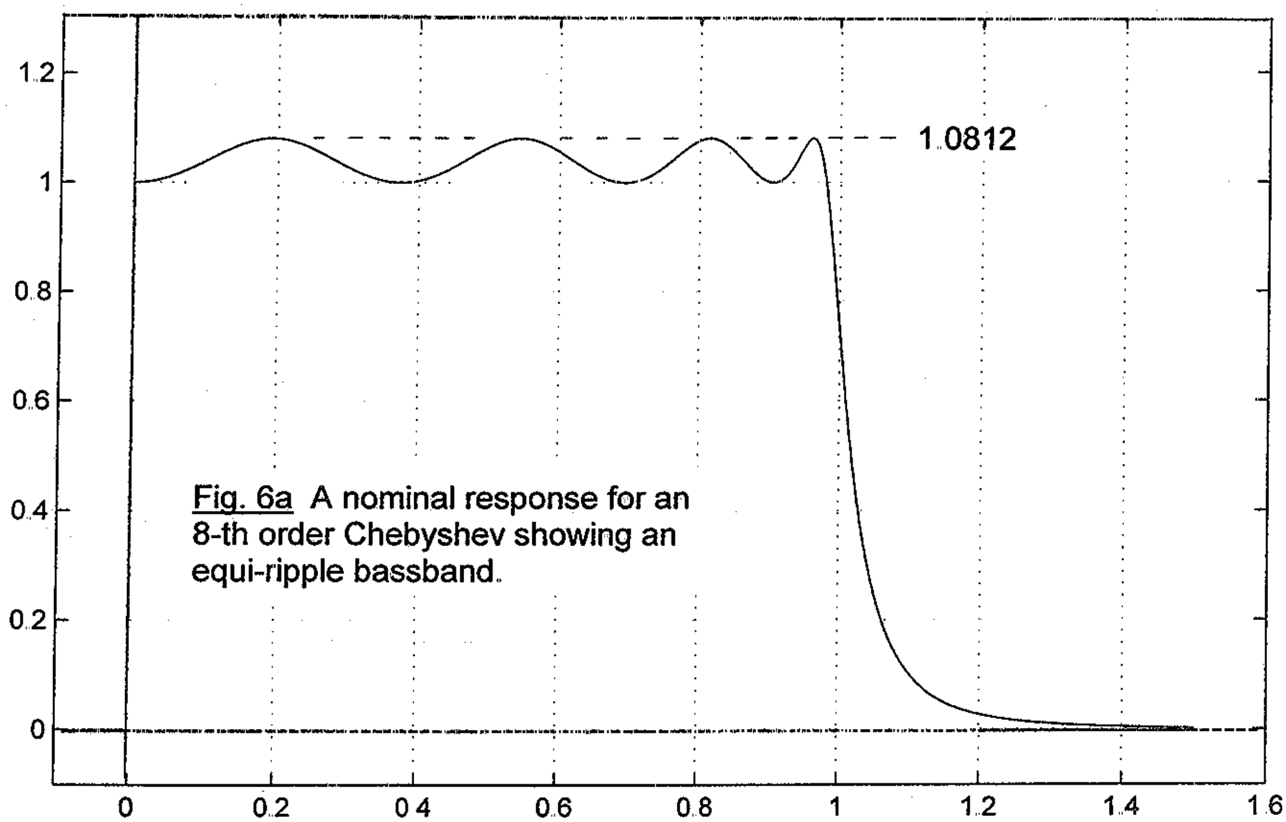



Fig. 5c By using active compensation, two op-amps with $G_n$=50 we find the response returns to near-nominal. This network would have eight op-amps in total.

Fig. 5a shows the nominal [ideal op-amp using $T_2(s)$] response normalized to a cutoff frequency of 1. With real op-amps [four 2nd-order sections like Fig. 4a, described by $T_3(s)$], the response of Fig. 5b is obtained. This response shows a peaking in the frequency response, reaching about 1.15 as the cutoff is approached, instead of the flat Butterworth response. Now, using active compensation [four 2nd-order sections like Fig. 4b, described by $T_4(s)$], the response of Fig. 5c is obtained, which seems indistinguishable from the nominal response (Fig. 5a). Actually it has a very slight rise - to 1.0043 and a slightly higher cutoff frequency.

For a second example we can look at an 8th-order Chebyshev low-pass. Fig. 6a shows the nominal [ideal op-amp, $T_2(s)$] response. This filter has a designed-in equi-ripple of 1.0812 as shown. Assuming a value of $G_n$=50 as in the Butterworth example, we get Fig. 6b for the real op-amp case [$T_3(s)$] and Fig. 6c for the compensated op-amp [$T_4(s)$]. The real op-amp case of Fig. 6b shows a rise to 1.54, while the compensated case of Fig. 6c shows a rise of 1.23. So while $G_n$=50 was enough for the 8th-order Butterworth, it does not work anywhere near as well for the Chebyshev (which has higher Q poles). Increasing $G_n$ to 100 (getting an op-amp that is twice as fast, or designing for half the cutoff frequency) works better. Fig. 6d shows the $T_3(s)$ case for $G_n$=100 while Fig. 6e shows the $T_4(s)$ case, with peakings of 1.33 and 1.11 respectively




<u>Fig. 6a</u> A nominal response for an 8-th order Chebyshev showing an equi-ripple bassband.

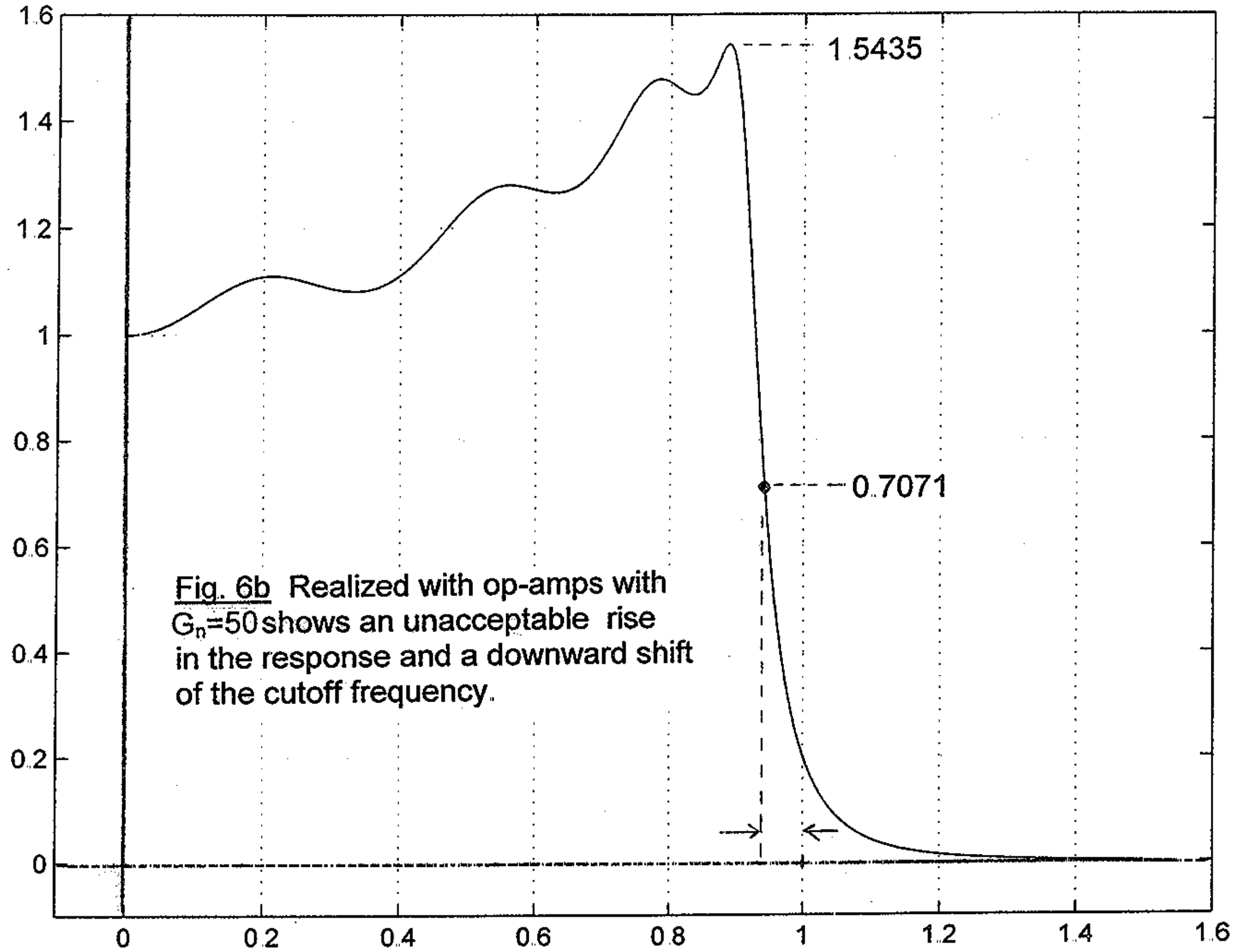



Fig. 6b Realized with op-amps with $G_n$=50 shows an unacceptable rise in the response and a downward shift of the cutoff frequency.

