ELECTRONOTES 199

NEWSLETTER OF THE

MUSICAL ENGINEERING GROUP

1016 Hanshaw Rd., Ithaca, NY 14850

Volume 20, No. 199 September 2001

GROUP ANNOUNCEMENTS

Contents of EN#199

In this issue, we conclude the presentation on Digital Filtering - the Basic Elements of Digital Signal Processing. Next issue we expect to move on to sampling. In the analog corner this issue, we consider the compensation of a second-order network for inherent gain.

Analog Signal Processing Corner

Reducing the Accumulating Gain - via Thévenin

-by Bernie Hutchins

As an instructor of electrical engineering, I often wonder at what point in our instructional program we manage to drive at least some good students away from the field. I suspect it is often when we present them with formal methods and apply them to highly contrived and exceptionally dull examples. Students come with some basic knowledge of electronics and computers, and often substantial knowledge of specific applications and products. They want to know about CD players, about image processing, about artificial intelligence, and that sort of exciting things Instead we give them complicated networks of many resistors and ask them to set up loop and/or node equations and solve them. Perhaps the goal is to show that the whole mess is equivalent to just one or two resistors. (Thévenin and Norton equivalents, etc.) Not too hard, but it sure doesn't look like anything real. In consequence, when in later courses (if they haven't left) when we actually show that something real and practical can be understood by "dropping the name" of something formal, they seem shocked. They had hoped to never hear the term again, and had supposed it only applied to something theoretical.

In analog filter design, it is frequently the case that the filter employed has a gain that is not what we want, and may even be intolerably high or low. In EN#192 (8) which is part of Chapter 3 of Analog Signal Processing we have one of our "see problems at end of chapter" for which we do not give the actual problem, and which we have been attacking individually in this Analog Signal Processing Corner in recent issues. There and in some examples [EN#192 (25)] we want to replace a resistor in the network with two resistors in order to reduce the gain. Generally when I discuss this in class, I can just say "This is just the same thing as the Thévenin equivalent - which you remember" and get away with it. In the general case, the students remember just enough about Thévenin equivalents to immediately grasp what we are doing (they seem amused that we have reached a practical example).

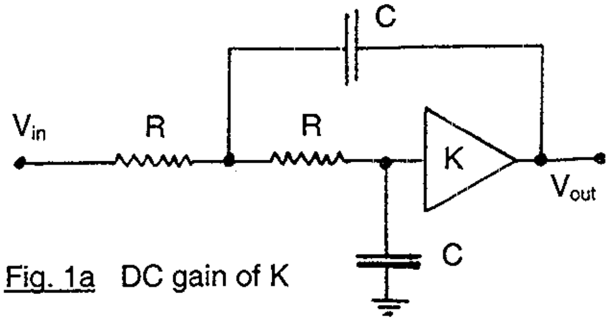


Fig. 1a DC gain of K

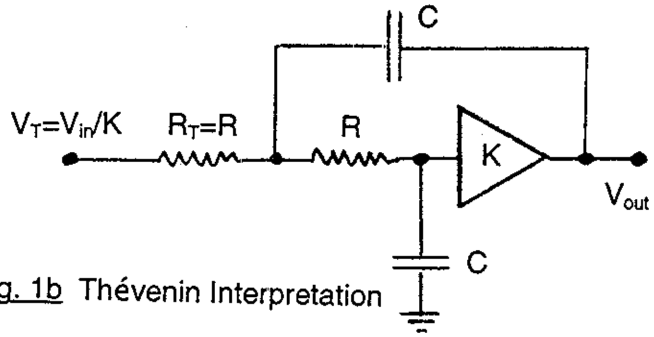


Fig. 1b Thévenin Interpretation

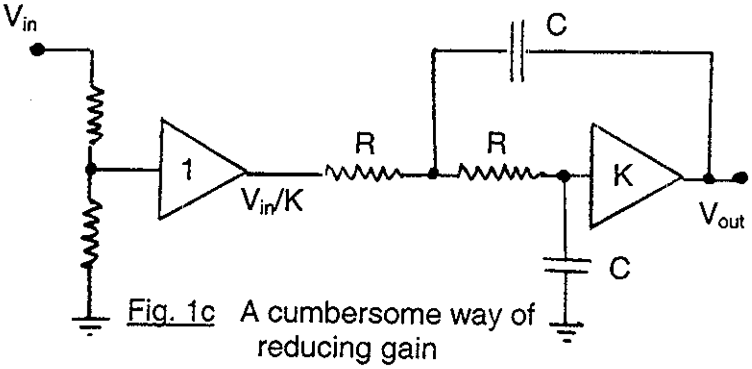


Fig. 1c A cumbersome way of reducing gain

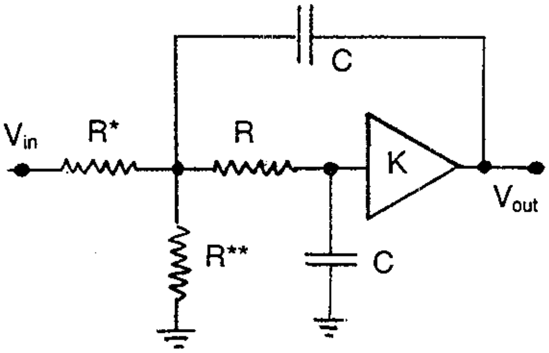


Fig. 1d A convenient network for reducing gain

Fig. 1a shows a conventional Sallen-Key low-pass which has a dc gain of K, where we expect that K is between 1 and 3 in practical cases. Particularly in cascade arrangements, these section gains can multiply to an overall gain that is so large we may be "clipping" against the power supply by the time we reach the output. To avoid this, we might think of arranging for the input to be of a compensating lower level. But this may result in problems with signal-to-noise (signal level decreased against a constant noise level) or with offset and drift (an offset of 1 mV is not much with a signal of 1V amplitude, but would likely be a problem if the signal were say, 3 mV).

(continues on page 32)

5. SPECIAL TOPICS IN FILTER DESIGN

5a. TYPE TRANSFORMATIONS

5a-1 Introduction

For the most part, the examples we have used above have involved low-pass frequency responses. In fact, low-pass filters are likely the most used in practical applications, and there is a tradition in filter studies to derive other types of filters from low-pass "prototypes." On the other hand, for most of the design methods we have looked at, it is not at all difficult to modify the input specifications so that high-pass, band-pass, and other desired responses are obtained directly. [Some responses are prohibited: for example, we can't have an IIV digital filter design that does not begin with an analog prototype that does not go to zero at infinity. Shortly we shall see that even length FIR filters that do not go to zero at half the sampling frequency should be avoided.]

Because we often prefer to just plug in the actual specifications, our interest in transforming one type of digital filter to another is more limited. For one thing, considerable insight into filters can be obtained by studying the relationships between different types. For another, it may be possible to easily or trivially implement a second needed response once one filter is in place.

5a-2 Methods of Type Transformation

When we consider the ways in which a low-pass filter might be converted to a band-pass or a high-pass, a number of possibilities come to mind. First, it might seem trivial to obtain a high-pass by subtracting a low-pass response from a constant. (When we consider the issue of phase however, the picture may become much more complicated.) Another simple idea would be to flip the response about half the sampling frequency somehow. (If the high-pass cutoff does not end up where we wanted it, we would likely be able to just choose a different "prototype" low-pass cutoff.) A third method (a generalization of this flip actually), would be a rotation of the response in the frequency domain, often thought of as a "modulation," usually obtained through the time-domain multiplication of the impulse response by a sinusoidal sequence.

Fig. 44 and Fig 45 illustrate a good number of the ideas that are most important here. The filters of Fig. 44 are length 25 (odd) while those of Fig 45 are length 26 (even). Fig. 44a shows the impulse response and the magnitude of the frequency response for a length 25 low-pass with a cutoff at 0.15 (relative to a sampling frequency of 1). Here we have used a least squared error method - any similar method would illustrate our points here. Glancing over at Fig. 45a, we see a very similar-looking length 26 filter.

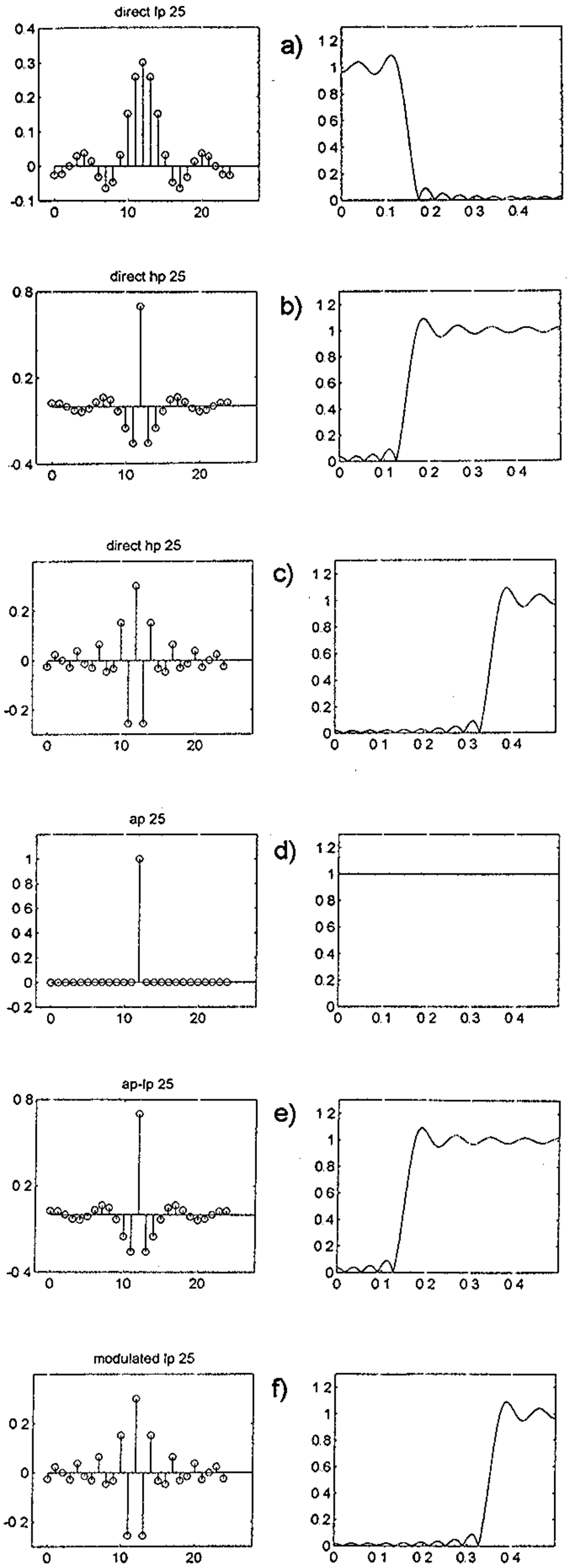


Fig. 44 Length 25 Filters

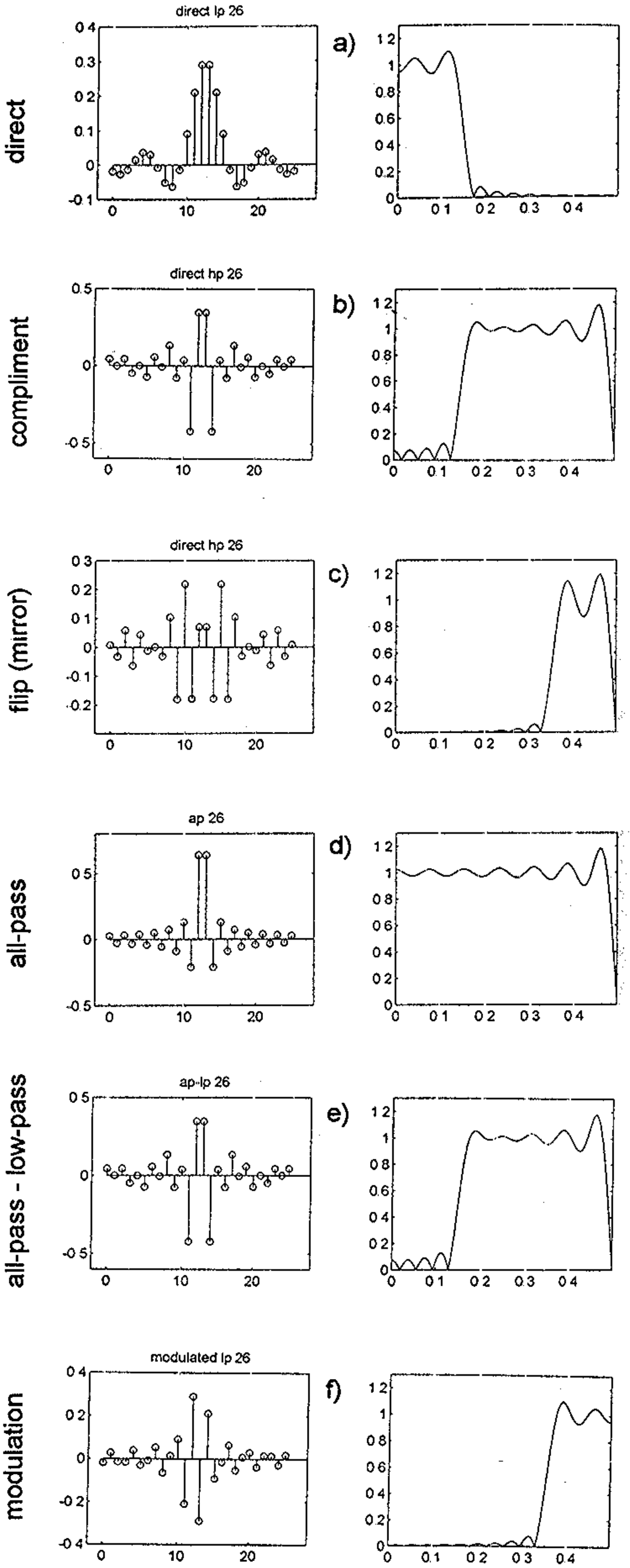


Fig. 45 Length 26 Filters

Fig 44b and Fig. 44c show the two notions of a "corresponding" high-pass. Fig. 44b can be thought of as a compliment (its cutoff is also 0.15) while Fig. 44c is the "flipped" version (its cutoff is 0.35). These two modifications were obtained simply by changing the input specifications to the design program. Results identical to Fig 44b and Fig. 44c are seen in Fig 44e and Fig. 44f, respectively. These two re-derivations illustrate methods of conversion of the Fig. 44a low-pass into these high-pass results.

