ELECTRONOTES 196

N E W S L E T T E R   O F   T H E

M U S I C A L   E N G I N E E R I N G   G R O U P

1016 Hanshaw Rd., Ithaca, NY 14850

Volume 20, No. 196 December 2000

GROUP ANNOUNCEMENTS

CONTENTS OF EN#196

This issue covers the final Chapters: 8, 9, and 10 of Analog Signal Processing. These chapters go into some of the more unusual areas of analog signal processing, voltage-controlled filters (unusual, but not to readers of this newsletter), delay line filters (related to digital filters), and analog adaptive filters.

In going over this analog material one more time after so many years, I have been surprised just how much material there was - much of it I had forgotten. Another annoying thing was how many "problems left to the reader" I had left hanging! In many cases, as previously stated, the actual problem statements were fairly obvious. Yet in some others, as I went through, I did not remember exactly what I intended to ask, and in still more, I had forgotten just how the problems worked out. On the happy side, I did however stumble across a fair number of the problems which I had actually typed up, but forgotten.

Looking at the whole thing now, I think it may prove useful if, from time to time, we work out and publish some of these problems, so we can perhaps have an "analog signal processing corner" as a repeating department. This is the current plan. Just below, in this spirit, although not in response to any particular text problem, a follow-up on the passive sensitivity discussion from Chapter 7 is offered.

TUNING EQUATIONS DERIVED FROM PASSIVE SENSITIVITIES

Passive sensitivity as discussed in Chapter 7 of Analog Signal Processing (EN#195) is important first of all in helping with our choice of one circuit configuration as compared to an alternative configuration that is nominally also capable of realizing a desired response. All other things equal (and this is itself a comparison that must be done very carefully), we choose the configuration with the lower sensitivity values. This use of passive sensitivity numbers is essentially "global" with respect to actual particular instances during production.

A second way to use passive sensitivity calculations relates to the actual "fine tuning" of individual instances (i.e., a particular circuit board off the production line). Suppose for example that we design a filter, choose a configuration, and construct 10 examples. Perhaps it is a low-pass with a nominal cutoff of 1000 Hz. Our global consideration of passive sensitivities assures us that we expect, perhaps, actual cutoff frequencies between 900 and 1100 Hz, and our tests indeed show measured cutoff frequencies well within this range. (In fact, we might well expect a balance between overvalued and undervalued components to keep us away from worst-case examples). So nothing is unexpected or wrong.

(continues on pg. 37)

CHAPTER 8

VOLTAGE-CONTROLLED FILTERS

8-1 THE NEED FOR VOLTAGE-CONTROLLED FILTERS:

The filters that we have discussed so far, and with which we are likely otherwise already familiar, are fixed filters, and we have in mind that they would be constructed by taking the correct parts off the shelf and soldering together the correct circuit. Such filters find wide application in cases where the specification parameters of the job the filter is to do are well established and expected to remain constant.

If we think a bit, it is clear that we can make variable filters rather than fixed filters by making some of the filter's components variable. In particular, we wish to make the filter's time constants vary to change the characteristic frequencies of the filter. To do this, we need to have variable resistors, variable capacitors, or both. If we look at the range of component values we usually encounter, it is clear that we can easily find suitable variable resistors, but that suitable variable capacitors are unlikely. (Consider that a standard radio-tuning variable capacitor is already fairly large, and is only in the hundreds of picofarad range.) We can easily find potentiometers in the range of a few ohms up to several megohms.

This opens up the idea that we can have variable or tunable filters. However if we restrict ourselves to potentiometers as the variable resistors, we are talking about manual tuning. Typically, a potentiometer is what we find used as a radio volume control, and it has a knob to be turned up or down. In this view, tunability is a step in the right direction, but there are still many things that we can't do.

For example, we might have a tenth-order filter needing a ten-way ganged potentiometer to tune all ten of its frequency controlling resistors. We might need quite an enclosure to house this control, and it would probably be somewhat difficult to turn the control shaft with the usual type knob. Secondly, manual tuning implies that we have to have someone close by to turn the knob, so we could not expect to tune the filter remotely (on a space satellite for example). In addition, computers can conveniently handle numbers and even voltages, but not turn knobs. Still additional problems with manual tuning come up when we need to adjust a filter faster, or with more precision, than can be done manually.

All of these additional jobs can be done with voltage-control or other electronic tuning. We can control almost unlimited controlling elements in parallel, control remotely, control by computer, control with rapidly changing voltage waveforms, and set parameters to an accuracy far greater than is possible with a manual knob. Thus we see a step up from simple tunability as being voltage-tunability or voltage-control as it has come to be called.

Most of the quality work in Voltage-Controlled Filters (VCF's) has come through the efforts of engineers working on the design of electronic music synthesizers. In their desire to achieve a dynamically changing spectrum, on a time scale compared with the shortest of individual notes expected, VCF's became one of the few possible approaches. It was found that excellent VCF's could be obtained if one went about it in the right way. This "right way" seems to have been to first consider the control elements and structures that were possible, and then see what sort of filter could be made tunable with these controls. In general, simply taking your favorite active filter and trying to put in voltage-variable resistors is not a productive approach.

What was found was that one particular control element, the so-called transconductance multiplier was a key control element, and that this was most useful in certain key **structures**. Fortunately, one of these structures was the first-order low-pass section, and another was the state-variable filter.

8.2 THE MULTIPLIER AS A FILTER-CONTROL ELEMENT:

A few devices (such as some configurations of FET's, and photoresistors) are useful as variable resistors, but only over a fairly limited range (no more than 10:1). A device that is more generally useful is some form of electronic analog multiplier. To see why this is, consider that a multiplier takes two input voltages (V_{in} and V_c) and produces an output voltage V_m as:

$$V_m = KV_{in}V_c \qquad (8\text{-}1)$$

where K is some constant of the multiplier. Normally we expect that if a voltage V_{in} is properly applied across a resistor R that a current V_{in}/R flows through it. If instead the voltage V_{in} is multiplied according to equation (8-1), and V_m is applied to R, then current $KV_{in}V_c/R$ flows, which is equivalent to applying the voltage V_{in} to a different resistor of value R/KV_c. This is the general idea - that by scaling a voltage, we can make things look as though a resistor has a different value. This is not always the way things turn out in practice, however.

Fig. 8-1a Successful VC Integrator

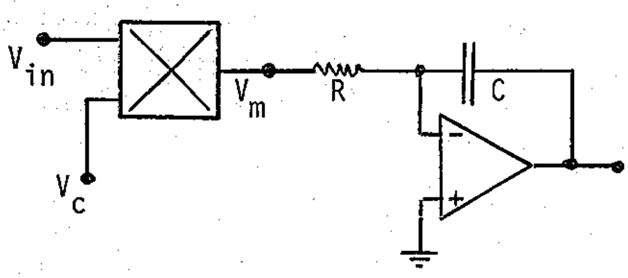
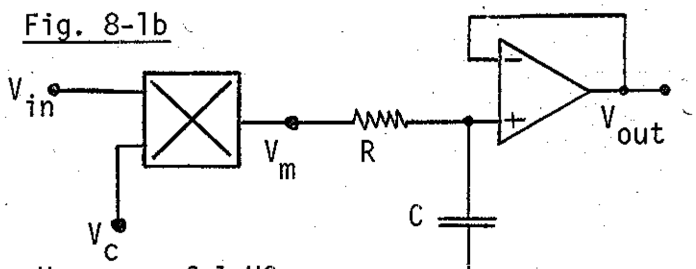