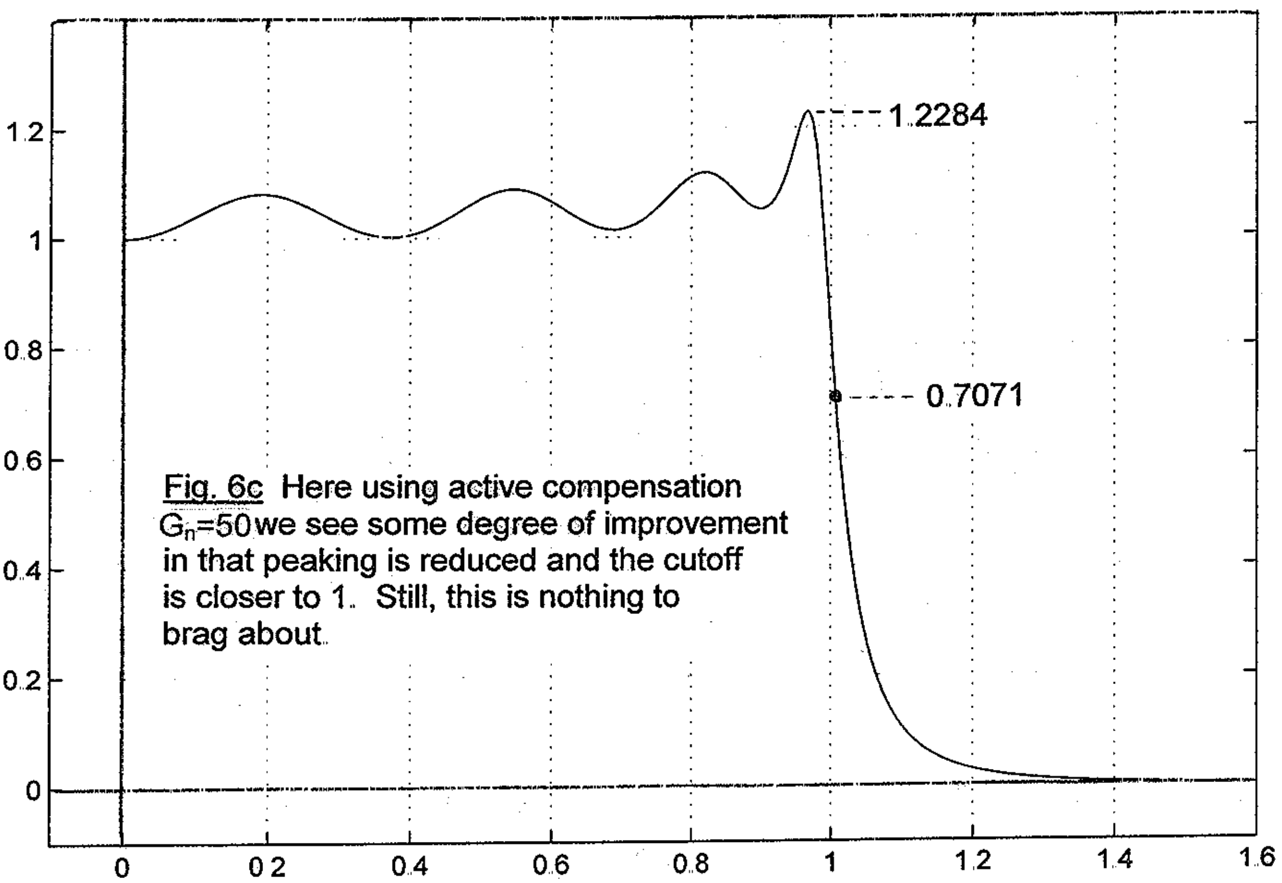



Fig. 6c Here using active compensation $G_n$=50 we see some degree of improvement in that peaking is reduced and the cutoff is closer to 1. Still, this is nothing to brag about.

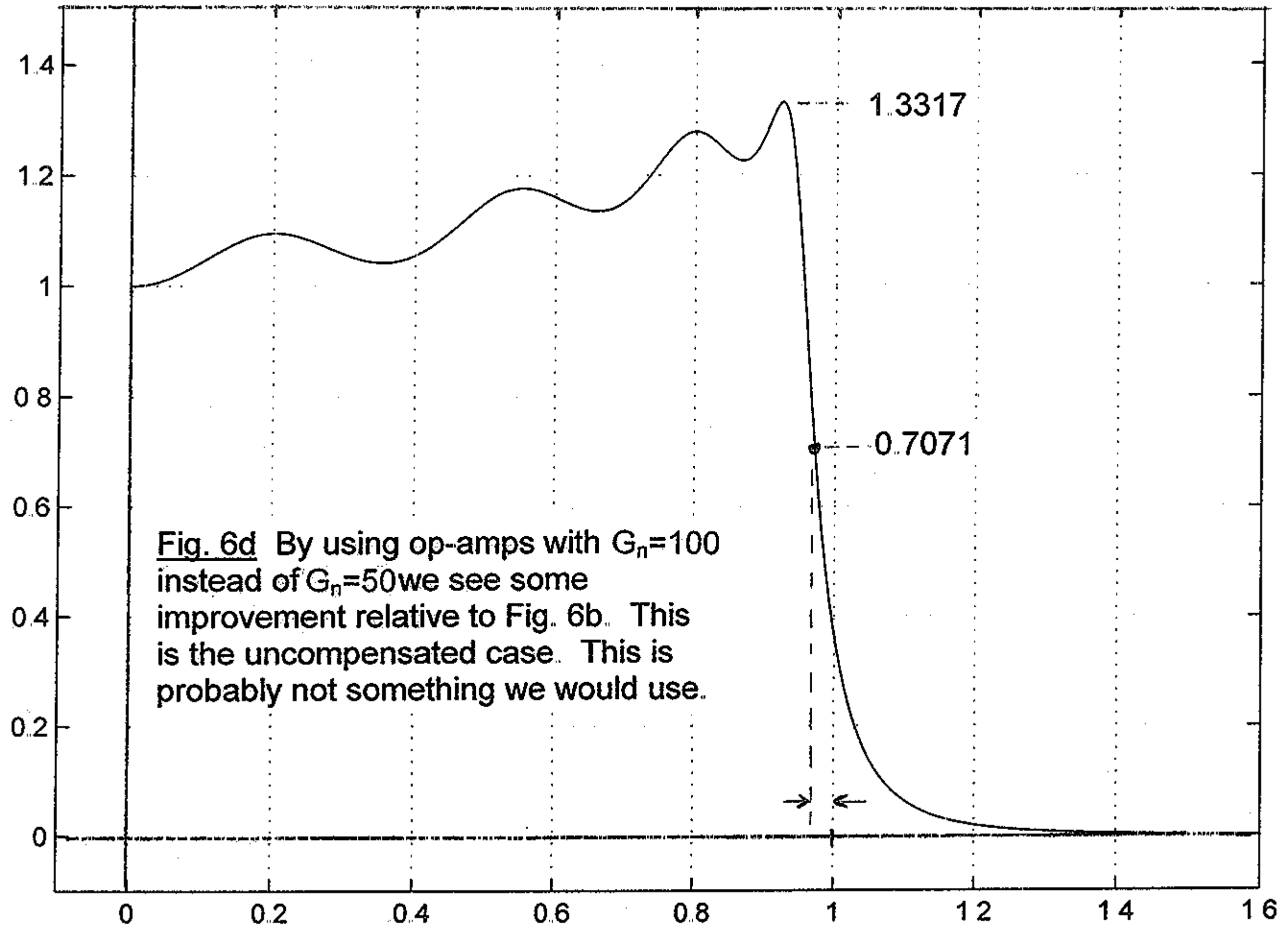



Fig. 6d By using op-amps with $G_n$=100 instead of $G_n$=50we see some improvement relative to Fig. 6b. This is the uncompensated case. This is probably not something we would use.