With some lack of caution, we might suppose that we could write:

$$H_H(e^{j\omega}) = 1 - H_L(e^{j\omega}) \qquad (81a)$$

This is perfectly correct, but we must be careful not to suppose that understanding this is merely a matter of considering the magnitude plots shown. In particular, we have seen that it is necessary when specifying a filter's frequency response to consider the phase as well as the magnitude. What this means is that there is a phase associated with $H_L(e^{j\omega})$ and we must choose the same phase for the "1" in equation (81a), and will end up with the same phase for $H_H(e^{j\omega})$. This need not be difficult in all cases. In the FIR examples here, the "1" has the same linear phase as the low-pass. This is merely the delay at the center of the impulse response. Accordingly we design the all-pass filter of Fig. 44d: just a delay of 12 units (the magnitude response is of course just 1). Corresponding to equation (81a) we have the time-domain equation.

$$h_H(n) = \delta(n-12) - h_L(n) \qquad (81b)$$

This we see by study of Fig. 44a, Fig. 44d, and Fig. 44e.

Fig. 44f was obtained from Fig. 44a simply by changing the sign of every other term of the impulse response (with the timing set here, the center tap is not inverted). One simple way to look at this is as a modulation. The time-domain sequence is multiplied by the sequence $(-1)^n$, which is a sequence corresponding to the sampling of a cosine of half the sampling frequency. Accordingly, the low-pass response (centered at 0 frequency) is convolved with a "spike" at half the sampling frequency, shifting it to be centered at half the sampling frequency, resulting in the high-pass.

Alternatively, and more generally, we observe that through the use of the z-transform relationship, we can find a "flipped" version of a transfer function, H(-z) as:

$$H(-z) = \sum_{n=-\infty}^{\infty} h(n)\,(-z)^{-n} = \sum_{n=-\infty}^{\infty} (-1)^n\,h(n)\,z^{-n} \qquad (82)$$

which equates alternating the signs in an impulse response with a flipping of the features (poles and zeros) in the z-plane, about the imaginary axis, relative to the usual notion of frequency as one goes around the unit circle.

So far, we only considered Fig 45a, the successful length 26 low-pass It might seem that the same findings from Fig. 44 would again appear. However, the even length case makes a difference. In particular, any even length, even symmetry, FIR filter must have a zero at z=-1 One easy way to see this is that any sequence representing half the sampling frequency, of which $(-1)^n$ is a good example, when weighted by filter taps of this symmetry, will always have a +1 on one tap and a -1 on the symmetric tap: in pairs, and collectively summing to zero output of the FIR summation.

Accordingly, we see a failure when trying to form high-pass filters in Fig. 45b and Fig. 45c In these direct designs, the specifications are asking for a response value of 1 at frequency 0.5, while the resulting filter must have a response of 0 at frequency 0.5. In addition to this local error around 0.5, relatively more error overall is clearly evident. We are therefore strongly warned against trying even length even symmetry for high-pass filters. And, as is also clear, the all-pass, Fig. 45d, does not work, defeating the subtraction method, Fig. 45e, which is nonetheless, equivalent to the direct design attempt of Fig. 45b.

Perhaps surprisingly now, we find that the modulation method seems to work - the magnitude response of Fig. 45a being flipped in Fig. 45f. To understand this, we must note that the impulse response in Fig. 45f now had odd rather than even symmetry, with the center of symmetry being 12.5. This adds a phase of $\pi/2$ to the phase response, which may or may not be of any consequence depending on the application.

Following up on the observation of Fig. 45f, we suppose that we might be able to get the subtraction method to work for cases of even length if we were to make adjustments to the phase response. Clearly for example, we get a perfectly flat all-pass for length 26 if we were to set the delay for 13 rather than for 12.5. In this case, the impulse response would look like Fig. 44d, except the tap that is non-zero (equal to 1 in fact) would be at 13 rather than at 12. At the same time, if we adjust the phase of the length 26 low-pass for an additional half unit of delay, we will find the subtraction method to work.

5a-3 Modulation Methods and Bandpass

Above we saw that a special case of modulation, multiplication by the sequence $(-1)^n$, shifted a frequency response by a frequency equal to half the sampling rate, which was the same as a flip of the z-plane. In a more general case, we could consider multiplying the impulse response of a filter, $h_1(n)$ by $\cos(n\omega_0)$ to give $h_2(n)$:

$$h_2(n) = \cos(n\omega_0)h_1(n) \qquad (83)$$

The DTFT of $h_2(n)$ is thus:

$$H_2(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \cos(n\omega_0)h_1(n)e^{-jn\omega}$$

$$= (1/2) \sum_{n=-\infty}^{\infty} (e^{jn\omega_0} + e^{-jn\omega_0})h_1(n)e^{-jn\omega}$$

$$= (1/2)\, H_1(e^{j(\omega-\omega_0)}) + (1/2)\, H_1(e^{j(\omega+\omega_0)}) \qquad (84)$$

This indicates that the multiplication (modulation) results in two copies of the low-pass, one shifted up, and the other shifted down. This is a method we can easily use to produce a band-pass response, for example.

Fig. 46 shows an example where a low-pass of length 29 with cutoff 0.1 is modulated to a center frequency of 0.3. Here we need to keep in mind that the low-pass response is centered about zero frequency, so the plot only shows half of it, which is one and a half ripples. Shifting this response so that it is centered about 0.3 now shows a full three ripples. We note here, as in the case of Fig. 45f, the symmetry of the impulse response is upset by the modulation. In some special cases, a simpler result is found. Fig. 47 shows the case similar to Fig. 46, except here the response is shifted to 0.25 Here the symmetry is maintained, and note also that half the taps are set to zero. This is the consequence of the modulating signal being $\cos(n\pi/2)$.

5a-4 Flipping the Response of an IIR Filter

Above we saw that it is possible to flip a frequency response by looking at $H(-z)$, which we associated with a multiplication of the impulse response by $(-1)^n$, equation (82). In the case where $H(z)$ is in fact an FIR filter, the filter coefficients are the same as the impulse response $h(n)$ themselves. In the case of an IIR filter, the filter coefficients of course determine the impulse response, but they are not the same thing. Typically we would write an IIR filter's transfer function as:

$$H(z) = (b_0 + b_1z^{-1} + b_2z^{-2} + b_3z^{-3} + \ldots) / (a_0 + a_1z^{-1} + a_2z^{-2} + a_3z^{-3} + \ldots) \qquad (85a)$$

$$= N(z) / D(z) \qquad (85b)$$

From this we observe two things. First, if we are looking for $H(-z)$, we find this by changing the signs of every other term in the numerator, and in the denominator The second thing is that we could, if we wished, flip just $N(z)$ or $D(z)$ individually, just to see what happens, if nothing else. In the case of the FIR, we thought of a flipping of the z-plane about the imaginary axis In the case of flipping $N(z)$, only the zeros are flipped. In the case of flipping $D(z)$, only the poles are flipped.

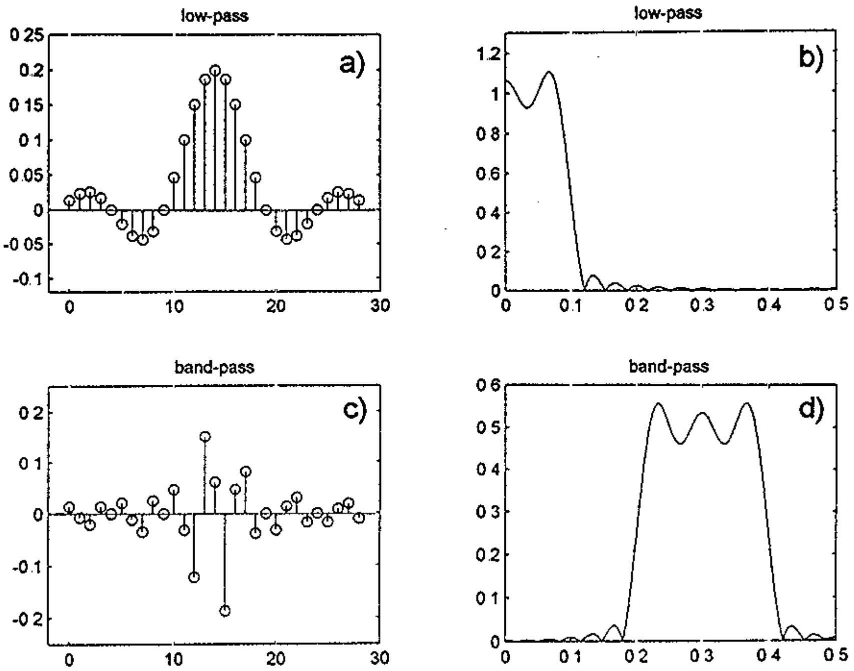


Fig. 46 Low-pass (a, b) modulated by cos(2π• 0.3•n) results in band-pass (c,d) centered about 0.3. This is a typical general case where the symmetry of the impulse response is upset. It may be a fine filter in many applications however.

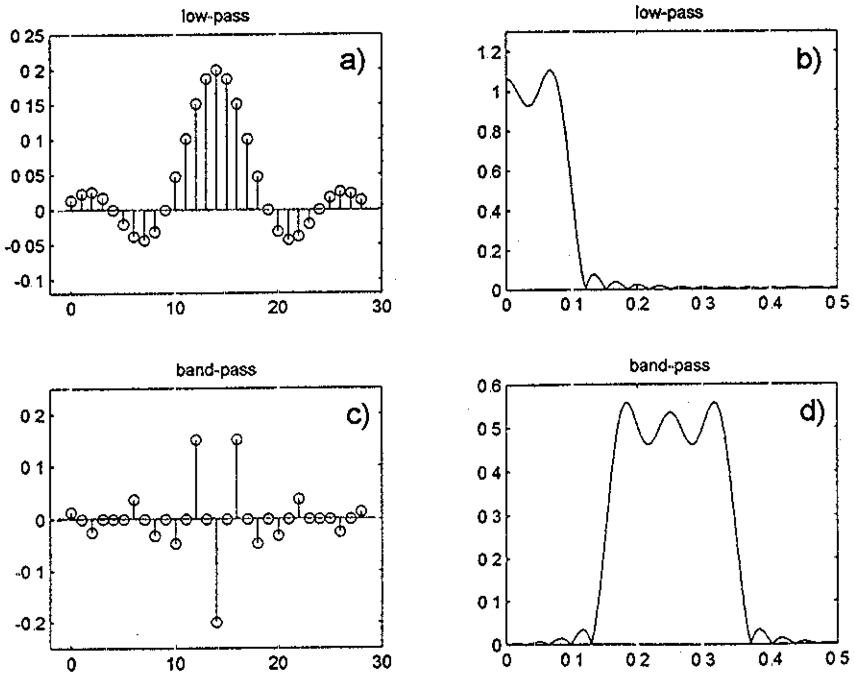


Fig. 47 Low-pass (a, b) modulated by cos(2π• 0.25•n) results in band-pass (c,d) centered about 0.25. This case preserves the linear-phase symmetry, and note that every other tap becomes zero.

Fig. 48a shows an interesting example of a 10th-order Butterworth low-pass designed by Bilinear z-transform method. In Fig. 48b, we have flipped the numerator N(z), making it N(-z). In this case, the 10 zeros which are at z=-1 in the low-pass filter move to z=+1, while the poles are not moved. Here we get a high-pass with the cutoff the same as the low-pass (the complimentary high-pass). We can think of the cutoff remaining in the same range as being a consequence of the poles not moving, and the poles being the major influence on the cutoff.

Fig. 48c shows the case of retaining N(z) while replacing D(z) with D(-z), flipping the poles while leaving the zeros alone. Now the poles move to the negative side of the unit circle, and the cutoff moves with them, while the zeros remain at z=-1, so the response remains low-pass. Finally, in Fig. 48d we have H(-z)=N(-z)/D(-z), so both the poles and zeros flip, and both the cutoff frequency and the type change. Not all other examples will be as clean as this Butterworth, Bilinear z-Transform example, but the general ideas will still be seen.

This example illustrates an advantage of implementation that comes from the type-transformation point of view. Suppose we want to separate a signal into high- and low-frequency bands using Fig. 48a and Fig. 48b. We would only need to implement the low-pass, and obtaining the high-pass would then only be a matter of forming a separate feed-forward sum - the feedback (poles) would be common to both. This simple soultion is obvious once we have studied the type-transformation idea.

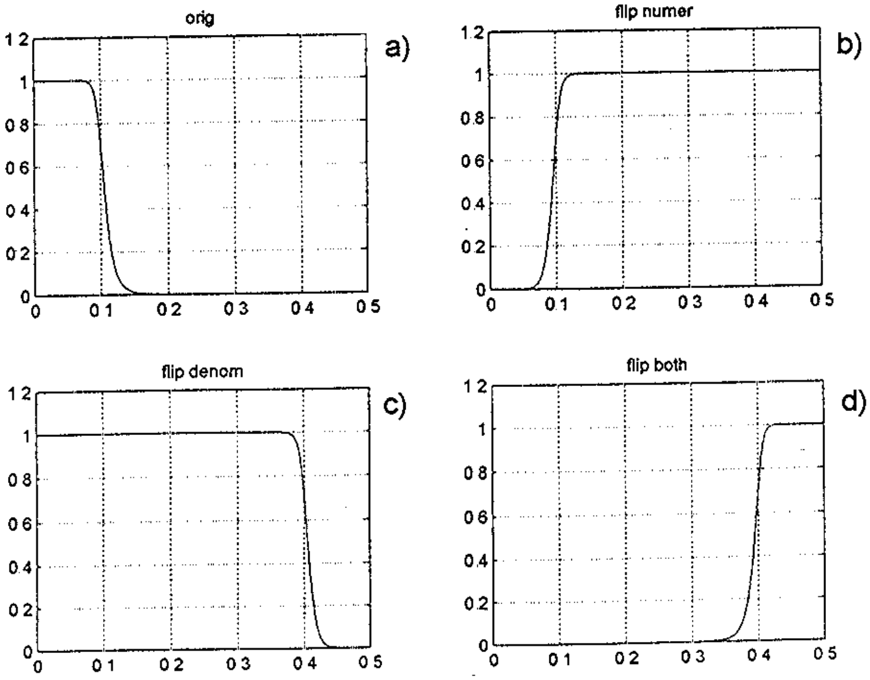


Fig. 48 Additional useful IIR filters can be obtaining by flipping every other sign in the numerator, denominator, or both.