Fig. 8-1b



Unsuccessful VC First-Order Low-Pass

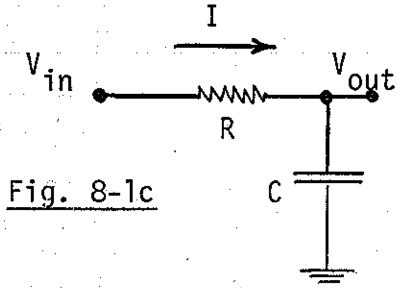


Fig. 8-1c

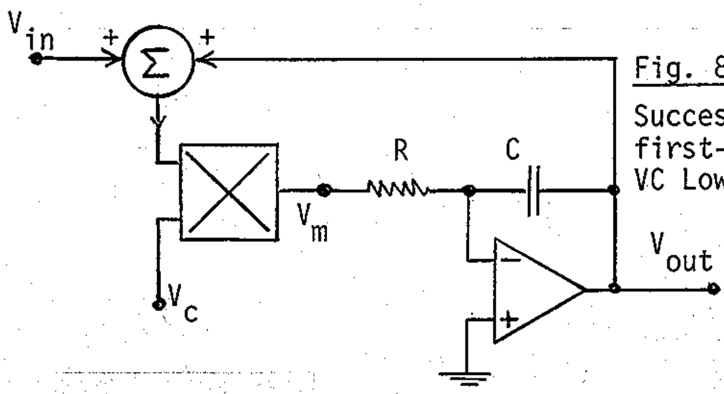


Fig. 8-1d Successful first-order VC Low-Pass

Fig. 8-1 shows four cases that serve as examples. Fig. 8-1a is a voltage-controlled integrator that does work. Fig. 8-1b and Fig. 8-1d represent an unsuccessful attempt, and then a successful attempt, respectively, at achieving a first-order voltage-controlled low-pass filter. Fig. 8-1a works by virtue of the fact that the lower end of the resistor R is at ground potential (a constant), so the only voltage that determines the current through R is V_{in} (for any fixed value of V_c). Thus a changing voltage looks the same as a changing resistance, considering the current that flows into the capacitor.

It is important to understand why the first-order low-pass of Fig. 8-1b does not work, since it at first sight seems so similar to Fig. 8-1a. In fact, the multiplier of Fig. 8-1b only changes the gain factor of the filter from 1 to KV_c, but does not change the time constant. This is obvious from just looking at the network as being composed of two parts: the multiplier to the left, and the very familiar first-order (fixed) low-pass to the right. In order to understand what is wrong, and how to fix it, consider that in the fixed first-order low-pass (Fig. 8-1c) the current through the resistor is not just a function of the input voltage V_{in}, but also of the output voltage V_{out} as $I = (V_{in}-V_{out})/R$. That is, the output gets to "fight back", and this is what is missing from Fig. 8-1b. Fig. 8-1d adds the missing feature. Here we first take the difference $(V_{in}-V_{out})$, and it is this voltage that is scaled by the multiplier before it is applied to the resistor. In addition, the first-order low-pass is now made an integrator so that the lower end of the resistor R is always grounded.

In fact, it is best to analyze Fig. 8-1d directly. The output of the multiplier is:

$$V_m = KV_c(V_{in}+V_{out}) \qquad (8-2)$$

and we have studied the inverting integrator and know it to give:

$$V_{out} = -V_m/sCR \qquad (8-3)$$

which are solved for the transfer function T(s) as:

$$T(s) = V_{out}/V_{in} = -1/(1 + scR/KV_c) \qquad (8-4)$$

From this, we have a resistor that is effectively scaled by $1/KV_c$, which is the same as scaling the cutoff frequency by KV_c.

Accordingly we now know how to approach voltage-controlled filters with two powerful weapons in our arsenal - the voltage-controlled integrator (making possible the state-variable approach), and the voltage-controlled first-order low-pass section. We will continue by looking at some possibilities for practical multipliers.

8-3 THE TRANSCONDUCTANCE MULTIPLIER:

We have seen above that the key element in our VCF approach will be the analog multiplier. Most all analog multipliers of practical interest will be based on some transconductance principle. For practical multiplier integrated circuits, transconductance multipliers are available in several forms. These we will separate into the full four-quadrant multipliers, and the two-quadrant multipliers or OTA`s (for Operational Transconductance Amplifiers).

To understand the difference in these devices, we can start with the idea that a multiplier should perform the operation:

$$Z = XY \tag{8-5}$$

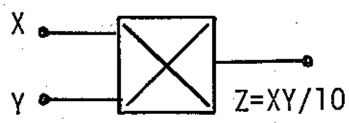
where X and Y are the inputs and Z is the output. In the case of a four-quadrant multiplier, both X and Y may take on negative and positive values. Also, Z is usually also a voltage in this case, and since we work with voltages in a convenient range such as ±10, the multipliers usually have a scale factor to bring the product into a usual range. For example, Z = XY/10 is common (Fig. 8-2a). True four-quadrant multipliers are also characterized by a cost in the range of $10 to $50 and/or a need for considerable individual trimming. In general, they are components that many engineers will avoid whenever possible.

As unpopular as the four-quadrant multiplier is, a two-quadrant multiplier in the form of an OTA has found wide application. This chip is available for about $1 or less, and is fairly easy to use. Being a two-quadrant multiplier, only one of the inputs can take on positive and negative polarity, while the other must be unipolar. The OTA chip happens to have a bipolar voltage input, while the second or unipolar input is a current rather than a voltage. In addition, unlike the usual four-quadrant multiplier, the output is a current. It turns out that all of these are things that the designer can exploit. The most popular and well known OTA for many years has been the RCA type CA3080, after which the OTA may be called a "3080" as often as it is an "OTA". The 3080 has been second-sourced by National as the LM3080, and new generations of OTA's have appeared, including dual versions such as the CA3280 and the LM13600. Other improvements have involved the use of a linearizing input stage (Gilbert input), to help with a linearity problem that will be described below.

Fig. 8-2b shows the conventional symbol for the OTA, which superficially resembles the op-amp. The most notable difference is the additional pin, the control pin for the control current I_c. This pin is pin 5 on the CA3080, and it is common for designers to refer to the control pin as "pin 5", even on OTA's with a different pin-out arrangement. The next difference to observe is that the output is a current source, and not a voltage as it is for op-amps. Finally, although it is not indicated in the diagrams, the input differential voltage should be limited to something like ±10 millivolts for linearity.

The point should be emphasized that the OTA is quite a different device from the op-amp. However, it is no more difficult to learn to use - the "rules" are just different. In order to understand better how the OTA functions, the structure should be understood (see Fig. 8-2c). The OTA consists of four "current mirrors" and a standard two-transistor differential inputs stage as shown. The current mirrors are configurations of transistors arranged so that when a certain

Fig. 8-2a Four-Quadrant Multiplier



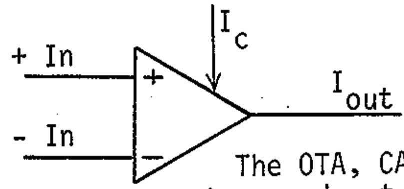


Fig. 8-2b

The OTA, CA3080, or two-quadrant multiplier

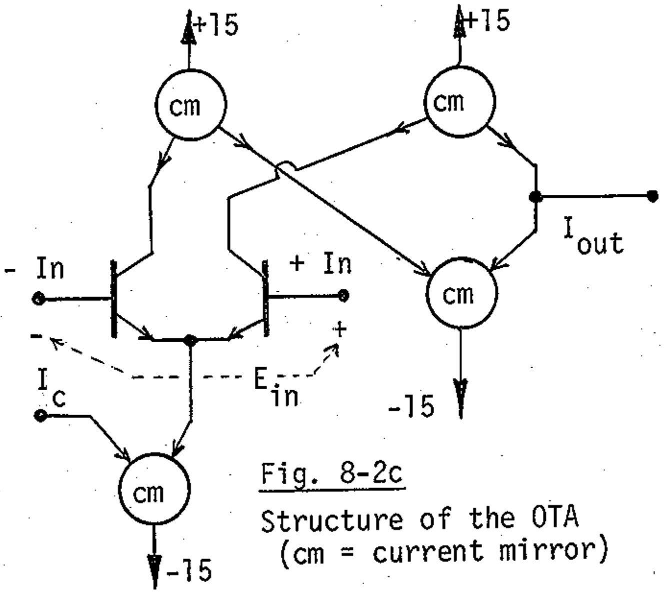


Fig. 8-2c

Structure of the OTA (cm = current mirror)