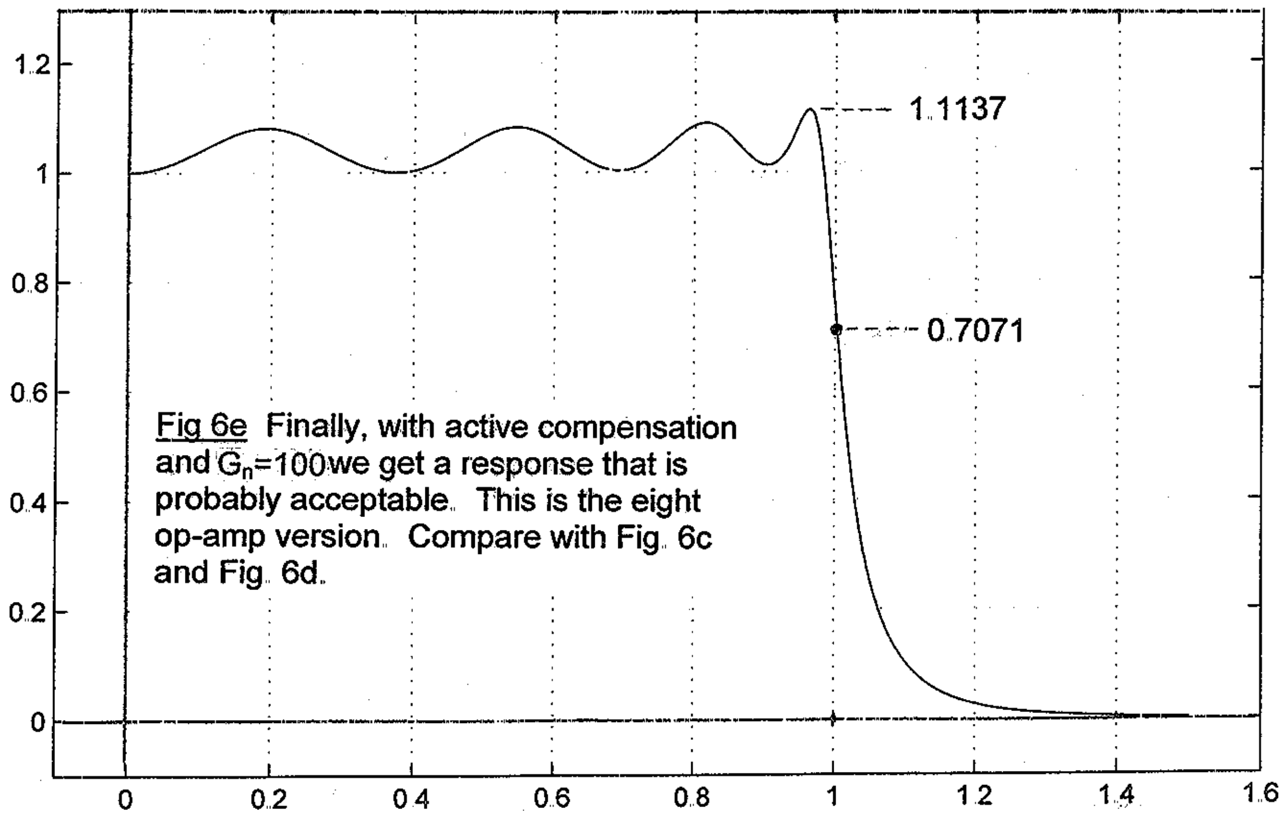



Fig 6e Finally, with active compensation and $G_n$=100we get a response that is probably acceptable. This is the eight op-amp version. Compare with Fig. 6c and Fig. 6d.

(relative to the 1.0812 that is designed in). So this works much better - about a 3% error While not shown, if we were to go to $G_n$=250, the rises are to 1.1852 (uncompensated) and to 1.0862 (compensated), again relative to 1.0812. This error is more like 1/2%. So we see that a very significant improvement can be achieved.

Yet the reader must be cautioned that calculations need to be made to be sure that the method is helpful. For some cases, while helpful, the improvement is not really enough (Fig. 6b and Fig. 6c). In some cases, for example if $G_n$ is only around 5 to 20, there may be no real improvement - the responses are different, but both are too bad to use.

# 5. COMPENSATION - THE POLE/ZERO VIEWPOINT

In Section 4 we looked at frequency responses as a way of usefully evaluating the active compensation method. Perhaps a more conventional view is the use of pole/zero plots: how much do nominal poles move, and what additional poles and zeros are now hanging around.

Consider for example the case of a 2nd-order Butterworth where the two nominal poles are:

$$p_2 = \{ -0.7071 \pm 0.7071j \} \qquad (9)$$

With $G_n$=20 for example, the real op-amp poles [$T_3(s)$] are now at:

$$p_3 = \{ -0.6252 \pm 0.6981j\ ,\ -14.3616 \} \qquad (10)$$

so we see that the nominal poles move considerably and we pick up an additional real poles (at -14.3616). In the case of active compensation [$T_4(s)$], we find four poles:

$$p_4 = \{ -0.7122 \pm 0.7128j\ .\ -7.0939 \pm 10.3128j \} \qquad (11a)$$

and now a zero appears at:

$$z = -12.6120 \qquad (11b)$$

In this compensated case, the nominal poles were far less disturbed, and we note that the two additional poles, along with the new zero, are a good approximation to the "magic" case (Fig. 7).

A conventional presentation of active sensitivity is that of displaying a single quadrant of the near-nominal poles for a range of values of $G_n$ [10]. Fig. 8a shows this view for our 2nd-order Butterworth example) where the nominal pole (of a pair) is

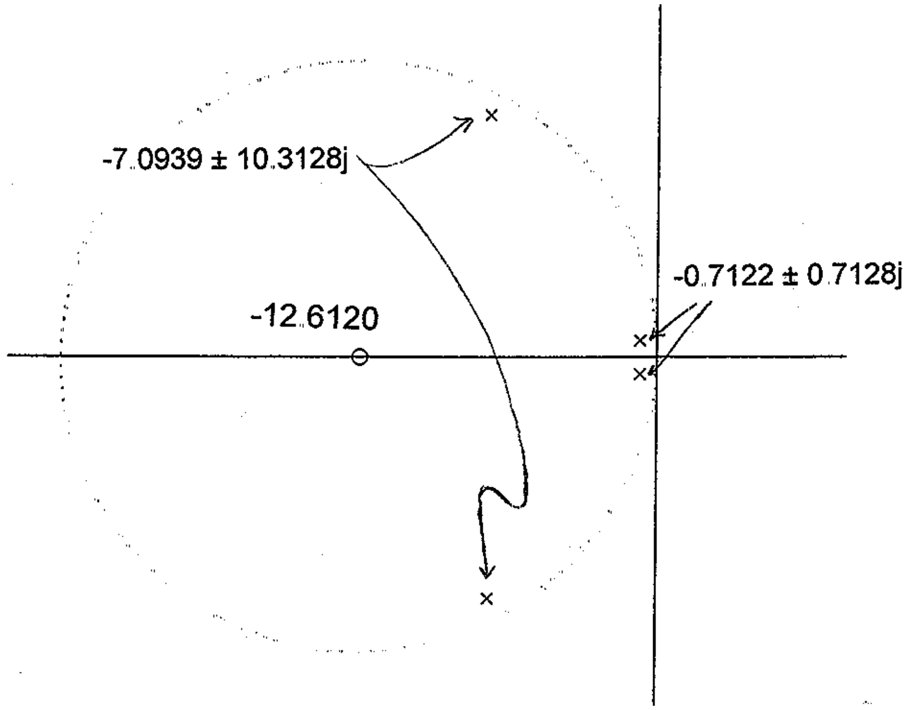



Fig. 7 Here we see the four pole one zero result using $T_4(s)$. We have two near-nominal poles and an array of two poles and one zero approximating the magic result.

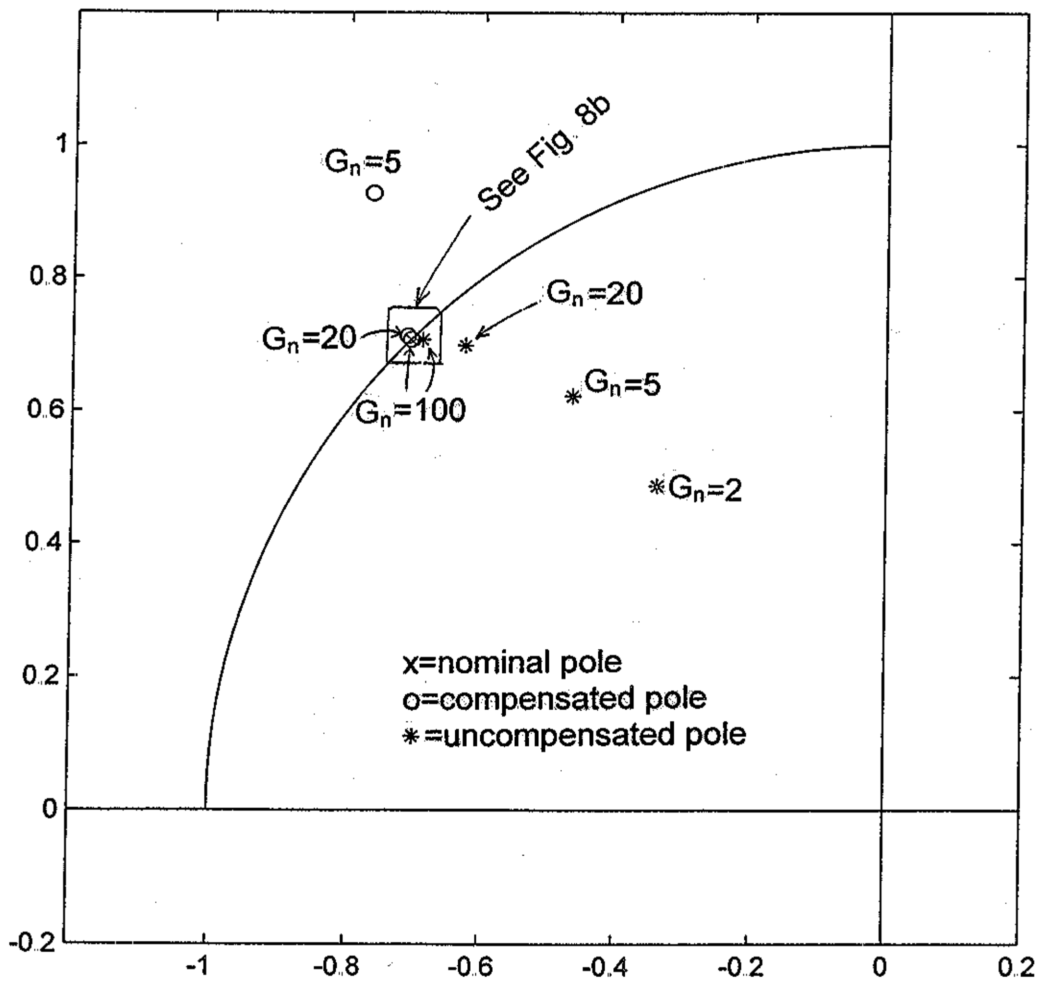



Fig. 8a A conventional plot showing active sensitivity for various values of $G_n$