5b. HILBERT TRANSFORMS

5b-1 Introduction

In our filter designs, we have paid primary attention to magnitude response, and when we have worried about phase, we have been primarily concerned with linear phase (although often, it was handed to us automatically in our design procedures). In addition, while we certainly have interest in such filters as high-pass, band-pass, and notch filters (in addition to the popular low-pass types), we often have to make up excuses for looking at all-pass filters. Clearly, since all-pass filters have a flat magnitude response, it is only their phase properties that are of interest. Typically, we are often concerned with adjusting or "equalizing" an existing phase response, without effecting the magnitude response, by cascading a compensating all-pass.

In one special case, we have a particular phase response in mind. This is the so-called "Hilbert Transformer" which offers a 90 degree phase shift (phase difference, actually) over a wide bandwidth. In essence, we would like to input a signal that contains a number of frequencies inside an allowed bandwidth, and have at the output, these same frequencies, at their original amplitudes, but with 90 degrees of phase added to each one. In simplest terms, we would input a sine and get out a cosine. Our purpose is to employ both the sine and the cosine (so-called "quadrature" components) to do something useful. For example, such networks are useful for shifting frequencies, and are extensively used in communications applications.

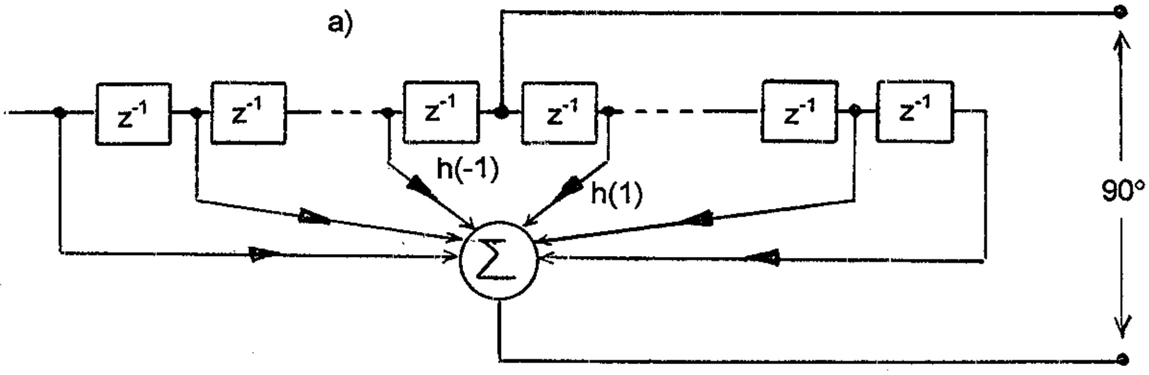
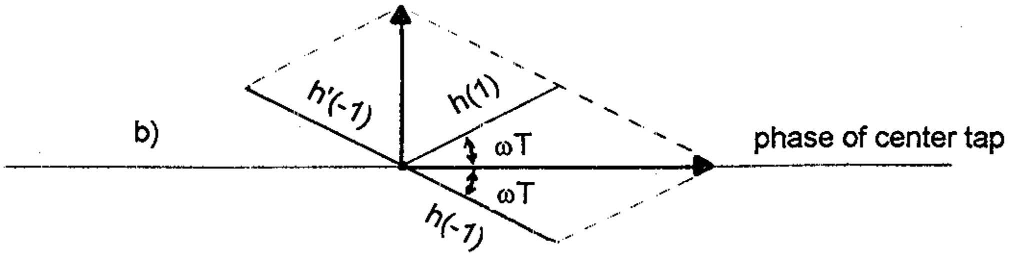


Fig. 49 Hilbert transformer structure (a) has odd symmetric taps resulting in a 90° phase difference (b).



5b-2 Standard Design of Hilbert Transformers

In practice, realistically, we look for a Hilbert transformer to take a single input, and to produce two outputs which differ in phase by 90 degrees (Fig. 49a). Typically one of the outputs is the center tap of an FIR delay line. In consequence, it is a linear phase all-pass with phase (N-1)ω/2. The other output is a more usual sum of many taps, similar to other FIR filters except for the odd symmetry of the taps (Fig. 50a). Note that the center tap must be zero for this sum, due to the odd symmetry.

Fig. 49b shows the general effect of even or odd symmetry on phase response. Suppose that the phase at the center tap is along the axis shown. Now consider the tap h(1) to be one tap to the right of center, which has an amplitude shown, and a phase delay relative to the center of ωT, where T is the delay between taps. With even symmetry, the tap h(-1) will have the same amplitude but an opposite phase (the "delay" is -T), relative to h(1). Completing the parallelogram of the vector (phasor) sum we see that the resulting phase is the same as the phase at the center. Now we suppose that the symmetry is odd such that h'(-1) is -h(-1), and here the vector sum is at 90 degrees relative to the center tap phase. It is easy to see that the phase results for either the even symmetry (linear phase) or odd symmetry (Hilbert transformer) are general for any set of symmetric taps, and to sums of symmetric taps. Accordingly, the 90 degree phase shift is perfect and automatic. Accordingly the interesting part of the design is not in the phase shift, but in obtaining a flat amplitude response. The output of the center tap is of course perfectly flat, but the sum is not

There are many ways of designing Hilbert transformers. (In fact, we can easily design for phase differences other than 90 degrees.) Instead of taking a more general view, we will restrict our example to the use of inverse DTFT, equation (4), with windowing. In specifying $H(e^{j\omega})$ we will set the magnitude to 1, so we have only to enter the phase. This we do by specifying the 90 degree phase as $e^{j\pi/2}$, which is just j. On the frequency interval -π to +π, we integrate with $H(e^{j\omega})=j$ from -π to 0 and $H(e^{j\omega})=-j$ from 0 to +π. So using the inverse DTFT we get:

$$h(n) = (1/2\pi)\int_{-\pi}^{0} j\, e^{jn\omega}\, d\omega \;-\; (1/2\pi)\int_{0}^{\pi} j\, e^{jn\omega}\, d\omega \;=\; (1/n\pi)[1 - \cos(n\pi)] \qquad (86)$$

These taps turn out to be (2/nπ) for odd values of n, zero for even n. (These are the same as the values of Fourier series coefficients for a square wave - the mathematics is the same.) Note that this is a non-causal filter, and that in most realizations, there will be some overall delay (Fig. 49a). But relative to the center tap, the result is what we want.

Fig 50 shows an example using the result of using equation (86). Fig. 50a shows the impulse response of a length 35 Hilbert transformer, while Fig. 50b shows the magnitude response. We note that this magnitude response approximates a value of 1, showing the usual Gibbs phenomenon. In most applications, the success of a system employing a Hilbert transformer depends on having both the phase and the amplitude within certain error bounds. Here the phase is perfect, but we probably have to do better with keeping the in-band amplitude flatter. [For example, in a frequency shifter (single sideband modulator), either amplitude or phase errors can result in unsatisfactory suppression of the unwanted sideband.]

Fig. 50c shows the Hilbert transformer taps of Fig. 50a windowed with a Hamming window. Very much like our other applications of windowing, we see from the magnitude response of Fig. 50d that the in-band response (say from about 0.05 to 0.45) is much flatter. At the same time, the bandwidth is probably slightly reduced relative to Fig. 50b, but the net improvement is extremely useful. In cases where there is still insufficient bandwidth, longer filters and/or more sophisticated design methods are available.

It is worth emphasizing again that it is the phase difference of 90 degrees between the two outputs that is exploited. We are not generally concerned with phase relative to the input. And, the phase relative to the input is a function of frequency. Thus if the phase of a particular component of frequency ω at the input is $\phi_0(\omega)$, the phase at the center tap will be $\phi_0(\omega)+(N-1)\omega/2$, and at the Hilbert transformer output, $\phi_0(\omega)+(N-1)\omega/2+\pi/2$. Only the phase difference $\pi/2$ is independent of ω.

Finally, our Hilbert transformers are used for broad-band phase differencing. As noted, Fig. 50d would be useful for frequencies from about 0.05 to about 0.45. That is, we expect frequency to vary but to be limited to some range. If on the other hand it were the case that we only needed quadrature components at a single known frequency, we would not need to go to the trouble of constructing a Hilbert transformer. Instead, if we have the original signal, and any other phase of the same signal (meaning in simplest terms, perhaps just a single delay of this signal), we can form an appropriate linear combination of the two available phases to form a 90 degree version

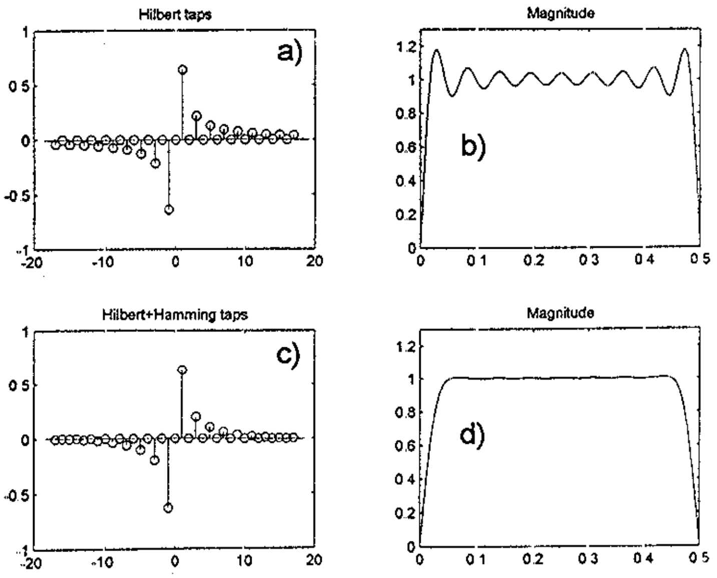


Fig. 50 Using inverse DTFT followed by Hamming windowing, a reasonably flat magnitude response results (d)

5c. MINIMUM PHASE

5c-1 Introduction

Linear phase filters are extremely popular and useful, offering a constant time delay at all frequencies and thereby minimizing phase distortion. Also, many (perhaps most) FIR filter design procedures automatically provide linear phase (or zero phase - simply converted to linear phase by using a delay) The delay of a linear phase filter corresponds to approximately half its length. In some sense, this is a fairly large delay, but any delay, within limits, is often irrelevant. For example, a delay of tens to hundreds of milliseconds in a CD player are of no importance, considering that the actual music was recorded in the past, and as long as music comes out within a few seconds of pushing a button. Delays may be important in other systems however, such as in control loops where delays may slow damping (potential instability).

5c-2 Properties of Reciprocal Zeros

Linear phase filters have zeros that come in reciprocals, and at the same time, complex conjugate pairs By this we mean that if there is a zero at angle θ and radius r, there must be three other zeros that have also been determined The four zeros are at angles ±θ for the angle and at r and 1/r for the radii. This is the most general case - zeros coming in quads. There are additional cases: the zeros may be on the unit circle (pairs - conjugates that are "self reciprocal"), may be real at r and 1/r (pairs that are reciprocal, with no conjugation), or we may have one or more zeros at z=-1 and at z=+1