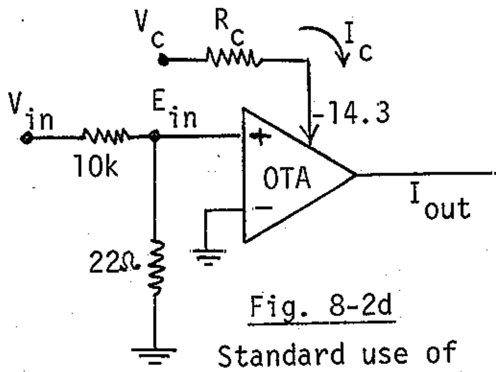


Fig. 8-2d

Standard use of attenuator with the OTA reduces V_{in} to E_{in}.

current is pulled from one branch, an identical current is sourced by the other branch; both currents being sourced or sunk by a connection to a power supply rail. One of these current mirrors receives the control current I_c and mirrors it, drawing this same current I_c from the tied emitters of the differential pair. Study of the remaining three current mirrors shows that they just "decouple" the two collector currents of the differential pair, so that the output is a current source that represents the difference between the two collector currents. When the two inputs are at the same potential, their collector currents are equal (to $I_c/2$, in fact), and the output current is zero. When there is a non-zero differential input voltage, the currents are out of balance, and the difference becomes the output current.

We do not want to go deeply into transistor theory, but the two-transistor differential input stage is well understood, and it is known that the difference between the collector currents is a hyperbolic tangent function, which can be considered approximately linear around zero for differential input voltages of no more than ±10 millivolts or so. In such a case, the output current is:

$$I_{out} = 19.2\ I_c\ E_{in} \qquad (8\text{-}6)$$

which we will consider the fundamental equation of the OTA. It is well to keep in mind that if we go all the way back to this equation for a start, we are unlikely to go wrong. Note that E_{in} is the actual voltage between the + and the - inputs. Clearly equation (8-6) implies a multiplication relationship between two electrical parameters, I_c and

E_{in}. From our study of the structure (Fig. 8-2c), it is clear that E_{in} can be bipolar, while I_c must be only positive (into the lower mirror). Thus the OTA is basically a two-quadrant multiplier here.

Since the input stage must be limited to about ±10 millivolts for linearity, and since we still want significantly larger voltages in the rest of our circuits, it is common to find an attenuator stage on OTA inputs, as seen in Fig. 8-2d. (Incidentally, the attenuation to this low level does imply problems with signal-to-noise ratio. One help is the "prewarping" or "Gilbert" input stage found in newer OTA's which permits input voltages of several volts.) With the addition of this attenuator, equation (8-6) becomes:

$$I_{out} = 19.2\ I_c\ (22/10022)V_{in} \tag{8-7}$$

or we can write an "equivalent resistance" as:

$$R_{eq} = V_{in}/I_{out} = 23.7/I_c \tag{8-8}$$

which amounts to 23.7 kΩ when I_c = 1 milliamp, and so on.

The concept of equivalent resistance and equation (8-8) should be used with considerable caution. As we have cautioned above, it is often better to go all the way back to equation (8-6). The problem comes up in assuming that R_{eq} implies that the OTA looks just like a "real" resistor. As we saw from our discussion of Fig. 8-1, we must be careful to look at voltage on both sides of resistors. Accordingly, it is best to think of R_{eq} as a notational convenience, and not as suggesting that the OTA can be treated as a resistor R_{eq} in all cases.

We have noted that the output of the OTA is a current rather than a voltage. In some OTA applications, we will drop this current through a resistor and then use an op-amp follower to buffer this voltage drop and to serve as a voltage source. In many VCF applications however, it is both possible and advantageous to just use the current directly.

Before going on to some actual circuits, it will be useful to discuss how the control current I_c is obtained. Note that the control pin is an input to a current mirror, and accordingly, its voltage will always be about one diode drop above the negative supply (about -14.3 with a -15 volt supply). I_c can be supplied with a number of current source arrangements. One simple way is to connect the control pin to a voltage source more positive than -14.3, through a suitable current limiting resistor. In such a case, for a control voltage V_c (see Fig. 8-2d), the control current is:

$$I_c = (V_c + 14.3)/R_c \tag{8-9}$$

As a practical matter, I_c should be limited to no more than 2 milliamps with 1 milliamp being a good design maximum. This means that on a standard ±15 volt supply, where the control voltage V_c might range up to +15, that if R_c is 30k or so, we are absolutely safe on control current. Note however that the control pin is a very sensitive part of the chip. If this pin is shorted to anything, there is a danger that the chip may blow. Shorting this pin to ground, or to the output of an op-amp, can blow the chip.

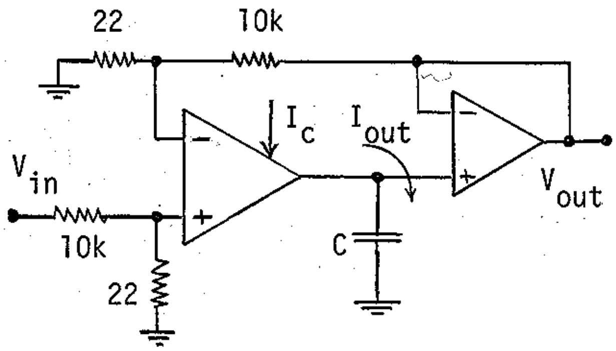


Fig. 8-3a VC Low-Pass

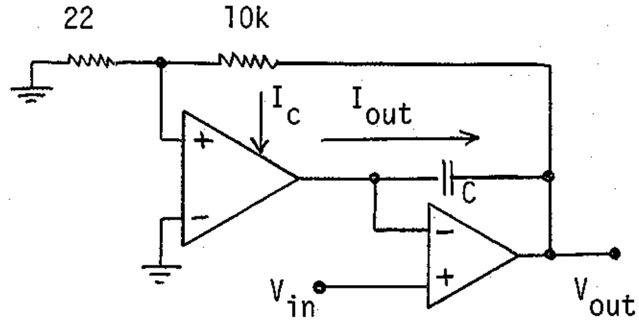


Fig. 8-3b VC High-Pass

8-4 FIRST-ORDER VOLTAGE-CONTROLLED FILTERS:

There are a large number of ways of achieving useful first-order voltage-controlled structures, which can then be combined into higher-order structures if desired. All of these involve some form of feedback from the output to the OTA input, for the same basic reason as was required in the development of Fig. 8-1d. Fig. 8-3a shows a form that is first-order low-pass. The analysis begins with equation (8-6) as:

$$I_{out} = 19.2\ I_c\ E_{in} = 19.2\ I_c(22/10022)(V_{in}-V_{out}) \qquad (8\text{-}10)$$

It is also clear that the output current flows through the capacitor C to ground, generating the voltage V_{out}, or:

$$V_{out} = I_{out}/sC \qquad (8\text{-}11)$$

These two equations can be solved for the transfer function as:

$$T(s) = V_{out}/V_{in} = 1/(1 + sCR_{eq}) \qquad (8\text{-}12)$$

where we have also used equation (8-7) for R_{eq}. Note that R_{eq} falls exactly into a position in the transfer function where we recognize its effect on the cutoff frequency. In fact, since the cutoff is at:

$$f_{3db} = 1/2\pi R_{eq}C = I_c/(2\pi\ 23.7\ C) \qquad (8\text{-}13)$$

we have a cutoff frequency that is proportional to the control current. Note that we did not try to begin with the idea of R_{eq} for the OTA and then treat the OTA as a normal type resistor. Instead we started with the basic equation for the OTA, equation (8-6), and put in the notation R_{eq} when it appeared naturally. Then we found that it fell exactly where it was most convenient for our being able to interpret the results.