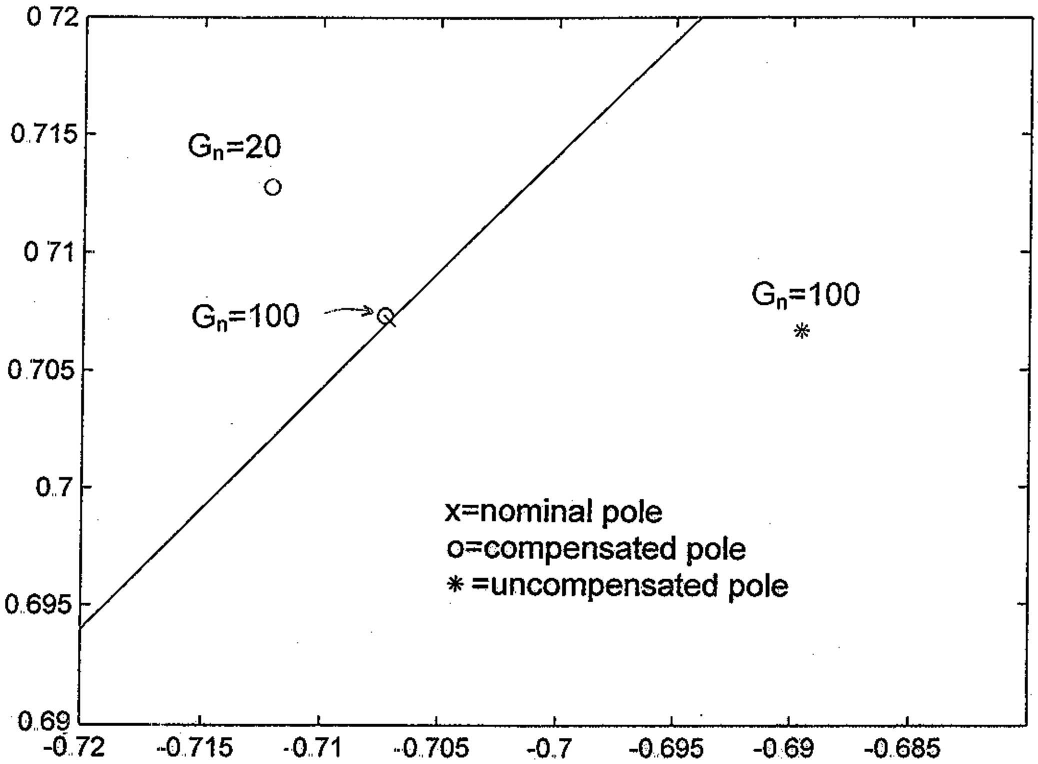



Fig. 8b A blow-up of Fig. 8a better shows the effectiveness of active cmpensation. Actually, no this scale, all three results are quite close to nominal (For 2nd order).

indicated by the x, the real op-amp pole is indicted by a '*', and the compensated pole by a 'o'. We note from Fig. 8a that the real op-amp results (which are inside the radius of the nominal poles) are entirely comparable with previous offerings [11,12] while the new compensated poles appear outside the radius of the nominal poles. From Fig. 8a we also notice that for $G_n$=2 and $G_n$=5 (not cases that we would be likely to try) all the results are poor. For $G_n$=20 and $G_n$=100, we see much better results, and we can see that active compensation is superior. A more zoomed-in view is provided by Fig. 8b. Note that $G_n$=20 compensated is superior to $G_n$=100 uncompensated (real op-amp). Also, $G_n$=100 compensated is very near nominal.

## 6. FREQUENCY RESPONSE OF THE "MAGIC" ARRAY

Previously we have praised the "magic" array (Fig. 3) for its near zero phase response at low frequencies [13]. As previously discussed [13], the magnitude response is not flat. Actually, it is quite flat for a while, but then peaks considerably, approaching1.5. This is seen in Fig. 9a where we have used the case of a zero at -1 and poles at $-0.5 \pm \sqrt{3}\, j / 2$. Other choices for the poles, all on the same circle shown, sometimes occur [14].

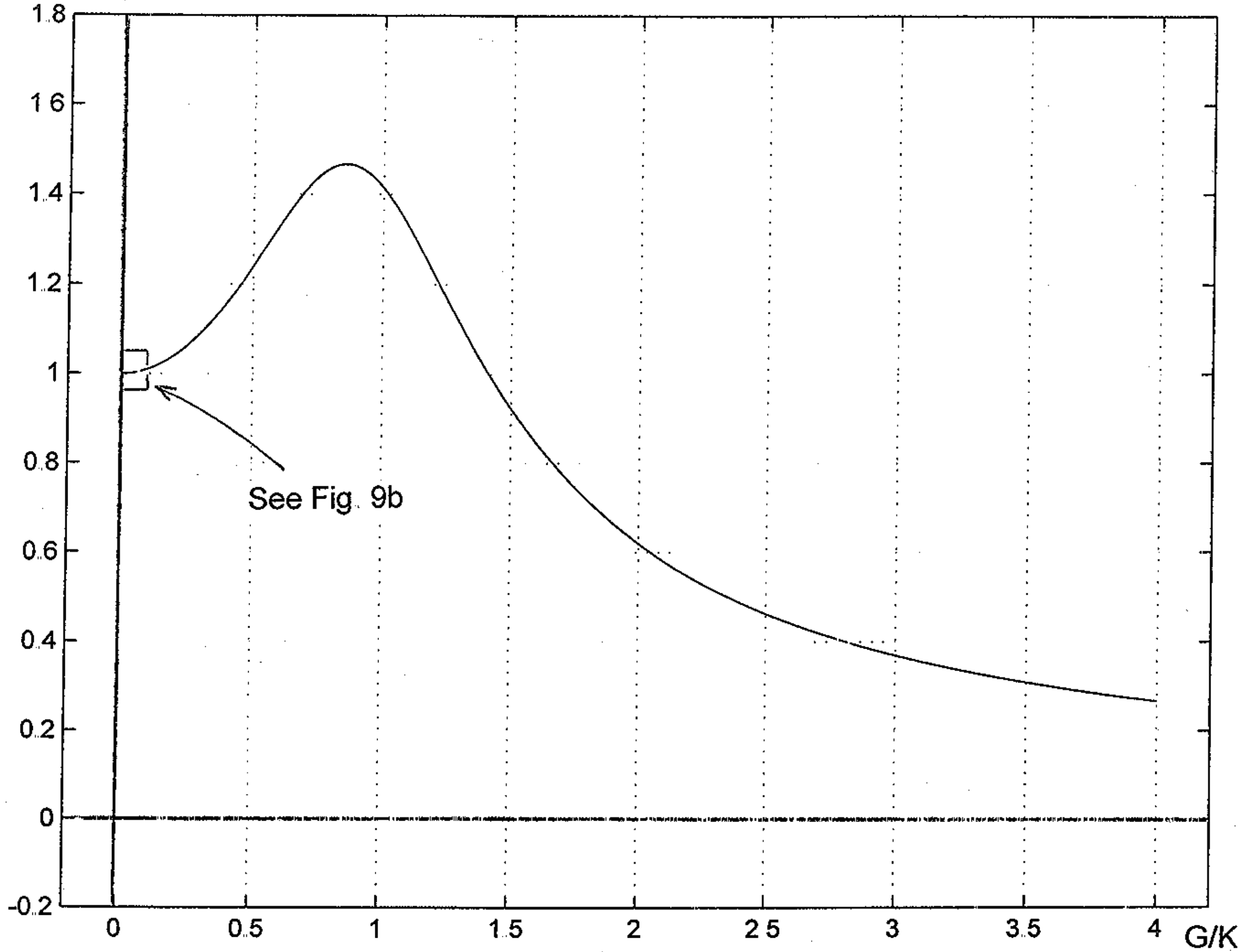
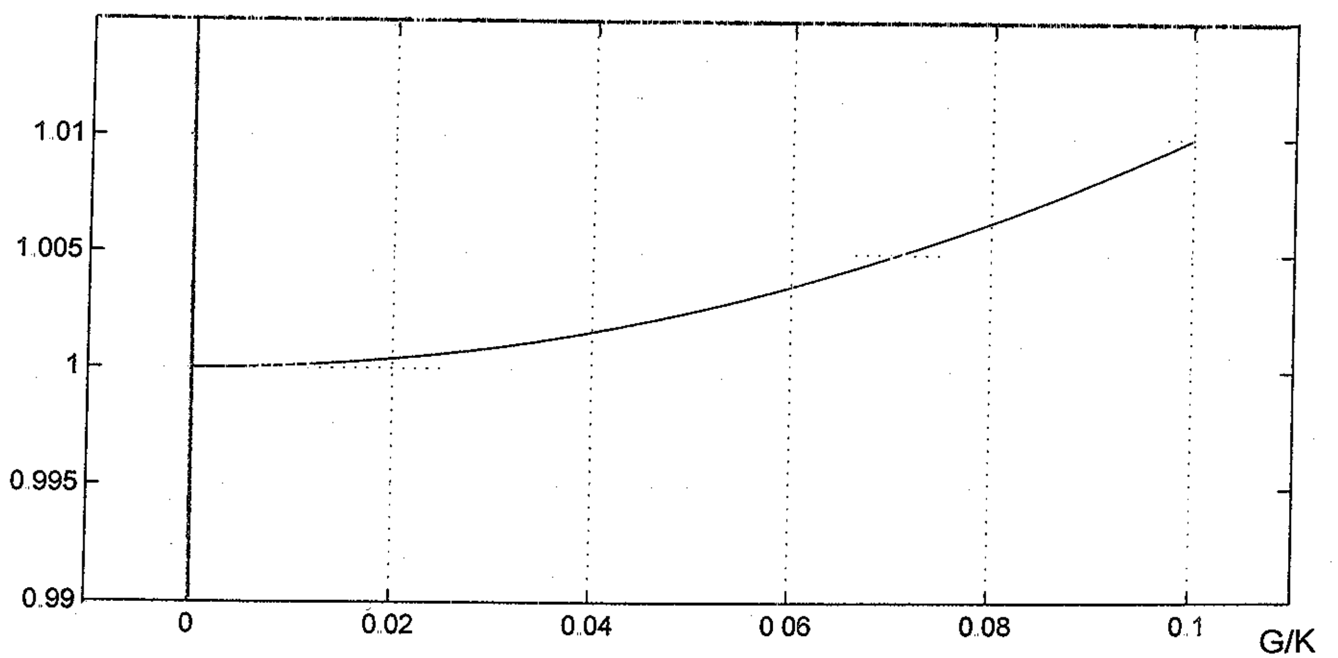



Fig. 9a The magic array has a not-so-magic bump reaching nearly 1.5. Yet we must keep in mind that the frequency is in very large units here - with 1 corresponding to the position of the zero (G/K in many of our examples). In Fig. 9b, (below) we see a zoom-in and we see that the peaking only reaches 1% at a normalized frequency of 0.1.