5c-3 Effect on Magnitude Response

We are interested in the effects of reciprocal zeros, on the magnitude response, and on the phase response. Fig. 51 shows a typical reciprocal pair. We know that an FIR transfer function can be represented (factored) in terms of its zeros (z_1, z_2, z_3, ...) as:

$$H(z) = (z-z_1)(z-z_2)(z-z_3)... \qquad (87a)$$

If we seek the magnitude of the transfer function (the magnitude of the frequency response), we can place magnitude bars on each factor:

$$|H(z)| = |(z-z_1)||(z-z_2)||(z-z_3)|.... \qquad (87b)$$

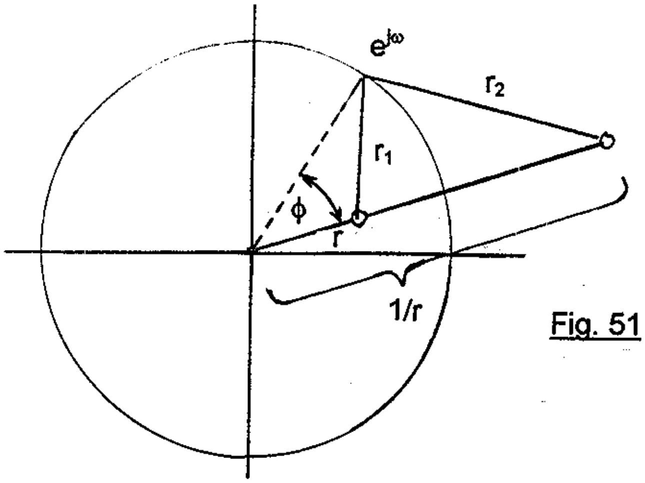


Fig. 51   A general case of reciprocal zeros

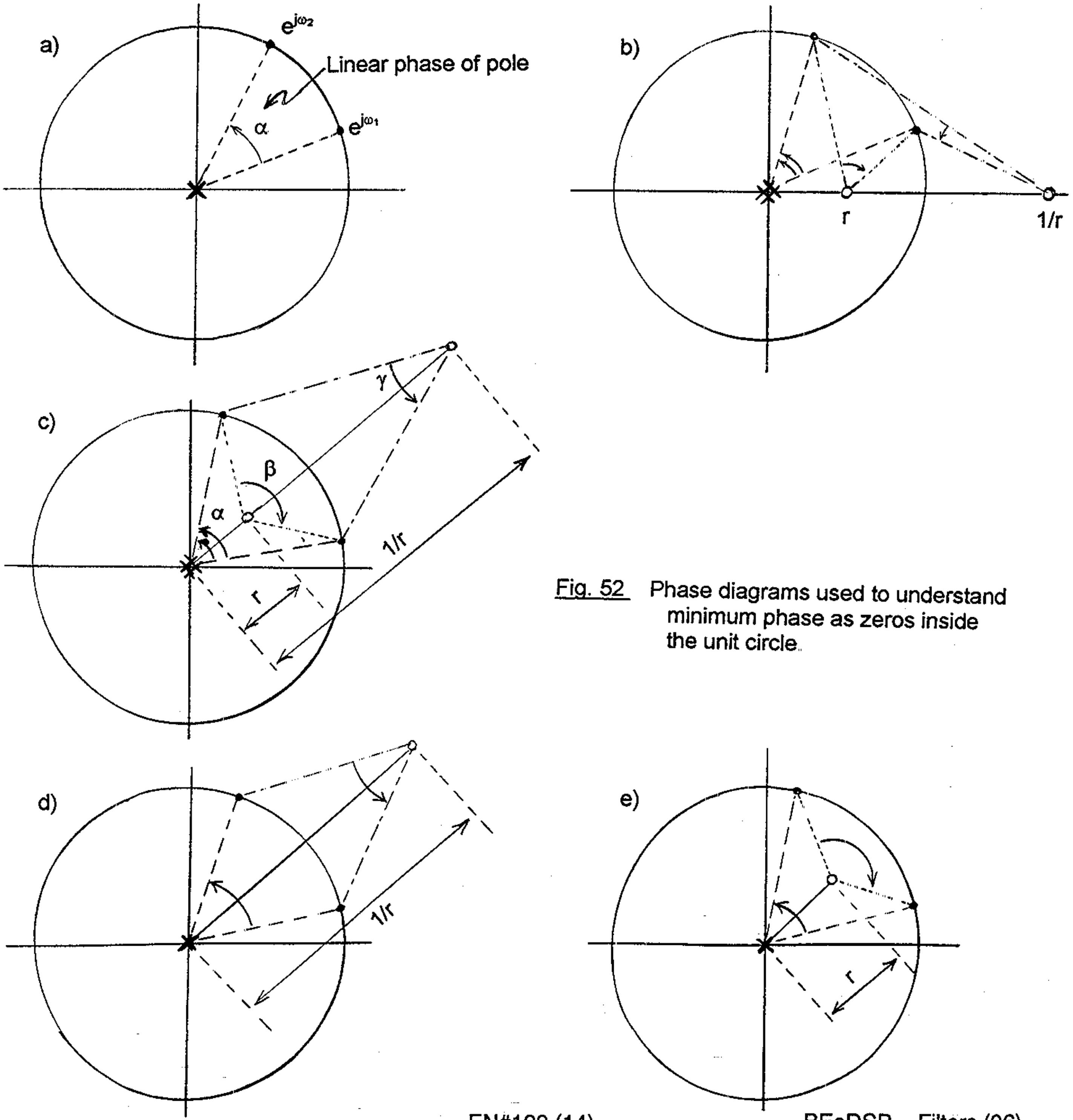


Fig. 52   Phase diagrams used to understand minimum phase as zeros inside the unit circle.

Accordingly, we recognize that the magnitude of the frequency response is proportional to the product of the distances from any point z of interest (in this case, a point on the unit circle corresponding to a frequency ω) to the zero. That is, an actual distance in the complex plane. If we seek to determine (or compare) the effect of any one zero, we need to see how this distance varies with frequency.

In Fig. 51 we show the pair of reciprocal zeros and a typical frequency point $e^{j\omega}$. The distances from this typical point to the zeros are r_1 and r_2 as shown. These distanced r_1 and r_2 are easily calculated in terms of the angle ϕ using the "Law of Cosines."

$$r_1^2 = 1^2 + r^2 - 2 \cdot 1 \cdot r \cdot \cos(\phi) \qquad (88a)$$

$$r_2^2 = 1^2 + (1/r)^2 - 2 \cdot 1 \cdot (1/r) \cdot \cos(\phi) \qquad (88b)$$

From these, it is clear that $r_1 = r_2 r$. What this means is that the effects of a reciprocal zero outside the unit circle, as compared to one inside the unit circle, differ only by a constant (r) as far as the magnitude response is concerned. Put another way, if we take a particular FIR filter and flip one or more zeros to reciprocal positions, we expect to change the overall gain of the filter, but not the shape of the magnitude in any way. Most simply, we anticipate a choice of reciprocal positions if we can make use of it.

It is well to mention here that we are accustomed to linear phase filters, so when we think of flipping a particular zero, we probably expect to find the new position already occupied. In such a case, we would end up with a second-order zero. There is nothing wrong with this, but likely we could better use the zero by putting it somewhere else. As we shall see, the practical design of minimum phase filters involve a "prototype" linear phase filter. Below we will also need to look carefully at linear phase in order to see how minimum phase might work. So while we are intimately involved with linear phase even here, we must keep in mind our minimum phase goal.

5c-4 Effect on Phase Response

In order to see the effect of the choice of reciprocal positions on phase response, we need to recognize that FIR filters, as we generally consider them, have poles at the origin of the unit circle. These we tend to ignore because they have no effect on the magnitude response, being at unit distance for all frequencies. Perhaps we can see this best by considering the simplest possible linear phase filter:

$$H(z) = z^{-1} = (1/z) \qquad (89)$$

This is a filter with unity gain and a linear phase equal to a unit delay. It's just a delay! Writing it as (1/z) makes the point that in the z-plane (not the z^{-1} plane) we have a pole at z=0. This zero is responsible for the linear phase, as is seen in Fig. 52a. Here we will find it sufficient to examine phase as the angle spanned by a region of frequency as seen from the singularity. The phase of a pole is the negative of that of a zero. Equivalently we can span angles for poles from low to high frequency, and for zeros from high to low frequencies.

Fig. 52b expands the view to a length three linear phase filter. This filter has two poles at z=0 and zeros at r and 1/r. The phase angles associated with the singularities are as shown by the arrows. The phase due to the two poles totals twice that of Fig. 52a. The phase of the zero inside the unit circle (clockwise) opposes that of the poles (counter-clockwise), while the phase due to the zero outside the unit circle (counter-clockwise) adds to that of the poles. Note that Fig. 52a is a special case of Fig. 52b where r=0 so that one zero cancels one pole, leaving one net pole (Fig. 52a) with a zero at infinity.

Fig. 52c considers the case of a reciprocal pair of zeros which we can consider a component of a larger linear phase filter. We can think of this as Fig. 52b rotated, and the basic geometry of the angles is the same. Since we know this is a linear phase filter, it must be true that the phase of the inside zero, minus the phase of the outside zero, cancels the phase of one of the poles, leaving the linear phase term of the single pole. While this can be shown mathematically, it is also convincing to draw a variety of cases (arbitrary angle, r, and circle arc length) with compass and ruler, and to measure the angles with a protractor.

5c-5 Choice for Minimum Phase

While our study of reciprocal placement of zeros, and the poles at z=0 reveals much about linear phase, here we are concerning ourselves with minimum phase. So when we make a choice about placing a zero at r or at 1/r for a particular case, we expect not to effect the shape of the frequency response (only the gain) but to have a major effect on the phase response. Thus we compare the choice of a zero outside (Fig. 52d) and a zero inside (Fig. 52e). It is thus clear that when the zero is outside, the phases add (maximal phase) while when the zero is inside, the phases oppose. Another simple way to understand this is that when the zero is inside, it is closer to the pole, and the pole and zero have a tendency to "sheild" each other. From Fig 52e we see that there is a net phase in favor of the zero.

5c-6 A Strategy for Minimum Phase

We might well thus suppose that a viable strategy for obtaining a minimum phase filter would be to first design a FIR filter that has the magnitude property we want Then we would find the zeros, and any zeros that are outside the unit circle will be replaced with zeros inside. Zeros on the unit circle will be kept. If we do this, we in fact find an impulse response that is weighted toward the input side, consistent with the notion of less delay.

If we actually do the suggested experiment, we will likely need to get an initial test filter by placing zeros in random (conjugate) pairs. As mentioned, if we instead start with standard FIR design programs, we likely get linear phase filters. To make these into minimum phase filters, we merely flip outside zeros to reciprocal zeros inside, placing them on top of existing reciprocal zeros. This is correct - just not that useful. Another way to look at this undesirable restriction is that in giving up linear phase (asking for minimum phase) we release the symmetry constraints on the impulse response, giving us potentially about twice the degrees of freedom, which then become available to help tailor the magnitude response to a better shape. To get minimum phase for this improved shape, we only have the lesser constraint of keeping all zeros inside or on the unit circle.

5c-7 From Equiripple to Minimum Phase

Given that we should be able to get about twice the degrees of freedom through the use of minimum phase, we might consider an approach where we start with an oversized linear phase design with the expectation of cutting it down. One popular scheme is based on an equiripple linear-phase filter

The first step is to start with an odd length linear phase filter of about twice the final size desired (Fig. 53). One then adds the stopband error (ripple amplitude) to the response (Fig. 54a). This is just a matter of adding the error to the center tap of the impulse response. The symmetry about the center remains, so we still have linear phase and reciprocal zeros. At the same time, the stopband zeros have been drawn together in pairs, becoming second-order (Fig. 54b). The entire response is now positive (or zero) and the zeros of the response are accordingly rounded as seen in Fig. 54.

We then reflect all the zeros that are outside the unit circle inside. This makes all these second-order as well. The magnitude response is the same as it was when the error was added (except for an overall multiplier). Since all zeros are second order, this response is the squared magnitude of same other filter of about half the length. This other filter is clearly equiripple. It is found by making all zeros first order.

In practical terms, once the error is added to the center of the impulse response, it is necessary to recompute the zeros from the modified h(n), all of which move at least a little. When these are found, it should be the case that some will be double zeros on the unit circle, while others will be in reciprocal pairs. It is then usually possible to just keep the zeros with magnitude less than one. This is first because this is what we want for the reciprocal pairs. Curiously, this likely even works for the unit circle pairs. Because of small errors and the second-order nature of these zeros, with the usual root finders, one of each pair seems to end up ever so slightly outside the unit circle while the other ends up ever so slightly inside, so a sorting with reference to one works.

Fig. 55 and Fig. 56 show the final minimum phase filter. We note from Fig. 55b that the response is equiripple. (Because this filter is only half the length of the original equiripple (Fig. 53b) its performance is far less impressive. The example here was chosen to most clearly illustrate the ideas involved, not for the final filter achieved.) From the impulse response (Fig. 55a) we see that the larger tap weights are skewed toward the input (low index) end, which helps us to understand how it is that the overall delay of the minimum phase filter is less. We are of course interested in looking at the phase of the minumum phase filter, and this we see in Fig. 56. For comparison, the phase response of a linear-phase filter of the same length is plotted. Our main region of interest is the passband, from 0 to about 0.15. In this region, we see that the minimum phase filter has reached about $\pi/2$ total phase. Note that this phase response is not a straight line, but wiggles slightly. In contrast, the linear phase filter has cycled through π, and continued to almost 3π total phase.

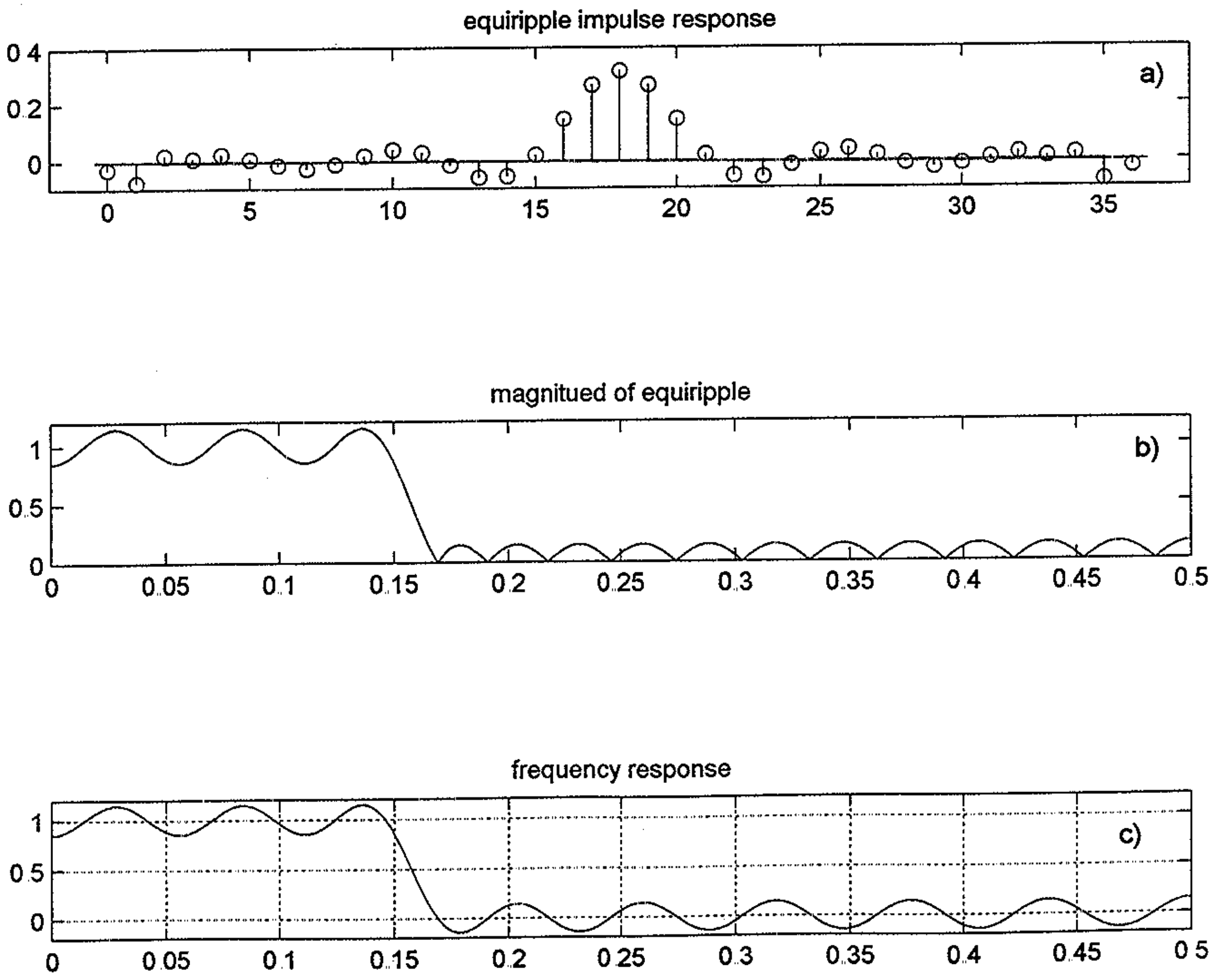


Fig. 53 Minimum phase design begins with an equiripple filter of about twice the final length desired. Note that (c) shows the actual frequency response (non-causal).