Fig. 8-3b shows a corresponding high-pass network. Again we start with equation (8-6):

$$I_{out} = 19.2\ I_c\ E_{in} = V_{out}/R_{eq} \qquad (8\text{-}14)$$

We also observe that I_{out} now flows out through the capacitor C, generating V_{out}, relative to V_{in} as:

$$V_{out} = V_{in} - I_{out}/sC \qquad (8\text{-}15)$$

and these equations result in the high-pass T(s) as:

$$T(s) = scR_{eq}/(1 + sCR_{eq}) \qquad (8\text{-}16)$$

These are only two of the possibilities with this general idea.

The control current I_c is often supplied by some form of current source, which may be a linear voltage-to-current converter, or even an exponential voltage-to-current converter in the case of electronic music circuits. However, for simple demonstrations and lab work, it is often sufficient to connect a pot between the + and - supplies, and feed the wiper voltage to the control pin through a 30k resistor (Fig. 8-4), using equation (8-8) to find the current.

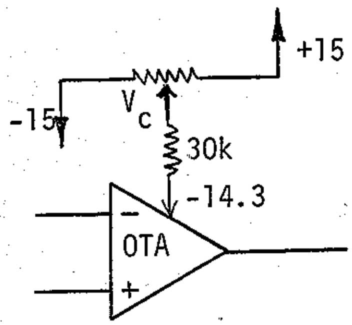
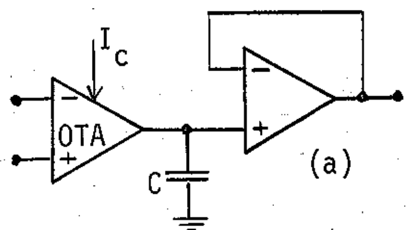


Fig. 8-4 A simple method of supplying the control current



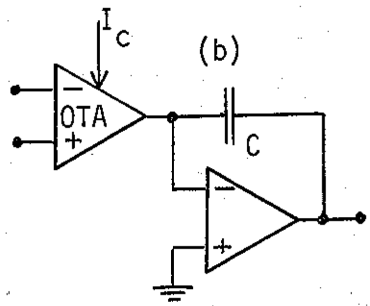


Fig. 8-5 Two forms of the VC Integrator

8-5 VOLTAGE-CONTROLLED INTEGRATORS AND STATE-VARIABLE FILTERS:

Fig. 8-5 shows two forms of an OTA controlled integrator. It is tempting to call the first a non-inverting integrator and the second an inverting integrator. However, because either the (+) or the (-) inputs of the OTA can be used in either case, we have free choice. In the case of Fig. 8-5a, the output voltage is:

$$V_{out} = I_{out}/sC \qquad (8\text{-}17)$$

while Fig. 8-5b has:

$$V_{out} = -I_{out}/sC \qquad (8\text{-}18)$$

both lead to transfer functions which have a denominator sCR_{eq}, which are integrators, and in this case, voltage-controlled integrators. With the voltage-controlled integrator available, we can achieve voltage-controlled versions of all our integrator based filters (Chapter 6). For example, Fig. 8-6 shows a voltage-controlled state-variable filter based on the ideas we have just discussed.

Fig. 8-6 shows one of the advantages of having both the (+) and the (-) inputs of the OTA available in our integrator designs. Here we have used the (-) input along with the inverting integrator structure,

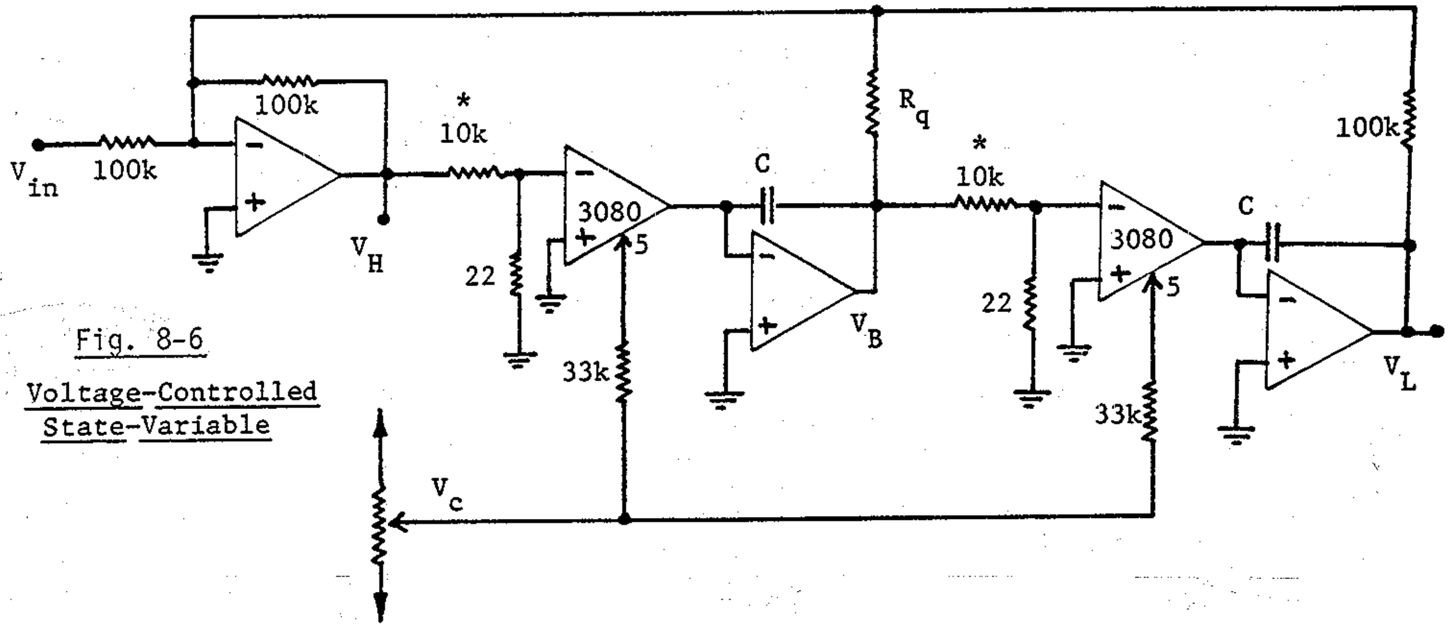


Fig. 8-6

Voltage-Controlled State-Variable

achieving a non-inverting integrator. This permits the feedback of the bandpass to the inverting input of the summer (through R_q), and the simple result that $Q = R_q/100k$ (compare Fig. 6-6).