The important thing to note is that the peaking that approaches 1.5 occurs for a frequency around 0.8 times the G/K. That is, it is very high up, and unlikely to be a factor. Perhaps a more realistic view is provided by Fig. 9b which shows that for normalized frequencies from 0 to 0.1, the peaking only gets to 1% at 0.1.

## 7. CONCLUSIONS

Most electrical engineers who are involved with signal processing, even those who concentrate mainly in DSP, know about the Sallen-Key configuration. Further it is generally recognized that while Sallen-Key is not robust when pushed to the limits, it is nonetheless used with good success away from these limits. At times, this leads to tendency to push the applications too far when some filter specification is increased in order and/or in frequency.

What we have shown here is that when a design falls short of what is required, we can sometimes rescue it with active compensation and still keep our comfortable Sallen-Key network and its simple design equations That is, we do not have to change any component values, but just add op-amps and a few more resistors. This is likely to be a useful and pragmatic solution for some design upgrades, and may even be "piggy-backed" in when quantities of a particular piece of equipment are small.

It is a good idea to mentioning that these actively compensated circuits may look strange to some engineers who are not familiar with the actual purpose and the implementation. That is, not unlikely many engineers (assuming ideal op-amp notions), will come to the conclusion that your "improvement" does nothing but waste components. We mention this in case you need to defend your design, and also, having seen these ideas, you are unlikely to forget that there was a good reason for adding the components should you happen across them in someone else's circuit.


## ACKNOWLEDGMENT

The author appreciates the assistance of Boon-Ping Lau at Cornell University who verified many of the calculations here, and who verified some of the basic finding experimentally.

## REFERENCES


[1] Analog Signal Processing (ASP): Series presented in Electronotes Vol. 19, No. 191, Dec. 1999 through Vol. 20, No. 196, Dec. 2000. The material here pertains most to Chapter 7 of ASP in Vol. 20, No. 195, July 2000.

[2] Special Issue F: B. Hutchins, "Compensation of Linear Circuit Blocks for the Effects of Real Operational Amplifiers," Electronotes, Vol. 15, No. 164-167, July 1986.

[3] ASP, Chapter 7, Section 7-6 "Sensitivity of State-Variable Filters," Electronotes, Vol. 20, No. 195, July 2000 pp 29-32 (pertains to active sensitivity compensation by overdesign).

[4] B. Hutchins, "Active Sensitivity in a Compensated State-Variable Filer," Electronotes, Vol. 20, No. 195, July 2000, pp 2, 49-58 (pertains to compensated integrators and summers).

[5] Special Issue F, (Reference [2]), pp 49-55

[6] ASP, (Reference [1]), EN#195, pp 33-48

[7] B. Hutchins, "Working with Finite Gain-Bandwidth-Produce Op-Amps - 1, Amplifiers," Electronotes Application Note 222, July 4, 1981.

[8] ASP, (Reference [1]), EN#195, pp 46-47

[9] ASP, (Reference [1]), EN#195, pp 17-20

[10] ASP, (Reference [1]), EN#195, pp 17, 20, 24, 31

[11] B. Hurchins, "Working with Finite Gain-Bandwidth-Produce Op-Amps - 2, Sallen-Key Filters," Electronotes Application Note 223, July 11, 1981.

[12] ASP, (Reference [1]), EN#195, pg 24

[13] Special Issue F, (Reference [2]), pp 51-55

[14] Special Issue F, (Reference [2]), pg 59

# THE CURIOUS AMBIGUOUS LENGTH OF A TRIANGULAR WINDOW

-by Bernie Hutchins

## Triangles!

Recently in supervising some student projects for a poster presentation in a DSP course, I had temporary custody of some blank poster boards in my office while students went to classes nearby. To my amusement, these boards had diagrams of suggested projects on the back. One of them was for a dinosaur project. I wondered if some student might become dissatisfied with his or her DSP project and flip the board over and go with the example provided! A second one had a math project on triangles as the example layout. Now, triangles are important in DSP (not just more important than dinosaurs) in that they are convolutions of rectangles, the impulse response of linear interpolation filters, and a simple class of data windows. This and a few other things brought to mind a neglected project of my own concerning triangular windows, and the results of finishing this are presented below.

## THE PROBLEM

Suppose we have a discrete-time rectangular window of, say, length 7. For example, we might well imagine it to be:

$$r(n) = \begin{cases} 1 & n=-3,-2,-1,0,1,2,3 \\ 0 & \text{all other } n \end{cases} \qquad (1)$$

How long is this window? Well - length 7! Simple enough. Now, what if we had a triangular window of length 7. We might think of:

$$t_7(n) = \begin{cases} 1/4 & n=-3 \\ 2/4 & n=-2 \\ 3/4 & n=-1 \\ 4/4 & n=0 \\ 3/4 & n=1 \\ 2/4 & n=2 \\ 1/4 & n=3 \\ 0 & \text{all other } n \end{cases} \qquad (2)$$

Again this seems simple enough. But there is an argument waiting here. Isn't this really a length 9 triangular window? After all, the actual triangle, from which our samples are derived, goes to and reaches exactly 0 at n=-4 and n=+4 (Fig. 1). Perhaps we should consider these two ends as part of the length. Of course, we could continue this game and claim it is a length 99 window, etc., simply be adding more zeros, although this becomes silly. But we might logically argue that the actual window simply does not exist outside of the range of the generating triangle. If so the only legitimate values for the length about which we care to debate seem to be 7 or 9. It may seem a philosophical point at best.

Here we will argue that the window is length 8. Let's see why this might be so. Fig. 2 shows how the exact same triangle used in Fig. 1 can be sampled to give a length 8 window. Here we sample as:

$$t_8(n) = \begin{cases} 1/8 & n=-7/2 \\ 3/8 & n=-5/2 \\ 5/8 & n=-3/2 \\ 7/8 & n=-1/2 \\ 7/8 & n=1/2 \\ 5/8 & n=3/2 \\ 3/8 & n=5/2 \\ 1/8 & n=7/2 \\ & \\ 0 & \text{for all other samples} \end{cases} \tag{3}$$

Actually we have sampled at half integers, but clearly we have eight non-zero values instead of just seven. Clearly, except for samples exactly at the integers, we always get eight non-zero samples (Fig. 3). [ In $t_8(n)$ above, as in $t_7(n)$, we have symmetric samples. All other sampling choices are non-symmetric (although length 8).] So we might like to say that $t_7(n)$ is neither length 7 nor length 9, but really length 8 like its infinite other cousins. While this is "cute," there is a bigger issue.

If in the curious case $t_7(n)$ is really length 8, it must have seven zeros. Of course, $t_8(n)$ already has seven zeros. We can easily find the zeros for any particular sampling by using any available root finder program on a case-by-case basis. For example, the six zeros of $t_7(n)$ as a length 7 window are at:

$$z = j,\ j,\ -1,\ -1,\ -j,\ -j \tag{4}$$

The seven zeros of $t_8(n)$, which is length 8, are at:

$$z = j,\ j,\ -1,\ -1,\ -1,\ -j,\ -j \tag{5}$$

Thus $t_8(n)$ has an additional zero (three total instead of just two) at z=-1 (Fig. 4).