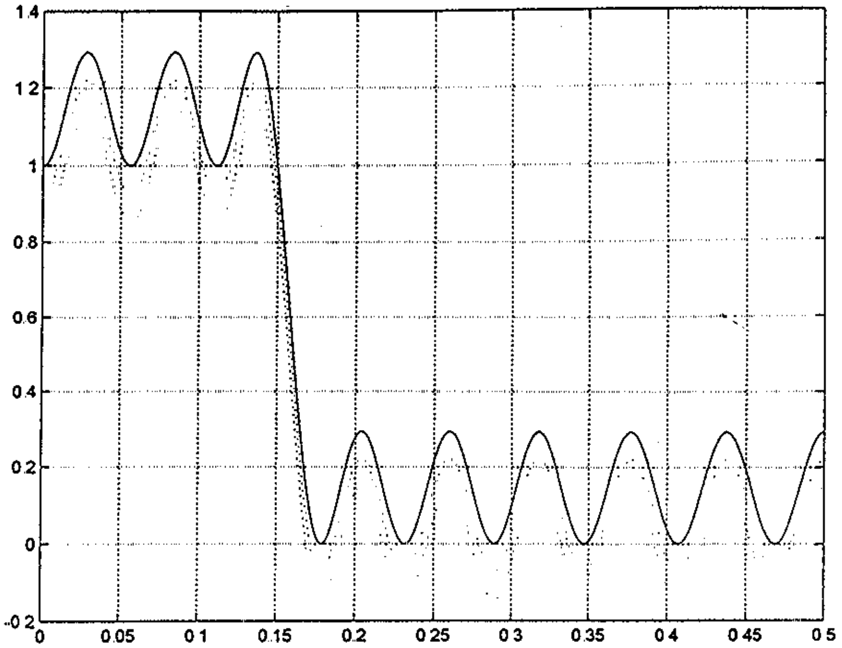


Fig. 54a By adding the magnitude of the stopband ripple to the center tap of the equiripple prototype, a filter with purely non-negative frequency response results. Here we show the original response (lower dotted line) and a case where only half the ripple magnitude has been added (middle dotted line).

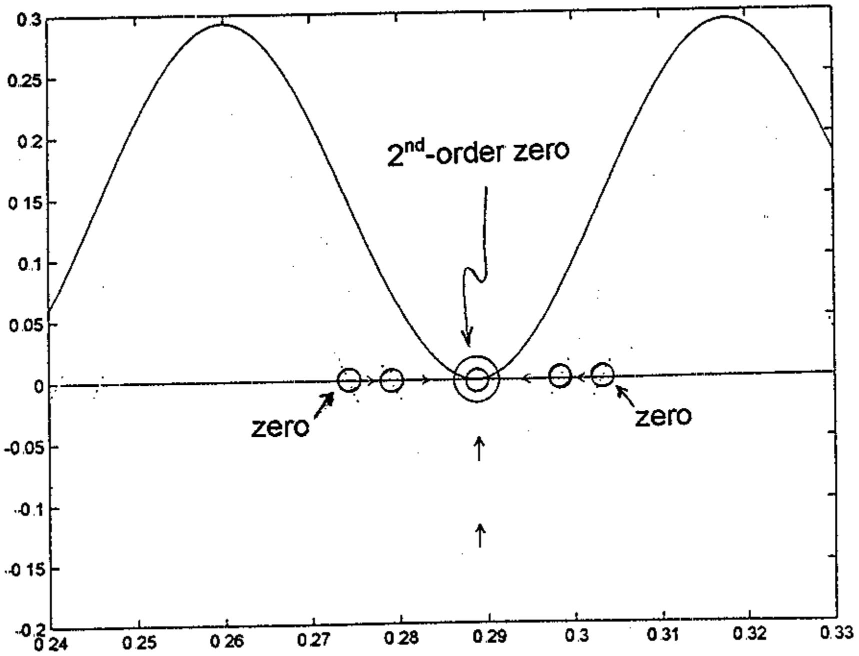


Fig. 54b Here we have a detail of Fig. 54a. As the response is raised, the zeros come together in pairs, forming a second-order zero

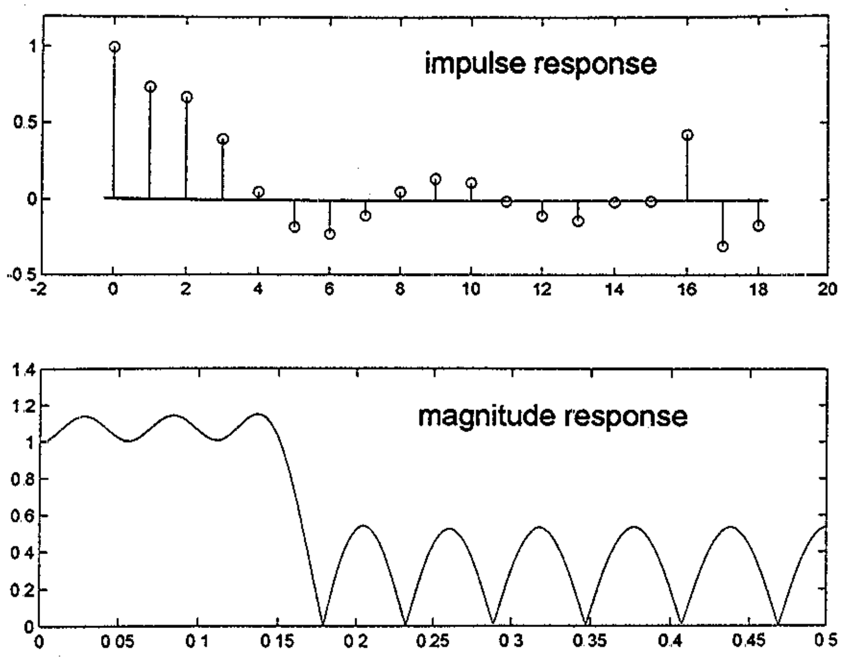


Fig. 55 In (a) we have the impulse response of the final minimum phase filter. Note that the taps are weighted toward the input end. The magnitude response (b) is equiripple.

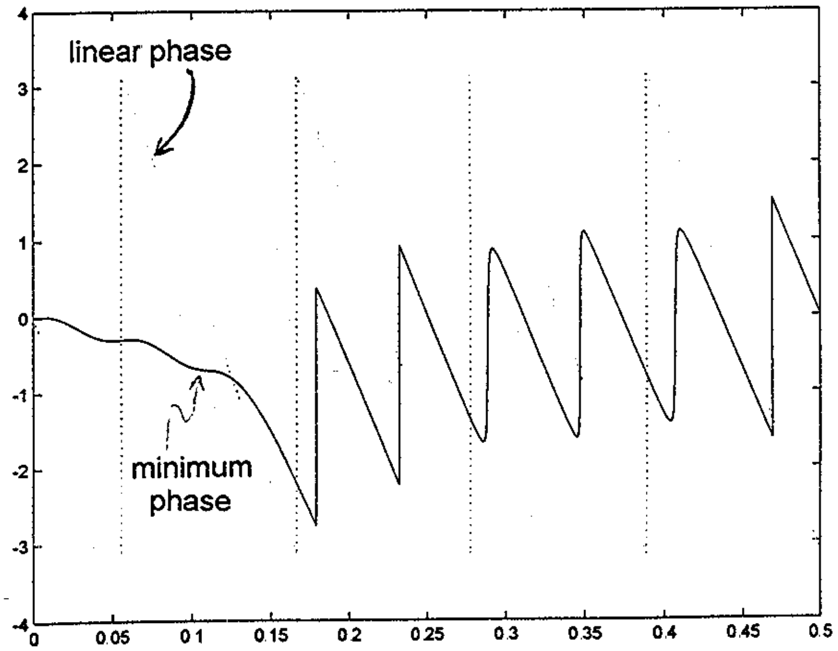


Fig. 56 The all-important phase response of the minimum phase filter (solid line) is not straight, but in the passband (0 to 0 15) it is far less than that of the comparable linear phase filter (dotted line).

5d INTEGRATOR SIMULATION

In Section 2d we looked at the Bilinear z-Transform (BZT) method of IIR filter design, which we there approached by just giving a substitution to be made. There we had a function of z to be substituted for the analog frequency variable s. Here we will recast equation (24a) as:

$$1/s \leftarrow (T/2)(z+1)/(z-1) \tag{90}$$

where, by flipping the equation over, we hope to thereby establish the point of view that we are substituting for an analog integrator. Is the right side of equation (90) in any sense a discrete version of an analog integrator? Here we will show that the Bilinear z-Transform is equivalent to a trapezoidal approximation to an analog integration. The BZT works extremely well (having several favorable properties which we had no right to expect.) To put this all in perspective, we will consider, in addition to the BZT, two forms of discrete rectangular integration, and an integrator approximated by cubics. All these are methods that are simple and fairly well understood in terms of numerical integration.

5d-1 Rectangular and Trapezoidal Integration

Fig 57a shows a sequence of samples. If we want to integrate this sequence, we might want the new value of the integration to be the old value plus the current update. (Again, these methods originate from cases where we are trying to approximate the integral of a continuous-time signal by using samples of the continuous-time signal.) For example, take y(n) to be the current value of the integration. We then want:

$$y(n+1) = y(n) + x(n)T \tag{91a}$$

Here we have what is really rectangular integration, since the update term x(n)T is really a rectangular area. If we z-Transform this equation we get:

$$Y(z)z = Y(z) + TX(z) \tag{91b}$$

or:

$$Y(z)/X(z) = T / (z - 1) = Tz^{-1} / (1 - z^{-1}) \tag{91c}$$

This is a simple integrator (Fig. 58a) or "accumulator." Clearly however (see Fig. 57b), a slightly different interpretation of a discrete-time integrator could be:

$$y(n+1) = y(n) + x(n+1)T \tag{92a}$$

Where here we again have an update term x(n+1)T. [Indeed, we really have no preference between these two possibilities, and this will eventually lead us to look at trapezoidal integration.] Continuing as above we have:

$$Y(z)z = Y(z) + TX(z)z \tag{92b}$$

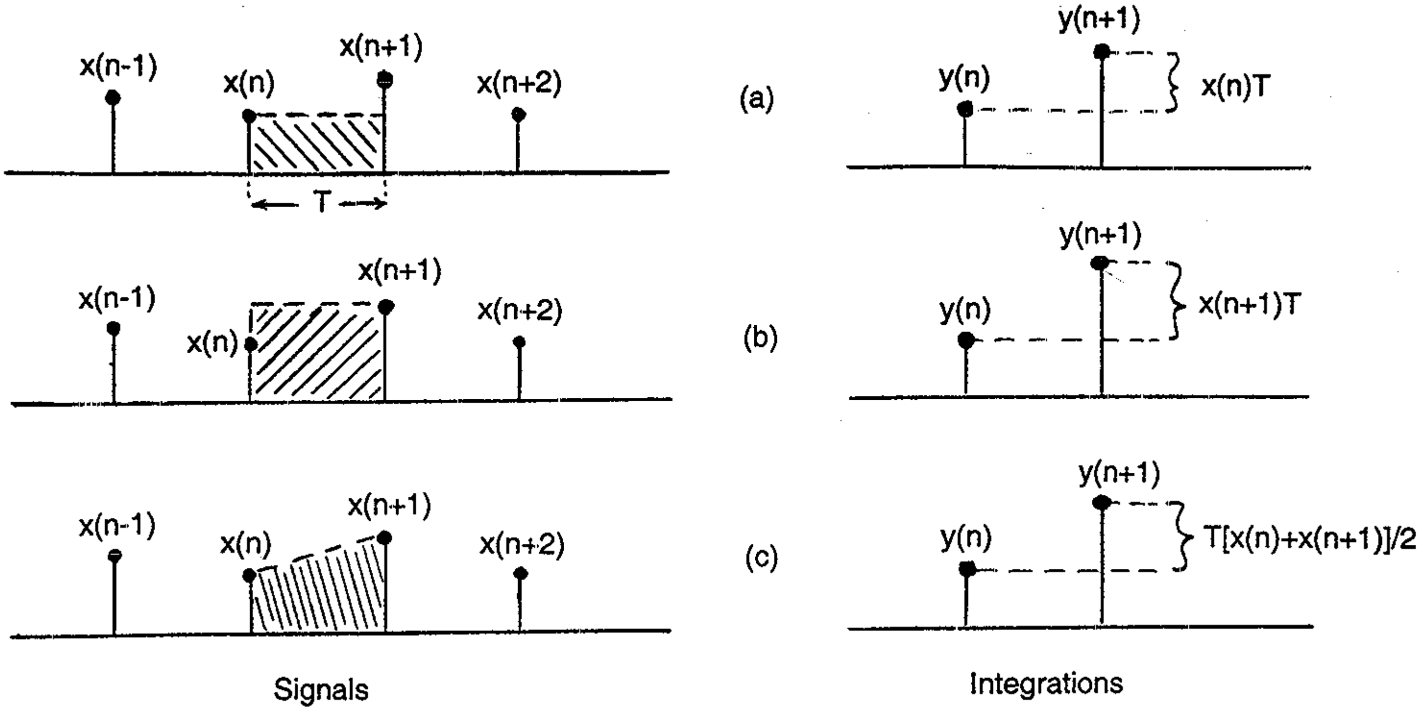


Fig. 57 Numerical integration.

or:

$$Y(z)/X(z) = Tz / (z - 1) = T / (1 - z^{-1}) \qquad (92c)$$

and we see that equation (91c) is just a delayed version of (92c), which obviously had to be the case. This leads to the integrator of Fig. 58b).