This VCF and the others we discuss are of course subject to the active sensitivity problems that we have seen with all our fixed filters. However, the problem may be even more complicated in the case of a variable filter. This is because the active sensitivity error is a function of the characteristic frequency of the filter, and in the case of a variable filter, this frequency changes. Thus we can not simply fix the filter by overdesign, for example. We must more or less "repair" the various blocks, much as we discussed in Chapter 7, being always aware that here the resistors may have changing values. The problem is not at all simple.

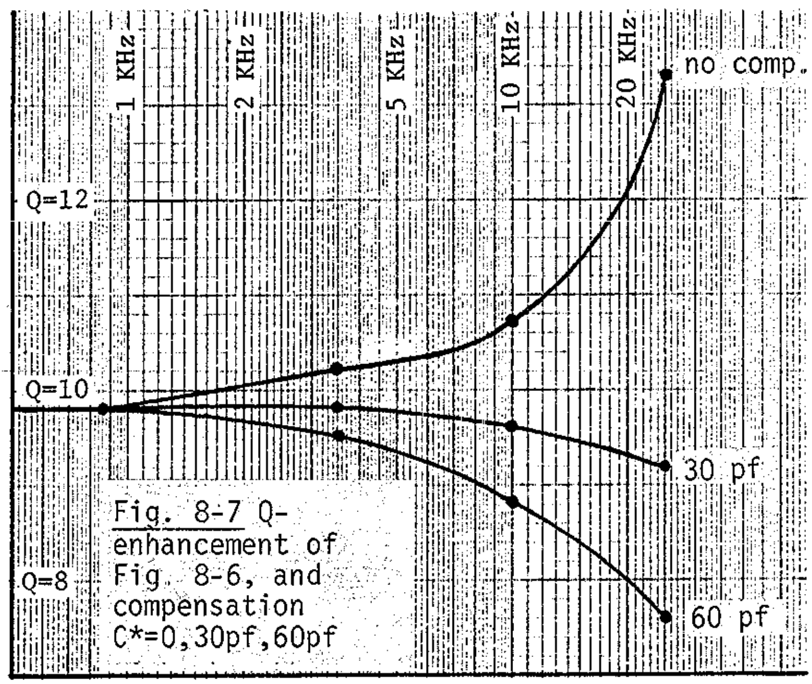


Fig. 8-7 Q-enhancement of Fig. 8-6, and compensation C*=0,30pf,60pf

Fig. 8-7 shows how the Q of the VCF of Fig. 8-6 changes with frequency, showing an enhancement with frequency. This can be so severe that the filter will break into oscillation when it is set to its upper frequency range. Traditionally, the "cure" for this has been to place small shunt capacitors across the OTA attenuator resistors (the 10k resistors marked with a * in Fig. 8-6). Capacitors in the range of a few picofarads to up to 50 picofarads were usually sufficient to level off the Q vs. frequency curve. In fact, this does seem to work, but a more exacting analysis (B. Hutchins, "Some New Results Concerning Q-Enhancement in OTA-Based VCF's," Electronotes, Vol. 14, No. 141, September 1982, pp 3-18) indicates that it is the summer of the state-variable rather than the integrators that is the real problem. A rather careful modeling of the whole network, and appropriate compensation methods on a block-by-block basis, can yield a very flat Q vs. frequency curve. Another unexpected result is that the integrator in the form of Fig. 8-5a is somewhat preferred over the integrator of Fig. 8-5b. In might have been thought that it would have been better to have the OTA driving into a constant ground potential (8-5b) than to have it drive into the variable output potential (8-5a). However, when stray capacitance effects are taken into consideration, the advantage goes to Fig. 8-5a.

ASP 8-10

CHAPTER 9

FILTERING WITH ANALOG DELAY LINES

9-1 INTRODUCTION TO DELAY LINE FILTERING:

In this chapter, we will look at various types of comb filters that can be realized using analog delay lines. The subject matter here is very closely related to digital filtering in that the essence of the filtering is time delay, and in that many of the same design and analysis techniques are employed. The main differences is that we work with a pure analog delay, and in a first approximation, no sampling is involved. This permits us to take advantage of the periodic frequency response in the case of analog delay line filters, while in the case of digital filters, only the first half of the unit circle in the z-plane is used, since the sampling theorem must be obeyed.

Comb filters find applications in cases where a number of harmonically related components must be filtered in a similar manner. For example, we might have a complex waveform, containing a fundamental and harmonics, which is to be notched out. We could envision a set of second-order notch filters in series, each responsible for its own harmonic. The comb filter however has a built-in periodic response (for example, Fig. 9-1) and thus one filter can takes out multiple harmonics. Moreover, once the filter is tuned to notch out any one harmonic, the others are automatically tuned. We are not restricted, however, to taking out harmonics with notch-like responses shapes - we can also enhance all harmonics with bandpass-like response shapes, and so on.

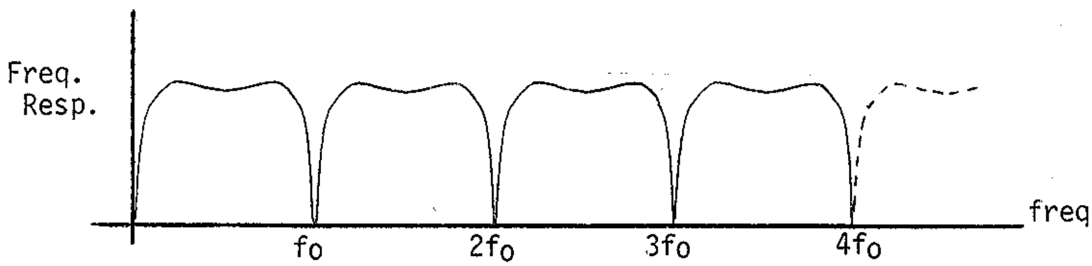


Fig. 9-1 A Typical Comb Filter

9-2 THE IDEAL ANALOG DELAY AND SOME REALIZATIONS:

The ideal analog delay line is shown in Fig. 9-2. A signal that is at the input of the delay at time t emerges after a delay of time T, at time t+T. Here T is not to be considered a sampling time, since we are not necessarily assuming that any sampling is taking place. However, the signal is only available to us at the input and at the output of the delay - at two discrete times separated by the interval T.

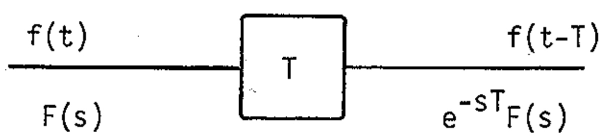


Fig. 9-2 Ideal Time Delay

Analog delays can be realized (or occur naturally) in a variety of cases. Transmission lines may be used for an analog delay (usually only for very short nanosecond delays), or we may need to deal with the analog delay of a transmisson line that we are working with. For delays from 100 milliseconds up to several seconds, magnetic tape recording can be used, with a delay corresponding to the tape speed and the distance between record and playback heads. Active filter all-pass or phase shift networks also can look like a pure delay over a limited range of frequency (see problems at end of chapter). Surface acoustic wave (SAW) devices may also be considered.

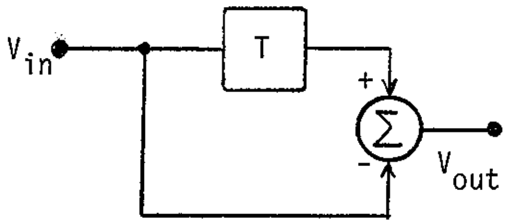
Probably the delay lines of most interest and most practical value are those which do involve sampling. These include digital delay lines (DDL), and charge-coupled device (CCD) delay lines. While these devices do involve sampling, specifically the sampling frequency f_S is not $1/T$, but rather some significantly higher frequency, usually an integer multiple of $1/T$. This means that there are, at any one time, not only a sample at the input, and another sample at the output, but also many samples in between, that are held internally. These are clocked along at the rate f_S. Thus if there are N samples between the input and the output of the delay line, $T = N/f_S$.