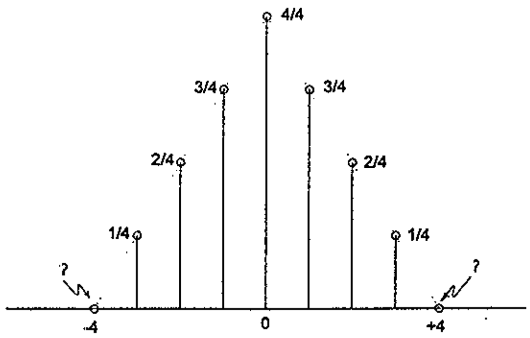



Fig.1 Length 7 $t_7(n)$

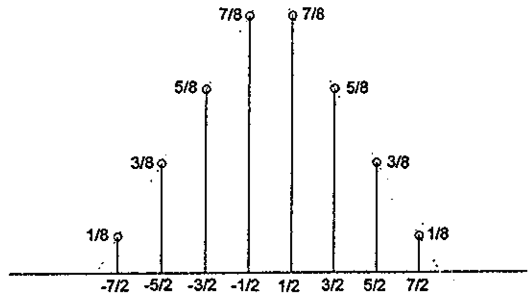



Fig. 2 Length 8 $t_8(n)$

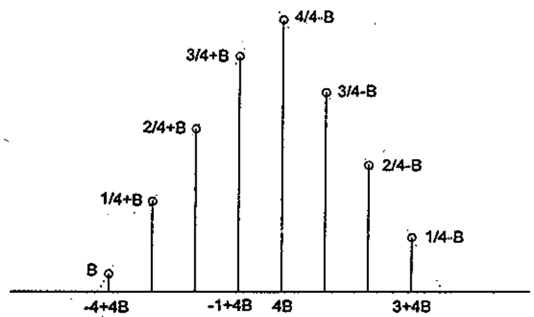



Fig. 3 Length 8, Offest Start

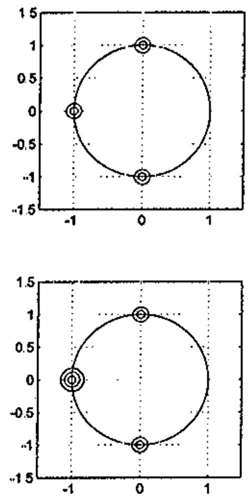


Fig. 4 Zeros: Length 7 (top)
Length 8 (bottom)

There are a lot of questions here. Where is the seventh zero of $t_7(n)$? Must it not be at infinity or at zero? What about all the other length 8 samplings of the triangle? Where is their extra zero? And - does any of this make any difference? Since all these are just sampling of the same triangle (Fig. 3), aren't the frequency responses (Fourier transforms of the windows) the same? Sampling theory suggests that the exact starting times for the samples aren't suppose to matter. If the starting time matters - why? A lot of questions.

Let's look at the last questions first. Although it may go without saying, the extra zero at z=-1 that is present in $t_8(n)$ makes a difference. (Indeed it is the same extra zero that always appears with even length, even symmetry.) This gives us a counterexample. Thus the starting time matters and sampling theory does not apply. Fig. 5 shows the responses involved, and in fact $t_8(n)$ seems to be a significantly better triangular window in that its sidelobe is much smaller.

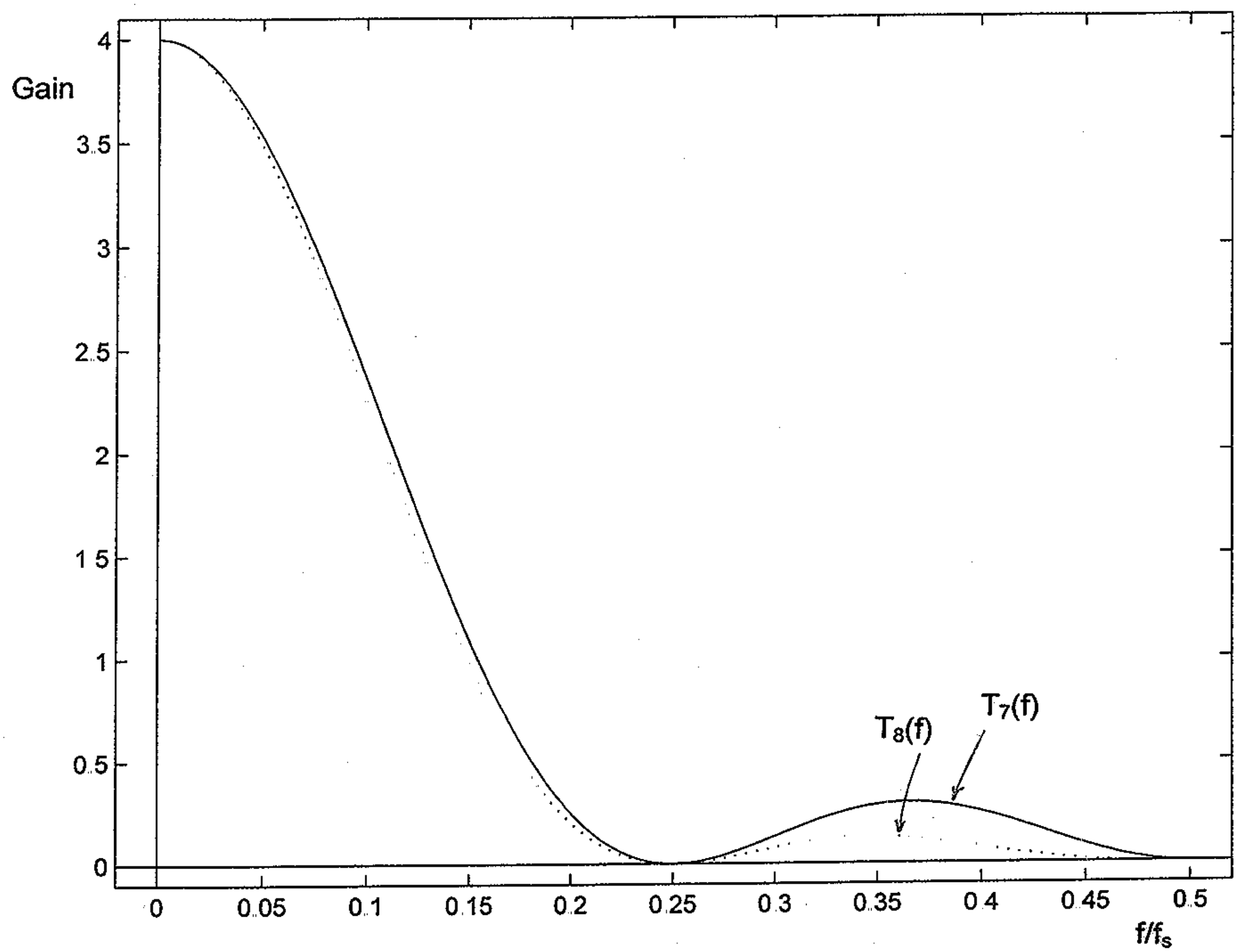



Fig. 5 The DTFT's of the two different samplings of the triangle are different. Note that $T_8(f)$ has a significantly smaller sidelobe.

## A BANDLIMITING FAILURE

Now, the reason why sampling theory does not apply to this problem is that the signal being sampled, the triangle, is not bandlimited. We knew this - we just forgot! That is, while $t_7(n)$ and $t_8(n)$ are samples of the exact same triangle, they are not samples of the same bandlimited signal. To illustrate, Fig. 6 shows the bandlimited function (bandlimited to a frequency of 1/2, the sampling rate being 1) that actually goes through the seven non-zero points of $t_7(n)$, and indeed, through zeros for all other sample points. This we find by convolving the samples with interpolating sinc functions [1]. Fig. 7 shows this same function with more information added.

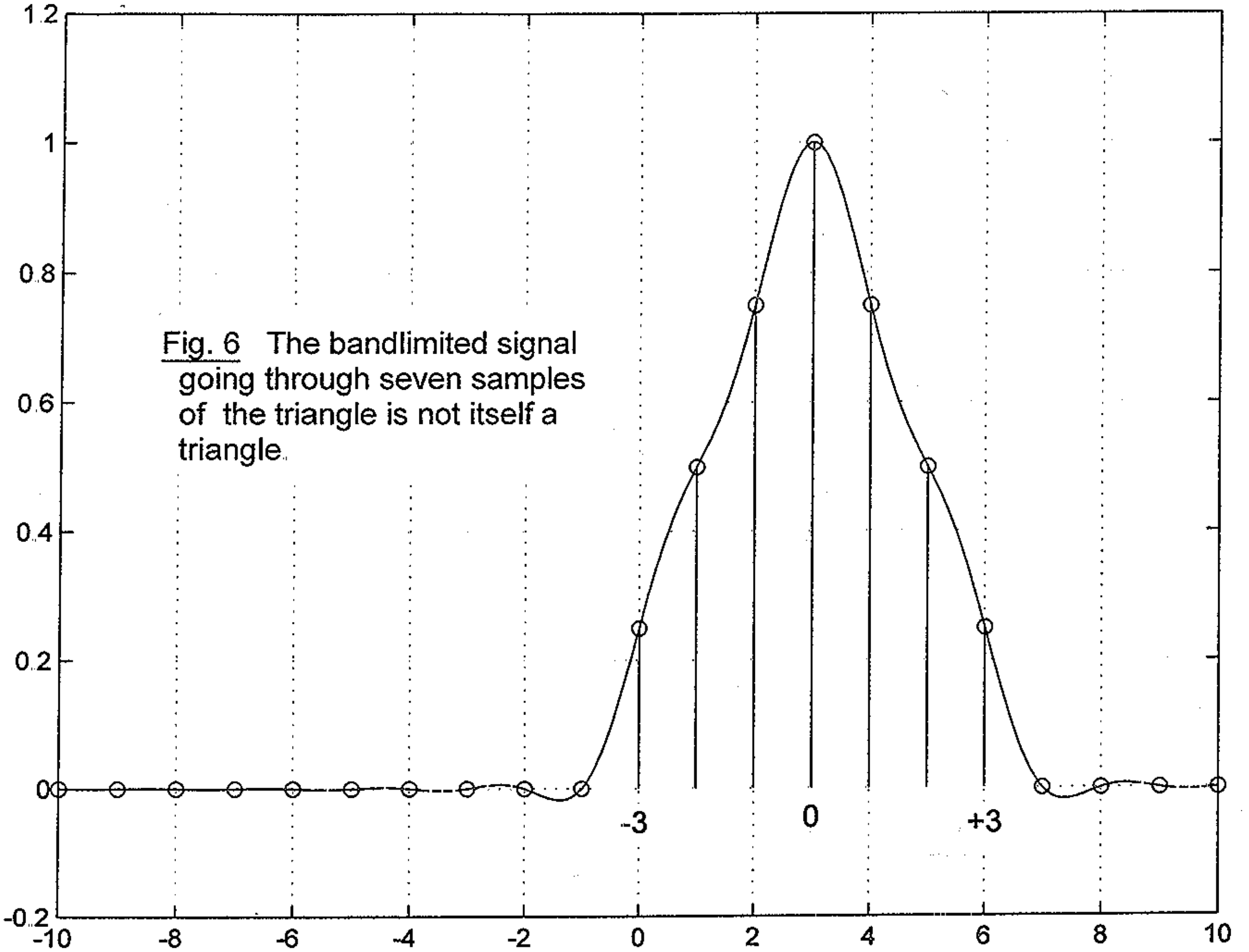



Fig. 6 The bandlimited signal going through seven samples of the triangle is not itself a triangle.