Without further delay, lets look at trapezoidal integration, which is shown in Fig. 57c. We resolve the ambiguity between choice of rectangular areas (forward or backward looking) by using the trapezoidal area instead. Thus:

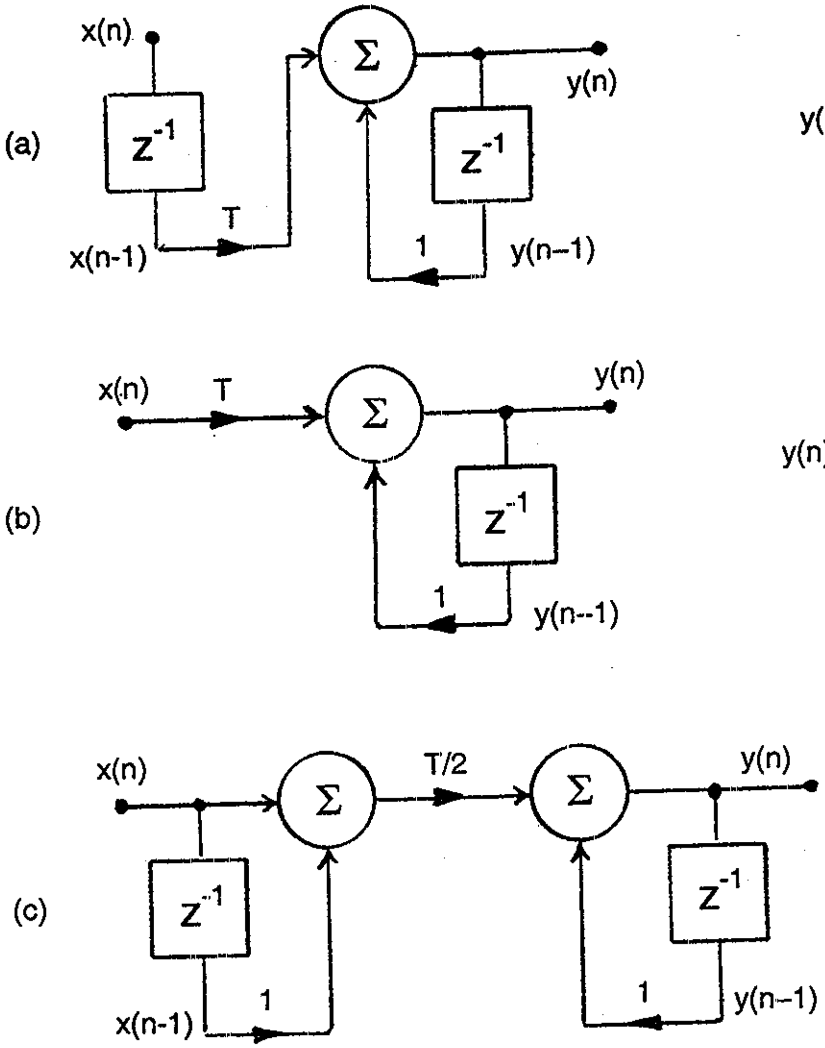
$$y(n+1) = y(n) + T[x(n) + x(n+1)]/2 \qquad (93a)$$

which has a transfer function:

$$Y(z)/X(z) = (T/2) [z+1] / [z-1] \qquad (93b)$$

Since this integration corresponds to 1/s, we can understand the BZT substitution of equation (90) [or of equation (24a)], as trapezoidal integration. The BZT integrator is shown in Fig. 58c.

Traditionally the two case of rectangular integration are not emphasized, although the corresponding cases of differentiation are: as the "forward difference" and the "backward difference." Both of these are used in control theory and (rarely) for filter design. It is natural



$$y(n) = y(n-1) + Tx(n-1)$$

$$\text{or: } y(n+1) = y(n) + Tx(n)$$

$$y(n) = y(n-1) + Tx(n)$$

$$\text{or: } y(n+1) = y(n) + Tx(n+1)$$

$$y(n) = y(n-1) + (T/2)[x(n) + x(n-1)]$$

$$\text{or: } y(n+1) = y(n) + (T/2)[x(n) + x(n+1)]$$

Fig. 58 Discrete time integrators, rectangular (a), (b); trapezoidal (c)

to approximate a derivative by a difference. [As with numerical integration, these things are much more believable if we suppose the sampling frequency is much larger than that required by the sampling theorem. That is, the discrete-time approximation is more likely valid when we have a dense set of samples.]

5d-2 Differentiating

We differentiate a function to obtain its "slope" and in the case of a discrete-time sequence, we take differences (between consecutive samples usually). If our discrete-time sequence consists of samples of a continuous-time function, the differences may approximate the derivative as:

$d_b(n) = (1/T)\ [\ x(n) - x(n-1)\ ]$ (94a)

called the "backward difference" or as:

$d_f(n) = (1/T)\ [\ x(n+1) - x(n)\ ]$ (94b)

called the "forward difference." Looking at the z-transform we have

$D_b(z) / X(z) = (1/T)\ [\ 1 - z\text{-}1\ ]$ (94c)

and

$D_f(z)/X(z) = (1/T)\ [\ z - 1\ ]$ (94d)

Since these represent derivatives, we seek to use the z-transforms of equations (94c) and (94d) as substitutions for the analog variable s. For the backward difference we would have:

$s \leftarrow (1 - z^{-1}) / T$ (95a)

for a mapping of:

$z = 1 / (1 - sT)$ (95b)

while for the forward difference we would have:

$s \leftarrow (z - 1) / T$ (95c)

for a mapping:

$z = 1 + sT$ (95d)

It is easy to show [12] that neither of these mappings is as attractive as BZT. As is obvious from equation (95d), forward difference merely scales and shifts the s-plane by 1. This is good as s=0 should be mapped into z=1, but clearly there are stable regions of the s-plane (any and all with negative real values) that are not mapped into the interior of the unit circle. Neither is the jΩ-axis mapped into the unit circle (Fig. 59a). The mapping of equation (95b) is less obvious, but the jΩ-axis is mapped inside the unit circle but not into the unit circle (Fig. 59b). Thus stable analog filters are mapped into stable digital filters, but restricting poles to the much smaller circle generally results in a restricted set of responses which are often just too "wimpy" to consider. With BZT, the jΩ-axis is mapped into the unit circle and the entire left half of the s-plane becomes the interior of the unit circle in the z-plane.

Note that the mapping of equation (95c), forward difference, is the same as that of one case of rectangular integration, equation (91c) while the mapping of equation (95a), backward difference is the other rectangular integration case, equation (92c). BZT is the average of (91c) and (92c).

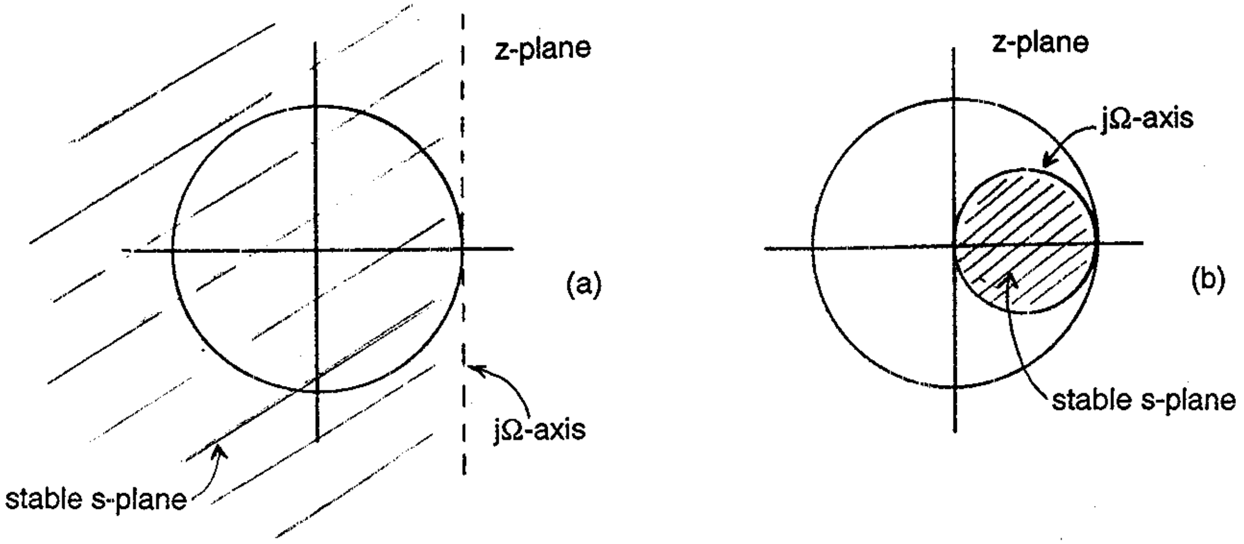


Fig. 59 Forward difference mapping (a) and backward difference mapping (b). Neither is as good as the BZT mapping (Fig. 15)

5d-3 Cubic Integration

As mentioned, BZT corresponds to trapezoidal integration, only a first-order formulation, and yet its mapping properties are near ideal (there is the frequency warping). What happens if we try a higher-order integrator approximation. While second-order would seem to be a logical next step, here we are going to try third-order (cubic). We do this because we expect that skipping to third-order might better indicate any improvements, and as we shall see, a cubic fit has a less ambiguous region of relevance relative to parabolic.

Here is what we will do: First we will choose four points of the sequence. For convenience, we will choose these points at times -1, 0, 1, and 2. The choice is arbitrary, and this one simplifies our math. We will then fit a (continuous time) cubic polynomial to these four points (Fig. 60a). How well do we expect this to work? Well, exactly as was the case with cubic interpolation (Section 4b-3), we feel pretty confident with the approximation near the middle (between 0 and 1), less so on the sides (-1 to 0, and 1 to 2), and don't trust the polynomial at all outside these regions. [Indeed, a polynomial runs off to infinity outside the points specified (it is a "vertical feature") while we think of signals as staying in range horizontally. See Fig. 60b.] Accordingly, if we were interpolating extra points, we might well want to stay in the range of 0 to 1. For our integration, we want to take the new contribution to the integral to be the area under the cubic polynomial from 0 to 1, not the trapezoid used with BZT (Fig. 60c).

So it is a simple matter of assuming a polynomial:

$$x(t) = at^3 + bt^2 + ct + d \qquad (96)$$

and we want it to fit the four given points x(-1), x(0), x(1), and x(2). Thus:

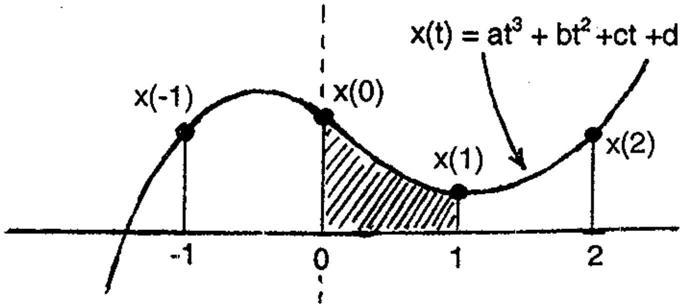


Fig. 60a Cubic polynomial fit to four points

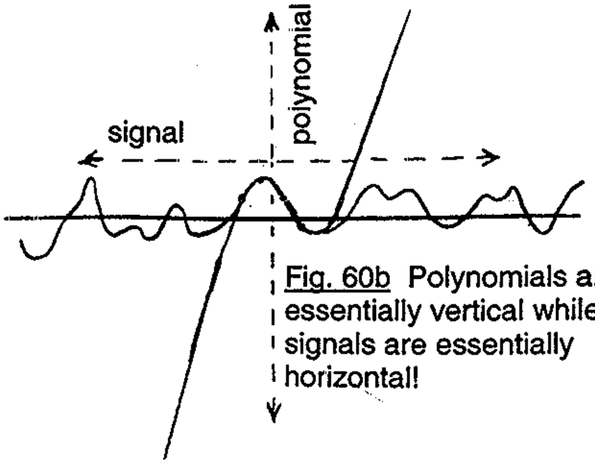


Fig. 60b Polynomials are essentially vertical while signals are essentially horizontal!