At this point, there are two notions of time interval that are of interest. The first of these, $1/f_S$, is the actual "sampling interval", and as with any sampled data system, we must not input frequencies in excess of $f_S/2$ or else we violate the sampling theorem, and aliasing can occur. The second time interval is T, and the frequency $1/T$ corresponds to a full trip around the unit circle in the z-plane. Not only can we input frequencies exceeding $1/2T$, but we can continue around the unit circle many times, until the frequency starts to approach $f_S/2$. In the limit of a perfect analog delay, equivalent to f_S becoming infinite, we can continue indefinitely to higher and higher frequencies, taking advantage of the repeating frequency response.

We will want to be able to write down networks involving analog delay lines, and to solve for transfer functions much as we have been doing. We need to know how a delay affects the Laplace transform of a signal. We can show, using equation (1-3) for the Laplace transform, that if F(s) is the Laplace transform of f(t), then $e^{-sT}F(s)$ is the Laplace transform of f(t-T). Thus when a signal passes through a delay T, it is equivalent in the Laplace domain to a multiplication by e^{-sT}, which is also often written as z^{-1}. We will in general assume that we are using perfect analog delays in the sections that follow, with the idea that various realizations of the delay may present individual problems to consider.

9-3 FIRST-ORDER NON-RECURSIVE COMB FILTER - FOUR METHODS OF ANALYSIS:

Fig. 3-3a shows a simple use of a delay line and a summer to form a non-recursive (no feedback) network. This network has the frequency response as shown in Fig. 3-3b. We will look at four ways to analyze this network.



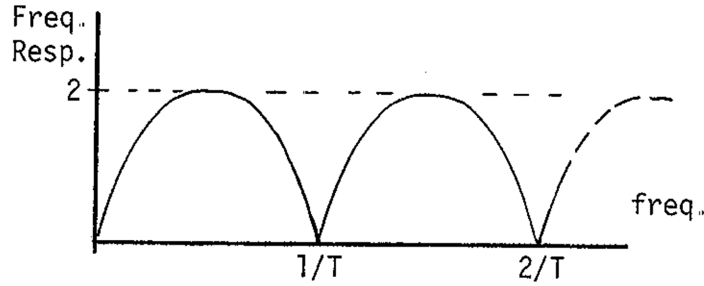


Fig. 9-3a Non-recursive   Fig. 9-3b Freq. Resp.

For the first method, we will begin with the idea that the frequency response is the ratio of the amplitude of a sinusoidal at the output to the amplitude of the sinusoidal at the input. We can take the input sinusoidal to be $\mathrm{Sin}(\omega t)$ in which case the sinusoidal at the output of the delay is $\mathrm{Sin}(\omega t - \omega T)$. Therefore, at the output of the summer, the voltage is:

$$V_{out} = \mathrm{Sin}(\omega t - \omega T) - \mathrm{Sin}(\omega t)$$

$$= -2\mathrm{Sin}(\omega T/2)\mathrm{Cos}(\omega t - \omega T/2) \qquad (9\text{-}1)$$

which is the result of the trig identity for the sum of two sines. Note that the $\mathrm{Cos}(\omega t - \omega T/2)$ term is the output "sinusoidal", having frequency ω and phase $-\omega T/2$. The term $2\mathrm{Sin}(\omega T/2)$ does not vary with time, and determines the amplitude of the output sinusoidal, and is accordingly the frequency response, which we will denote by the digital filter frequency response notation:

$$|H(z)| = |2\ \mathrm{Sin}(\omega T/2)| \qquad (9\text{-}2)$$

This is the function plotted in Fig. 9-3b.

In the second method, we will look at the transfer function $H(z)$, where, using the z^{-1} notation, for the Laplace transform of a delay, we have:

$$V_{out}(z) = -V_{in}(z) + V_{in}(z)z^{-1} \qquad (9\text{-}3)$$

or:

$$H(z) = V_{out}/V_{in} = z^{-1} - 1 \qquad (9\text{-}4)$$

We have seen that the frequency response can be obtained from a transfer function using equation (1-18), and this can be applied here, giving:

$$|H(z)| = \left[H(z=e^{j\omega T})H(z=e^{-j\omega T})\right]^{1/2}$$

$$= \left[(e^{-j\omega T}-1)(e^{j\omega T}-1)\right]^{1/2}$$

$$= \left[2 - 2\mathrm{Cos}(\omega T)\right]^{1/2}$$

$$= |2\mathrm{Sin}(\omega T/2)| \qquad (9\text{-}5)$$

which is the same result we got with trigonometry.

The third method is to use a geometric interpretation of frequency response, which is very similar to that used for the s-plane filters, except here we are looking for the response on the unit circle in the z-plane. This we can understand in terms of the $j\omega$-axis in the s-plane becoming the unit circle in the z-plane, since:

$$z = e^{sT} = e^{(\sigma + j\omega)T} = e^{\sigma T}\left[\mathrm{Cos}(\omega T) + j\mathrm{Sin}(\omega T)\right] \qquad (9\text{-}6)$$

Now the geometrical interpretation follows as seen in Fig. 9-4. From equation (9-4), we see that H(z) has no poles, but does have a zero at z=+1. The frequency response is proportional to the distance from this zero, which is the distance r shown in Fig. 9-4. From simple trig:

$$r = 2\ \mathrm{Sin}(\Theta/2) \qquad (9\text{-}7)$$

Further, once around the unit circle corresponds to a frequency of 1/T, so here the frequency is:

$$f = (\Theta/2\pi)(1/T) \qquad (9\text{-}8)$$

or $\omega = 2\pi f = \Theta/T$, and:

$$|H(z)| \propto r = 2\mathrm{Sin}(\omega T/2) \qquad (9\text{-}9)$$

which is again the same result.

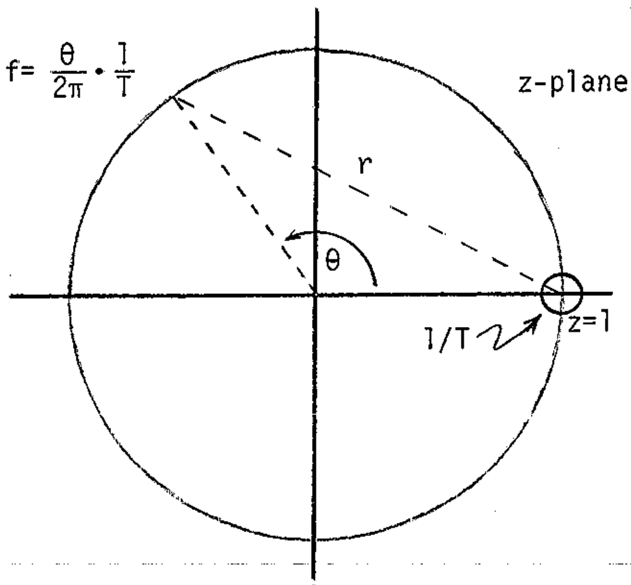


Fig. 9-4 Geometric Method

The fourth method involves inversion of the impulse response of the network as an alternative way of obtaining H(z). Clearly if we put in an impulse, it comes out inverted immediately and then right side up after a time T. From this we have for the inverse Laplace transform the result $-1 + e^{-sT}$, which is the transfer function, the same result as we got from equations (9-3) and (9-4). From this, the frequency response can be obtained as before.