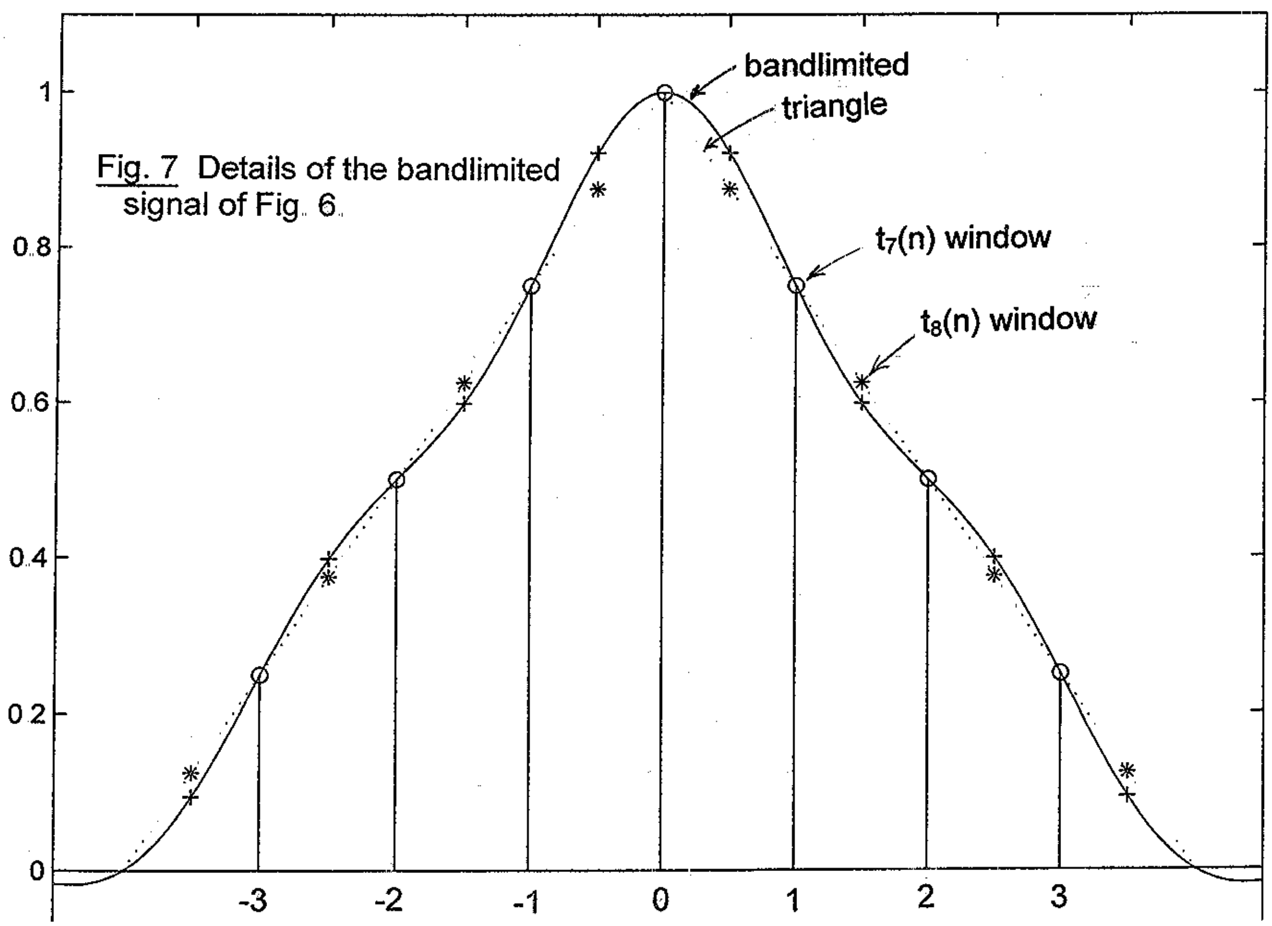



Fig. 7 Details of the bandlimited signal of Fig. 6.

In Fig. 7, the solid curve is the bandlimited function, while the dotted curve is the triangle. Note that the seven window samples of $t_7(n)$ (circles on top of a stem) are the intersections of both these curves. Indeed we thought of the circles as being the result of sampling the triangle, and we used these samples to find the bandlimited curve that went through the same points. Now, in taking the $t_8(n)$ samples, we shifted the sampling times by 1/2 (the $*$ in the triangle in Fig. 7). We note that these are different from the samples at the same points in time that correspond to the bandlimited curve (the points marked with a +).

Fig. 8 is an extension of our presentation and is similar to Fig. 7 although it is a little more cluttered. Here we begin with the $t_8(n)$ window samples (circles atop stems) and find the bandlimited curve that goes through these. Note well that this is not the same bandlimited curve that went through the $t_7(n)$ points. [In fact, it is significantly smoother and is barely different from the triangle. This can be understood by considering that the difference between the triangle and the bandlimited function in Fig. 7 is principally the third harmonic. Fig. 5 shows significantly less third harmonic energy (vicinity of f=3/8 relative to a f=1/8 fundamental - eight samples per cycle of the triangle) for the $t_8(n)$ window relative to the $t_7(n)$ window. These same conclusions about harmonic content also follow easily by taking the length-8 FFT's of $t_7(n)$, padded with one zero, and of $t_8(n)$.] Continuing with the procedure we now take samples of the bandlimited curve (marked with a +) and we see that these are not the same as the $t_7(n)$ window (marked with a $*$) as is most evident near the middle and at the ends.

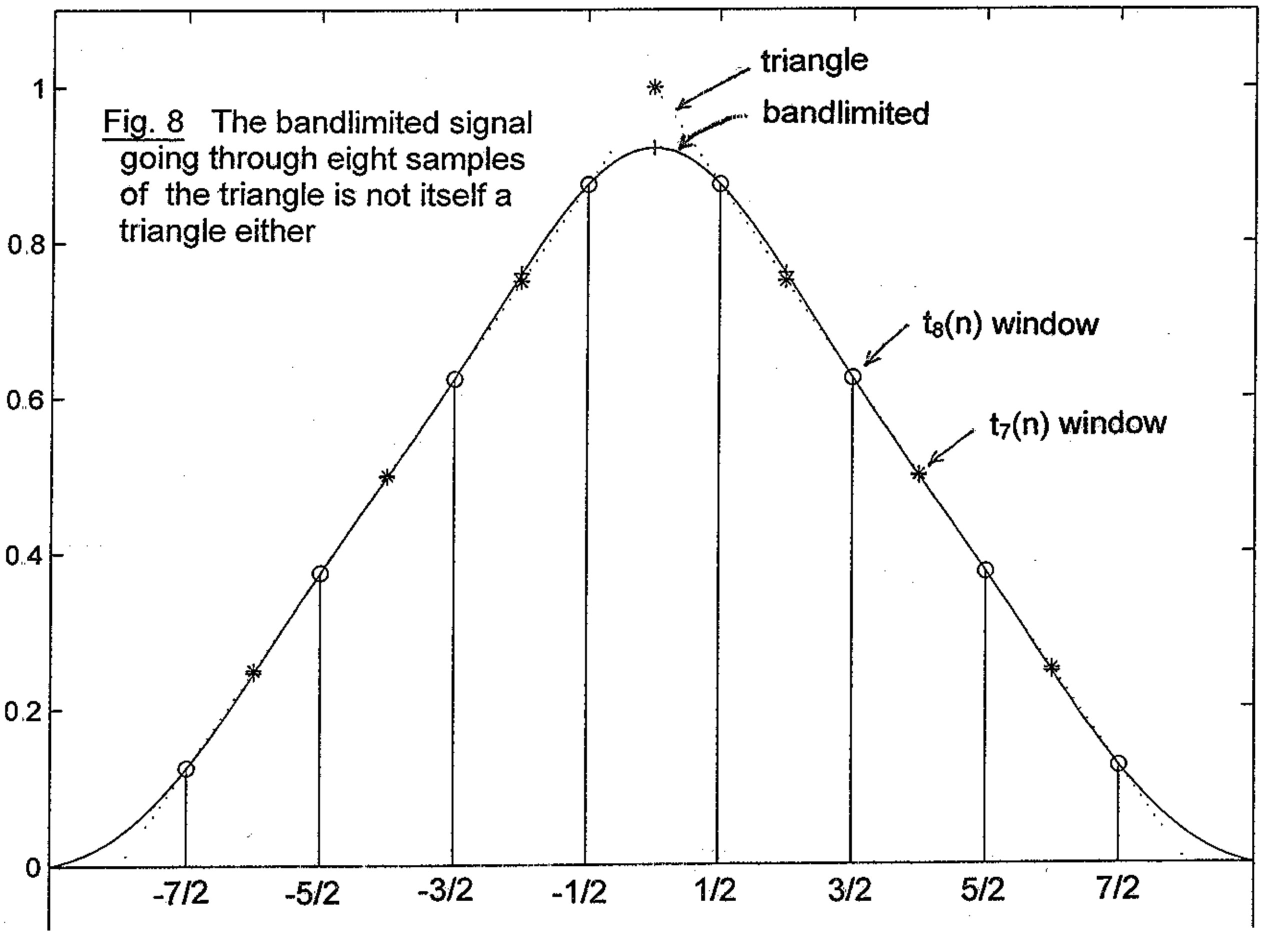


Fig. 8 The bandlimited signal going through eight samples of the triangle is not itself a triangle either