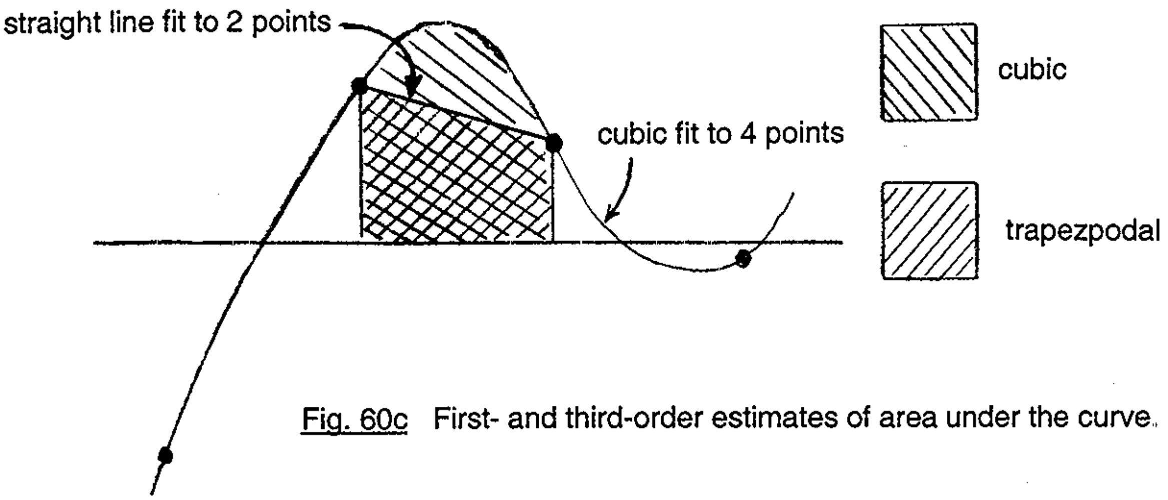


Fig. 60c First- and third-order estimates of area under the curve.

$$x(-1) = -a + b - c + d \qquad (97a)$$

$$x(0) = d \qquad (97b)$$

$$x(1) = a + b + c + d \qquad (97c)$$

$$x(2) = 8a + 4b + 2c + d \qquad (97d)$$

which are easily solved for a, b, c, and d:

$$a = (1/6)\,[\,-x(1) + 3x(0) - 3x(1) + x(2)\,] \qquad (98a)$$

$$b = (1/2)\,[\,x(-1) - 2x(0) + x(1)] \qquad (98b)$$

$$c = (1/6)\,[\,-2x(-1) - 3x(0) + 6x(1) - x(2)\,] \qquad (98c)$$

$$d = x(0) \qquad (98d)$$

We now know x(t) in terms of a, b, c, and d, which are determined by the given points x(-1), x(0), x(1) and x(2). The area under the cubic from t=0 to t=1 is a simple matter of integrating x(t) and we obtain:

$$A = a/4 + b/3 + c/2 + d \qquad (99)$$

Putting in the results of equation (98) we obtain:

$$A = -x(-1)/24 + 13x(0)/24 + 13x(1)/24 - x(2)/24 \qquad (100)$$

We now want our integration to be:

$$y(n+1) = y(n) + A(n) \qquad (101a)$$

$$= y(n) - x(n-1)/24 + 13x(n)/24 + 13x(n+1)/24 - x(n+2)/24 \qquad (101b)$$

where we have generalized equation (100) to an arbitrary n. When we z-transform equation (101b) we arrive at our integrator as:

$$Y(z) / X(z) = [- z^{-1} + 13 +13z - z^2] / 24(z-1) \qquad (102)$$

This gives us our new substitution:

$$s \leftarrow 24(z-1) / [- z^{-1} + 13 +13z - z^2] \qquad (103)$$

Notice that this result is not all that much different from the BZT. If we can justify throwing out the first and last terms of the denominator (1 is small compared to 13) and if 24/13 is a good enough approximation to 2, we get BZT from equation (103). Thus equation (103) can be thought of, as we would hope, as a refinement on the trapezoidal integration.

5d-4 Using Cubic Integration

Having arrived at the substitution of equation (103) we might suppose we should pursue a development paralleling that which followed the BZT substitution. What are our prospects in this regard? One immediate thing to note from equation (103), taken to be a mapping, is that each analog pole results in three digital poles (not just the one we get with BZT), along with three zeros. Further, it is easy to show, by example if in no other way, that stable analog poles can be mapped into unstable digital poles. In addition, because zeros at analog infinity are mapped into zeros at z=-1, we understand that there is apparently a frequency warping similar to that with the BZT [which resulted in equation (24d)]. But it is not clear if the corresponding warping equation can be developed (it is surely much harder than BZT), or if perhaps, the warping might be so severe as to rule out the cubic integration approach. The prospects do not look especially promising.

However we would be prepared to argue that because we are using a superior integration method, somehow the difficulties with the cubic integration (CI) method must be resolvable, and the resulting digital filter should be a better approximation to the analog filter. Another way to look at this is to mention that we now have three poles to manipulate to try to get a better approximation to the analog filter. (In fact, we shall see that the extra two poles provide minor adjustments contributing to a better approximation.) Accordingly we can here set aside the issue of obtaining an analytic warping equation, and instead try to determine the warping empirically after we complete the initial steps of a design. The only seemingly insurmountable problem is that of the unstable poles we are likely to encounter. This we will get around by noting that an unstable pole can always be "reflected" inside to a reciprocal position, giving the same effect on the shape of the frequency response magnitude. This important result was worked out in detail for the case of zeros in Section 5c-3. The same mathematics applies to the case of reciprocal poles. This is a convenient and often overlooked way to achieve stability.

So at this point we are prepared to just try an example. Of course, the algebraic complexity we saw with BZT [in actually plugging the function of z into a high order T(s)], will be even worse with CI. We will accordingly try to see how much we can learn from just a first-order example. Once we have what appears to be a correct filter, we will back-calculate the frequency warping from the magnitude response of the first-order example. To help show that this is valid, we will apply the same back-calculation to BZT where we know the tangential warping well.

Consider the first-order analog low-pass:

$$T(s) = 1 / (s + \Omega_c) \qquad (104)$$

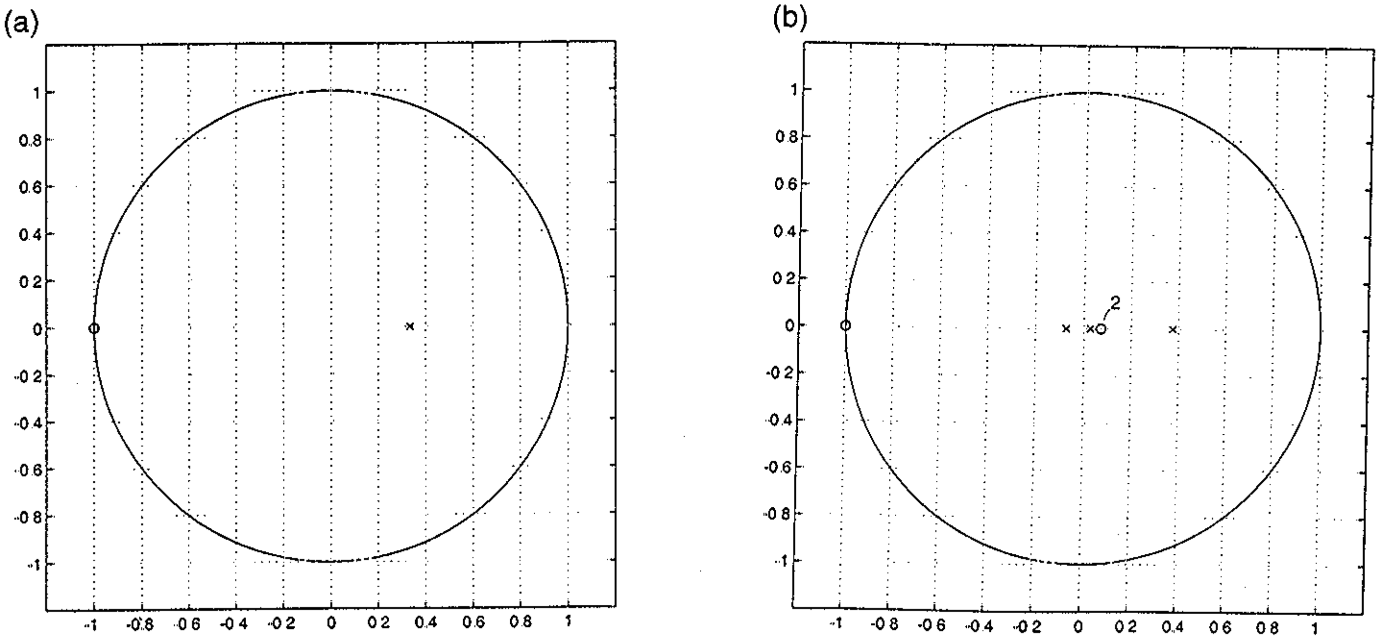


Fig. 61 BZT design (a) and CI design (b)

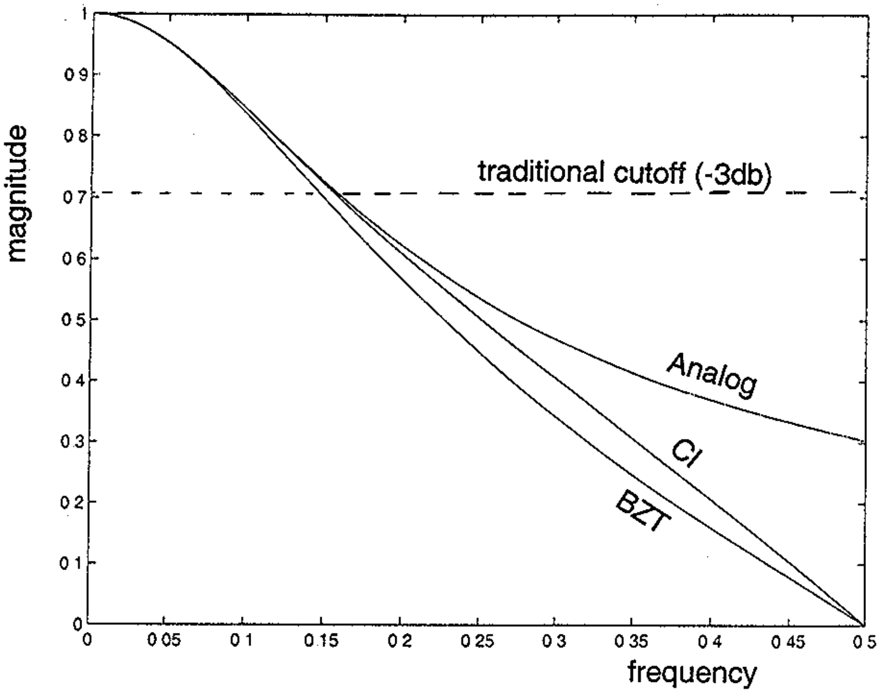


Fig. 62 Analog, BZT, and CI designs for a first-order low-pass with cutoff at 1 rad/sec.

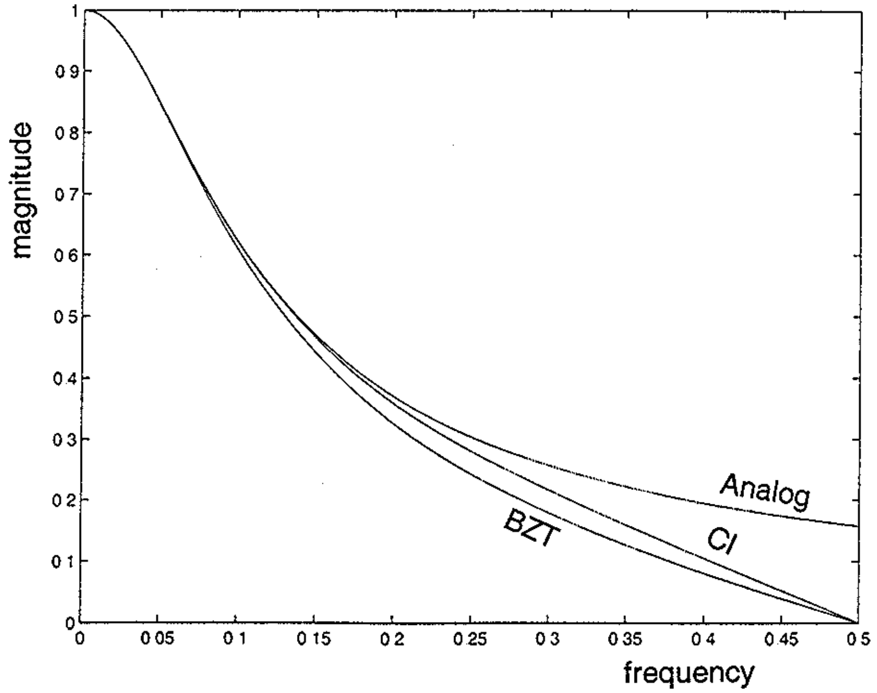


Fig. 63 Analog, BZT, and CI designs for a first-order low-pass with cutoff at 0.5 rad/sec.

where Ω_c is the 3db cutoff frequency. Making the BZT substitution of equation (21) with T=1, we obtain a first-order transfer function:

$$H_{BZT}(z) = (z+1) / [\, z(\Omega_c + 2) + (\Omega_c - 2)] \qquad (105)$$

which has a zero at z=-1 and a pole at $z = (2-\Omega_c)/(2+\Omega_c)$. For the case $\Omega_c = 1$, the pole is at z=1/3 (Fig 61a). The CI substitution of equation (103) yields a third-order transfer function:

$$H_{CI}(z) = (-z^{-1} + 13 + 13z - z^2) / [-\Omega_c z^{-1} + (13\Omega_c - 24) + (13\Omega_c + 24)z - \Omega_c z^2] \qquad (106)$$

Because this is third order, finding the poles and zeros must be done numerically. For $\Omega_c = 1$ we find real zeros at 13.9282, -1, and 0.0718 and real poles at -13.7051, 2.6778, and 0.0272. The only actual concern here are the two poles outside the unit circle. We can however reflect