We see from the frequency response that it consists of sinusoidal lobes, with equally spaced notches at frequency intervals of 1/T. Accordingly, it is capable of cancelling all frequencies that are harmonics of 1/T. In the case of a periodic waveform with fundamental 1/T, that means that all harmonics are cancelled. Thus the entire waveform is cancelled. This result is at first impressive, but less so if we simply consider Fig. 9-3a in the time domain. For a periodic waveform of period T, we always have exactly the same voltage at the input and the output of the delay line. Subtracting these of course results in zero output at all times.

Other forms of non-recursive comb filters are sometimes found. The form shown in Fig. 9-5a sums rather than subtracts the two signals. This results in a different frequency response, as seen in Fig. 9-5b. This form is sometimes called the "delay-add" type (or Cosine type) of comb filter, as opposed to the "delay-subtract" type (or Sine type) of Fig. 9-3a. The spacing of nulls is the same (1/T) in both cases, but they are displaced by 1/2T relative to each other. The delay-add type here can be seen to remove a fundamental at 1/2T, and all odd harmonics of 1/2T. It is easy to remember which type is which, simply by considering what happens at dc. At dc, the time delay of the delay line "expires" and the input and output of the delay line are both the same dc voltage. If we subtract these, we get zero (Sine type frequency response) while if we add them, we get 2 (Cosine type frequency response).

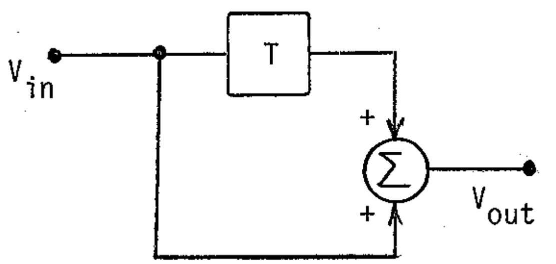


Fig. 9-5a "Delay-Add"

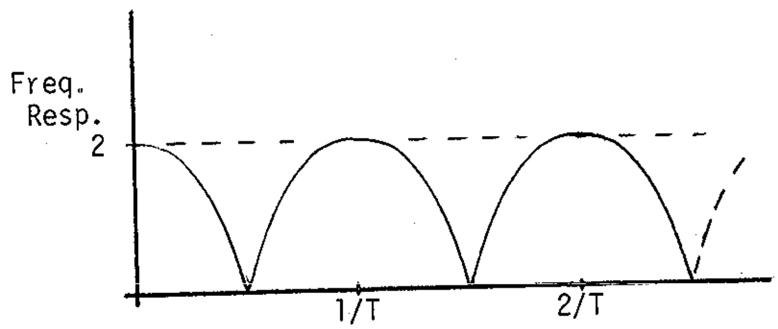


Fig. 9-5b Freq. Resp.

Another variation would be to make an unequal weighting of input and output, which moves the zero off the unit circle. This can result in attenuation in certain frequency regions (valleys of the response), but not a complete null. Since the summation is usually a matter of op-amp summers with summing resistors, the incomplete null is actually what we have in practice, although trimming of resistors can be used to get very good rejection.

9-4 THE FIRST-ORDER RECURSIVE NETWORK:

The non-recursive networks above have resulted in notch-like responses since they are based on zeros, and no poles. We can use a recursive structure (with feedback) to give poles, and corresponding responses that are more bandpass in nature. (Eventually, or course, we consider both together). Fig. 9-6a shows the first-order recursive network, while Fig. 9-6b shows the position of its pole, and Fig. 9-6c its frequency response.

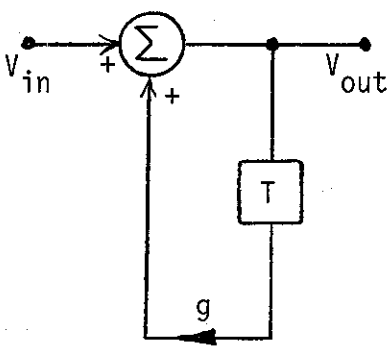


Fig. 9-6a
Recursive Structure

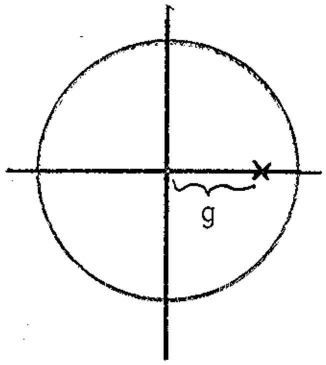


Fig. 9-6b
Pole Plot

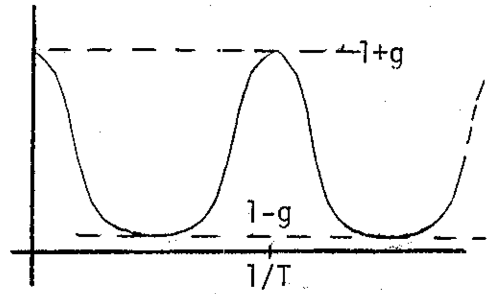


Fig. 9-6c
Freq. Resp.

The transfer function of the recursive filter is given by:

$$H(z) = 1/(1 - gz^{-1}) \qquad (9\text{-}10)$$

which has a pole at z = g. For stability, $|g|$ must be less than 1 so that the poles is inside the unit circle. For positive values of g, the pole is approximately as shown in Fig. 9-6b, and the response peaks at dc and at multiples of 1/T, as shown in Fig. 9-6c. For negative values of g, the pole is on the left side, and the response peaks at 1/2T and at odd harmonics of 1/2T. Clearly the recursive filter is suited to cases where frequency components are to be enhanced. It is easy to obtain the frequency response by any of the methods suggested above for the non-recursive network, and the result is:

$$|H(z)| = 1/\sqrt{(1 - 2g\mathrm{Cos}(\omega T) + g^2)} \qquad (9\text{-}11)$$

From this, or from a geometric interpretation evaluated at z=+1 and z=-1, we can see that the "peak-to-valley" ratio in the response is given by:

$$|H(z)|_{max}/|H(z)|_{min} = (1+g)/(1-g) \qquad (9\text{-}12)$$

which is valid for positive g (and is the "valley-to-peak" ratio for negative g).

9-5 NOTCH AND ALL-PASS RESPONSES:

Fig. 9-7a shows a first-order delay line notch filter, while Fig. 9-7b shows the pole/zero plot, and Fig. 9-7c the frequency response. Here the response is reminiscent of the non-recursive comb filter of Fig. 9-3 in being notch-like, due to the zero on the unit circle. However, as in the case of active filter notch circuits, the pole that appears here is useful in sharpening the notch and in flattening the passbands, relative to the sinusoidal lobes of Fig. 9-3b. This is because the pole, being brought up close to the zero here, tends to "hide" the zero until frequencies get very close to the zero.

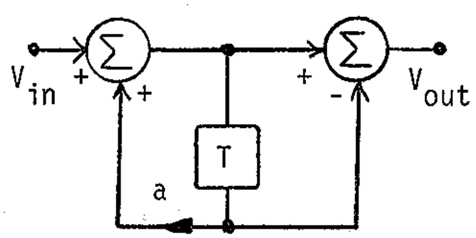


Fig. 9-7a Notch Network

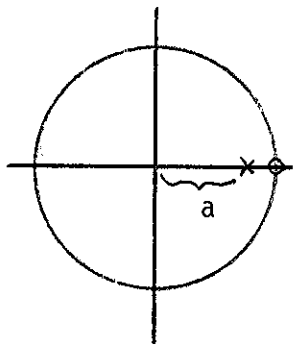


Fig. 9-7b Pole/Zero

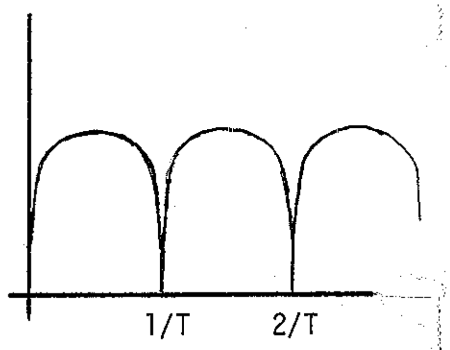


Fig. 9-7c Freq Resp.