## FINDING THE EXTRA ZERO

Finding the roots of polynomials of order greater than two is trivial when done numerically, and usually impossible without numerical help. Thus we find it trivial to work with many different samplings of the triangle as in Fig. 3, and compute the roots (the zeros). We find that there are seven roots, six of which are at j, j, -1, -1, -j, and -j while the seventh is somewhere else. To understand this better, and to in this case arrive at a closed-form formula for the seventh zero, write down the polynomial from the zeros. Assume that the seventh root is at z=p.

$$A\,T(z) = A(z-j)(z+j)(z-j)(z+j)(z+1)(z+1)(z-p)$$

$$= Az^7 + A(2-p)z^6 + A(3-2p)z^5 + A(4-3p)z^4 + A(3-4p)z^3 + A(2-3p)z^2 + A(1-2p)z - Ap \quad (6)$$

Here we have multiplied out all seven factors of the form $(z - z_x)$ where $z_x$ are the zeros, but we have also added an overall multiplier A. This is necessary as the roots of a polynomial identify the polynomial only up to an arbitrary multiplicative factor.

We note that we have eight coefficients that are functions of A and p, and that these are weights of the samples of the window. We know these samples for any choice of sample timing. For example, if we have the $t_8(n)$ window, we have at the ends (the $z^7$ and the $z^0$ terms):

$$A = 1/8 \quad (7a)$$

$$-Ap = 1/8 \quad (7b)$$

which leads to A=1/8 and p=-1. There are other equations of course that lead to the same answer. For example, for the $t_8(n)$ window, the $z^6$ and $z^1$ terms give:

$$A(2-p) = 3/8 \quad (8a)$$

$$A(1-2p)=3/8 \quad (8b)$$

which again gives us p=-1 and A=1/8. The redundancy in the equations is a consequence of our choice of the triangle in the time domain, or the zero pattern in the frequency domain. The problem is highly structured. Note that if we look at $t_7(n)$, either the first or the last coefficient must be zero. The first, A, cannot be zero or the polynomial disappears. Hence p must be zero.

For a general sampling of the triangle we have a sample B for one end and a corresponding sample (1/4 - B) for the other (Fig. 3). Again, using the ends:

$$A = B \tag{9a}$$

$$-Ap = (1/4 - B) \tag{9b}$$

These equations give the zero p as:

$$p = 1 - 1/4B \tag{10}$$

which is the general answer (Fig. 9).

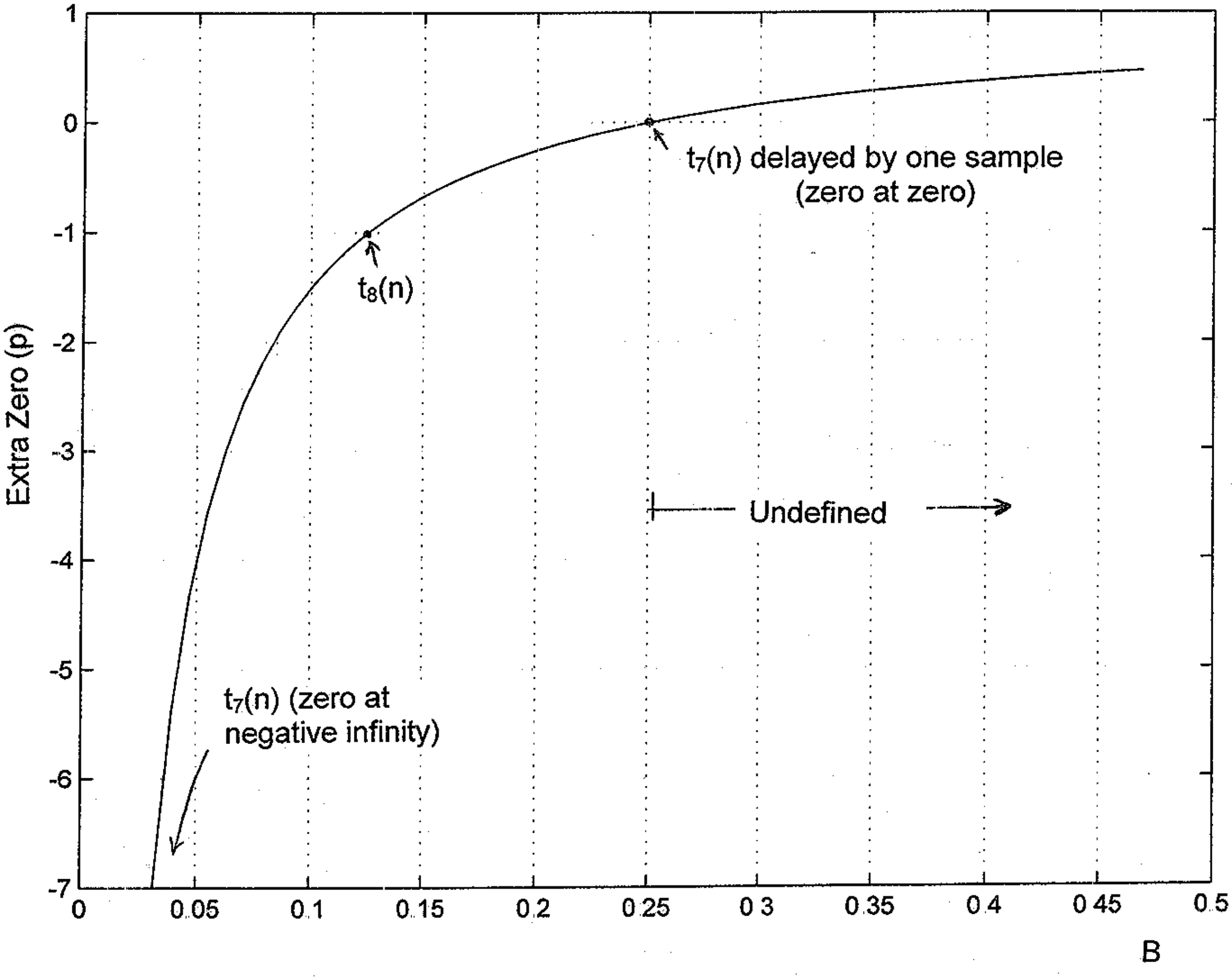



Fig. 9 For values of B of 0, 1/8, and 1/4 (se also Fig. 3), we find corresponding zeros at negative infinity, at -1, and at zero, which in turn corresponds to $t_7(n)$, $t_8(n)$ (Fig. 4), and $t_7(n)$ delayed by 1. We find that the useful range of Fig. 9, a plot of equation (10), is for B = 0 to B=1/4. Beyond this, the results repeat periodically.


REEFERENCE [1] B. Hutchins "The Importance of the Notion of Bandlimiting to Sample Reconstruction, Electronotes, Vol. 18, No. 184, pp 3-20

BOOK REVIEW - ANALOG DAYS (Continued from page 2)

For me, and for the long time readers of this newsletter, the book divides historically into three parts: prior to the mid-60's, the late 60's until the mid-80's, and beyond the mid-80's. The middle period is the one we know best: the (seemingly) halcyon "golden era" where new ideas and products seemed to infiltrate an eager marketplace. But as much fun as this era was, it was not an easy time, as this book relates, although it is doubtful many of the participants would have taken the option to opt-out. It was worthwhile doing, and even though mistakes were made, it was a proud era.

The earliest time period tells us mainly what Bob Moog was doing, and emphasizes that there certainly was no "master plan" as we often suppose, looking back at what we view as a reasonable success. In fact, it was pretty disorganized at times, and as is often the case, things happen by accident. But it does answer some interesting questions of the "Who thought of doing....?" type. Someone would express the need for something (usually in imperfect terms), someone else would run across the street for a doorbell switch, and presently major conceptual pieces of the puzzle came into focus, leaving everyone to wonder if it should not have been obvious all along.

The last era (post-golden-era) is perhaps the least known and is in some ways a bit distressing. This is not just a story of the "best laid plans ....." going wrong (indeed as we note, the plans were often non-existent), but also of a grand vision caving in to the reality of the marketplace and of less admirable ideas about the place of the synthesizer relative to a continuing musical art: big companies taking over and producing lowest-common-denominator (almost toy) machines; the synthesizer producing commercial jingles but not concert music; a preference on the part of the music-consuming public for prosaic results that did not challenge the synthesizer's capabilities.

The book gets into few technical details. There are a few technical misconceptions that do not detract much from the story being told.

Trevor and Frank have written a great book, from which we can learn a lot, and it gives us much to be proud of. And its fun too. Most highly recommended.

-Bernie Hutchins

*********************************************************************

ELECTRONOTES, Vol. 21, No. 202 (August 2003)

Published by B. Hutchins, 1016 Hanshaw Rd., Ithaca, NY 14850 (607)-257-8010