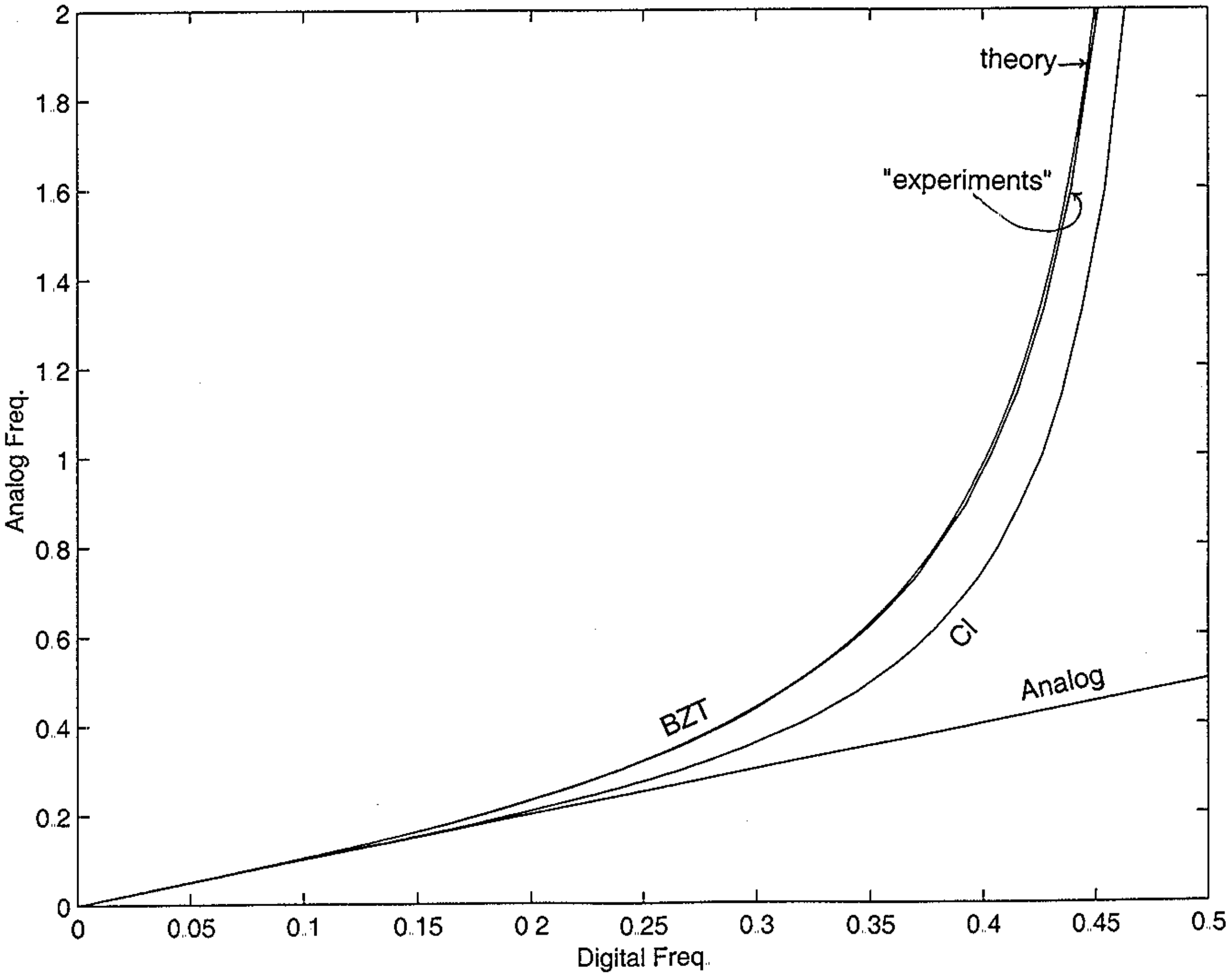


Fig. 64 Warping curves for analog (no warping), CI, and BZT (curve for BZT shows theoretical as well as "experimental" plot.

the zeros as well as the poles inside the unit circle. Once this is done, we have zeros at 0.0718, 0.0718, and -1 while the poles are at -0.0730, 0.3734, and 0.0272 (Fig. 61b). Note that two of the zeros and two of the poles are relatively weak (close to z=0). What remains is the zero at z=-1 (same as BZT) and the pole at 0.3734 (close to the pole at 1/3 from BZT).

The actual test is of course to compare the frequency responses of BZT and of CI to each other, and both to the analog response, and this is shown for Ω_c=1 in Fig. 62. (Appropriate gain normalizations have been made.) True enough, the CI result matches the analog result better; clinging to it up until a frequency of about 0.15 while BZT gives up at a frequency of about 0.08. Both BZT and CI depart and head for zero response at half the sampling frequency. [Note that the analog cutoff is $\Omega_c /2\pi = 1/2\pi = 0.159$ Hz here.] Fig. 63 shows a second example where the cutoff is lowered to Ω_c=0.5.

From the plots of Fig. 62 and Fig. 63 it is clear that a frequency warping occurs both in BZT (which we knew about) and in cubic integration (which we suspected). It can also be seen that the warping with CI is less severe, relative to BZT, in the low-frequency region, for example, around the cutoff frequency. (Both warp severely as frequencies approach 0.5). In order to get a fuller picture of the warping, we can back-calculate from the response curves. To do this, we choose a series of response values, and search the magnitude curves for these levels and then note their corresponding frequencies. For example, if we used Fig. 62 and a level of 0.4, we would have an analog frequency of about 0.365 while CI has a frequency of about 0.3 and BZT a frequency of about 0.27. A computer program can then fairly easily sweep a wide range of response levels, and plot the BZT and CI frequencies as a function of the analog frequency. [Note that we could not use levels less than 0.3 with Fig. 62 or 0.16 with Fig. 63, and would not be able to scan digital frequencies of much more than 0.3 or 0.4. Instead we must choose a much lower analog cutoff, perhaps 0.01. This gives a very narrow plot (not shown), but is perfectly valid.] Fig. 64 shows the warping curves. While we have chosen first-order and a very low-cutoff (seemingly a special case), the result is general (simple first order, monotonic, is perhaps the ideal choice). We are able to verify that the warping with CI is noticeably less than with BZT, as we supposed should be the consequence of using a better integrator approximation. Fig. 64 also shows a plot for the theoretical value [equation (24d)] for BZT, in close agreement with "experiment." The theoretical warping for CI was not established.

REFERENCES

(Sections 1-3 of Filter Element)

[1] T.W. Parks and C.S. Burrus, Digital Filter Design, pp 63-67, Wiley (1987)

[2] B. Hutchins, "Subtle Design Considerations for Hamming Windows," Electronotes Application Note No. AN-319, March 1992

[3] T.W. Parks and C.S. Burrus, Digital Filter Design, pp 69-71, Wiley (1987)

[4] B. Hutchins, "Least-Squared Error Filter Design," Electronotes Application Note No. AN-315, November 1991

[5] B. Hutchins, "Weighted, Integrated, Least-Squared Error for FIR Filter Design," Electronotes Application Note No. 332, August 1995

[6] B. Hutchins, "A General Review of Frequency Sampling Design," Electronotes Application Note No. AN-337, March 1996

[7] T.W. Parks and J.H. McClellan, "Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase," IEEE Trans. Circuit Theory, Vol. CT-19, pp 189-194, March 1971

[8] T.W. Parks and J.H. McClellan, "A Program for the Design of Linear Phase Finite Impulse Response Filters," IEEE Trans. Audio and Electroacoustics, Vol. AU-20, No. 3, pp 195-199, August 1972

[9] B. Hutchins, "Extensions of Equiripple FIR Filter Design," Electronotes Application Note No AN-311, Nov. 1990

[10] B. Hutchins, "Mathematical Considerations Relating to Equiripple Filters," Electronotes Application Notes No. AN-342, Feb. 15, 1997

[11] B. Hutchins, "A Simple Equiripple Filter Obtained by Iterating by Hand," Electronotes Application Note No. AN-343, Feb. 28, 1997

[12] L.R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing, pp 212-216, Prentice-Hall (1975)

(End of Filter Element)

****

(Thévenin - continued from page 2)

Accordingly we often desire to effectively turn down the input level, for convenience, or on a stage-by-stage basis for a cascade arrangement (out of necessity). Fig. 1b shows the idea of reducing the input level by a factor of K so as to end up with unity gain overall. Here we denote this reduced input voltage as V_T, and named the input resistor R_T, but insist that it be equal to the original R (so as not to change filter cutoff and/or characteristic). The "T" stands for Thévenin.

There are a number of cumbersome ways to achieve a reduction of input level; for example, a voltage-divider followed by a buffer (Fig. 1c). We want to just use the approach of replacing the input resistors with two resistors R* and R** (Fig. 1d). Consider the sub-network of Fig. 1d, which is

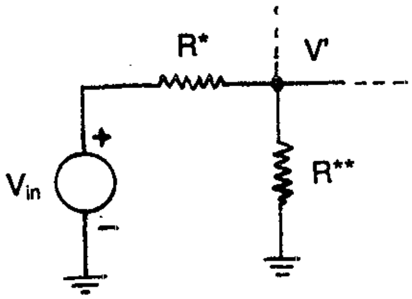


Fig. 2a Gain-reducing sub-network

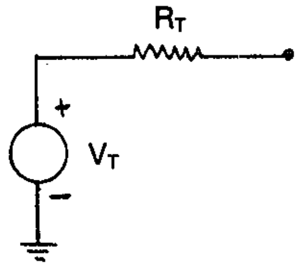


Fig. 2b Thévenin Equivalent

shown in Fig. 2a. We seek a Thévenin equivalent of this, as suggested in Fig. 2b. Thévenin's theorem tells us that we find the voltage V_T by finding the open circuit voltage of Fig. 2a. That is, find the voltage V' in Fig. 2a for the case where nothing else is connected to the junction point of R* and R**. This is just a voltage divider:

$$V_T = V_{in} R^{**} / (R^* + R^{**}) \qquad (1)$$

and the Thévenin resistance R_T is found by shorting all voltage sources and determining the resistance looking back in from the V' junction. This is just the parallel combination of R* and R**:

$$R_T = R^*R^{**}/(R^* + R^{**}) \qquad (2)$$

We want V_T to be the reduced voltage:

$$V_T = V_{in}R^{**}/(R^* + R^{**}) = V_{in}/K \qquad (3)$$

and we want to maintain the input impedance at R:

$$R_T = R^*R^{**}/(R^* + R^{**}) = R \qquad (4)$$

We easily solve equations (3) and (4):

$$R^* = KR \qquad (5)$$

$$R^{**} = KR/(K-1) \qquad (6)$$

While we have used Thévenin's theorem, all that we really needed to remember was that the voltage divider should result in 1/K and the parallel combination of R* and R** must maintain R. (This seems quite reasonable, but is not exactly obvious.) If we remain unconvinced, we can always just solve out the network of Fig. 1d directly, and calculate its design equations. We would find that the transfer function would be the same, except the gain would now be unity.

However, let's do a simpler problem. What usually bothers people is that they know it is perfectly all right to suppose that a voltage divider is "valid" as long as there is nothing adding or subtracting current from the output node. (That is, the divider is not "loaded.") True. And if there is something attached to the node (other than a high-impedance buffer), the divider is pushed or pulled about, at least to some degree. In one senses, this can be thought of as a

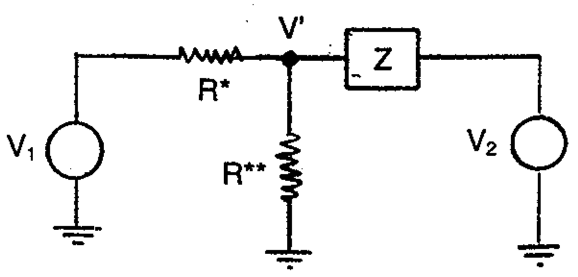


Fig. 3a Gain reduction divider with Thévenin on right representing "the rest" of the network.

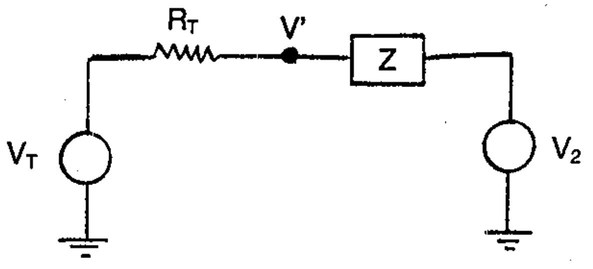


Fig. 3b Two "Fighting Thévenins"

failure to achieve an exact division. Sometimes we want this well-defined division, and we must resort to keeping the resistors in the divider very small (having a large standing current) or use a buffer. But this is not always necessary or even what we want. In many cases, the actual non-zero impedance represents a desirable flexibility at the node. We want it. That's what the non-zero Thévenin resistance takes care of. Fig. 3a shows our gain reducing resistors connected to a impedance Z and a voltage-source V_2. That is, we are assuming Thévenin's theorem is correct (at least for the right side) and representing "the rest" of the circuit by a Thévenin equivalent. Note that we are assuming the Thévenin impedance is Z, and it need not be purely resistive. This would in fact be the case if we were solving the complete low-pass filter. From Fig. 3a, we sum currents at V':

$$\frac{V_1 - V'}{R^*} = \frac{V'}{R^{**}} + \frac{V' - V_2}{Z} \qquad (7)$$

This we solve for V' as:

$$V' = [V_1R^{**}Z + V_2R^{**}R^*] / [R^{**}Z + R^*Z + R^*R^{**}] \qquad (8)$$

Is this the same result we get by assuming a Thevenin equivalent of the left side (Fig. 3b). Fig. 3b is simple to solve for V'. We could set up one current, but it is convenient to see the result as the superposition of two voltage dividers:

$$V' = (V_TZ + V_2R_T) / (R_T + Z) \qquad (9)$$

FInally substitute for V_T and R_T the expressions of equations (1) and (2). This gives us back equation (8).

This method of reducing gain is fairly general. However, keep in mind that the technique is most useful when applied in a distributed manner (correcting each stage) rather than trying to do it all in one stage. The price to be paid for this distribution effort is just one additional resistor (3 cents!) per stage.

****

Electronotes, Vol. 20, No. 199, September 2001
Published by B. Hutchins, 1016 Hanshaw Rd., Ithaca, NY 14850 (607)-257-8010