Fig. 9-8a shows a delay line all-pass network, while Fig. 9-8b shows the pole/zero plot, and Fig. 9-8c the (completely flat) frequency response. Here the poles is at a position a, within the unit circle for stability, while the zero is outside the unit circle, and at a radius 1/a that is reciprocal to that of the pole. It can be shown (see problems at end of chapter) that any point on the unit circle is at relative distances to the pole and to the zero that are always proportional, hence the all-pass magnitude response.

It is probably evident how the transfer functions and pole/zero plots are derived for these networks, and we have left this out, since it is really a combination of the derivations found above. It should also be recognized that there are a number of variations on these circuits that are sometimes seen.

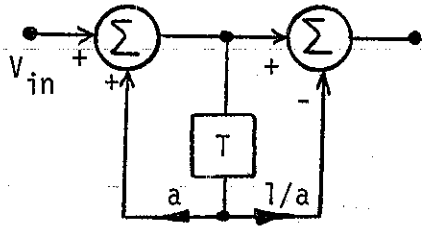


Fig. 9-8a All-Pass

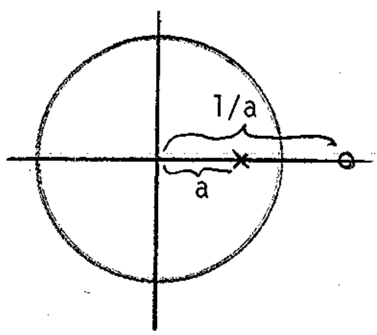


Fig. 9-8b Pole/Zero

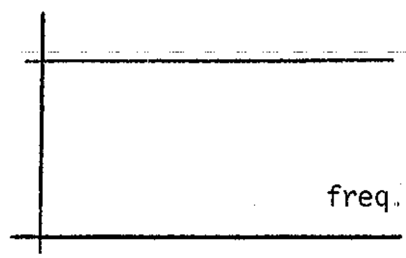


Fig. 9-8c Freq. Resp.

9-6 SECOND ORDER NETWORKS:

In the case of first-order analog networks, there was relatively little of interest that could be achieved, and we had to go to somewhat higher order networks to achieve interesting results. In this delay line case, things are somewhat different in that the first-order networks already have the interesting and useful property of a periodic frequency response already built in. Nonetheless, second-order and higher order delay line networks can be of interest to us. In going over to these second-order designs, we will need to make use of some digital filter design methods, and in particular, the "Bilinear-Z Transform" method is useful.

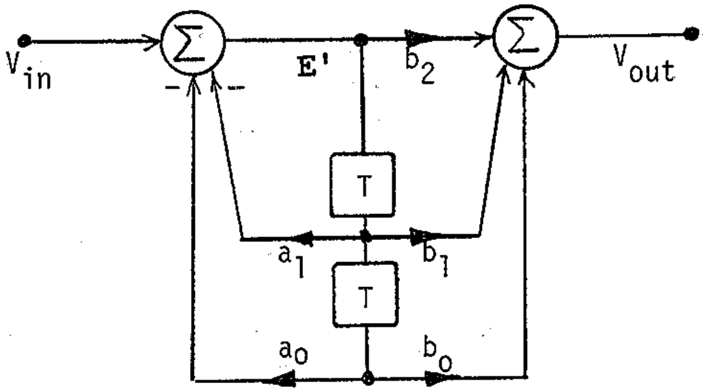


Fig. 9-9a 2nd-Order Network

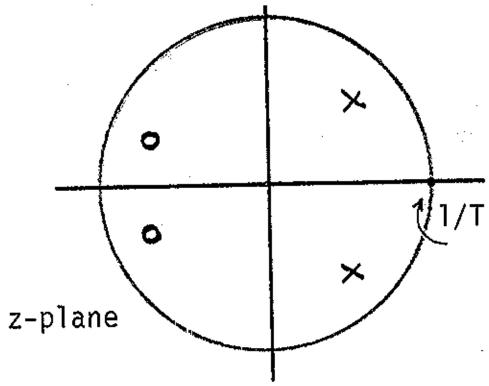


Fig. 9-9b Pole/Zero

Fig. 9-9a shows a second-order delay line network (note the two delay lines), while Fig. 9-9b shows a typical pole/zero plot, by way of showing that it is capable of producing two poles and two zeros. It is easy to derive the transfer function be noting that the voltage E' is given by:

$$E' = V_{in} - a_1 E' z^{-1} - a_0 E' z^{-2} \qquad (9\text{-}13)$$

The output is given by:

$$V_{out} = b_2 E' + b_1 E' z^{-1} + b_0 E' z^{-2} \qquad (9\text{-}14)$$

from which the transfer function V_{out}/V_{in} is given as:

$$H(z) = (b_2 z^2 + b_1 z + b_0)/(z^2 + a_1 z + a_0) \qquad (9\text{-}15)$$

The values for the coefficients can be determined by starting with an analog prototype transfer function and plugging into the bilinear-z transform. That is, we start with T(s) and make the substitution:

$$s \longleftarrow F(z-1)/(z+1) \qquad (9\text{-}16)$$

For example if we substitute into a normalized low-pass transfer function $T(s) = 1/(s^2 + Ds + 1)$, we arrive at H(z) given by:

$$H(z) = A(z^2 + 2z + 1)/(z^2 + Bz + C) \qquad (9\text{-}17)$$

where:

$$A = 1/(F^2 + DF + 1) \quad (9\text{-}18)$$

$$B = A(2 - 2F^2) \quad (9\text{-}19)$$

$$C = A(F^2 - DF + 1) \quad (9\text{-}20)$$

In terms of the network of Fig. 9-9a, we have:

$$b_2 = 1 \quad (9\text{-}21)$$

$$b_1 = 2 \quad (9\text{-}22)$$

$$b_0 = 1 \quad (9\text{-}23)$$

$$a_1 = B \quad (9\text{-}24)$$

$$a_0 = C \quad (9\text{-}25)$$

which completes the network.

As with the second-order active network, relating the network parameters to a transfer function was only one step in moving toward a useful network. Here we need to find how to choose T, D, and F in order to achieve the frequency response that we want. Two of these we can deal with easily. First, 1/T remains the spacing between periodic repetitions of the response. Secondly, it is the property of the Bilinear-z transform that the response shape is carried over from the analog to the discrete time case. This means that if we choose a value of D that gives, say, 2db passband ripple in the case of T(s), this same value of D will give 2db ripple in H(z).

This leaves us with the parameter F to manipulate to our advantage if possible. In addition, we should keep in mind that we could also have used other second-order T(s) instead of just the low-pass we actually chose. For the low-pass, we get two zeros at z=-1, which can be seen from equation (9-17). If we instead choose a high-pass T(s), we get two zeros at z=+1 . (Incidentally, a bandpass T(s) gives one zero at z=-1 and the other at z=+1.)

In considering our options, it is perhaps well to keep in mind that we are generally after only one or two types of response. This is because we are taking advantage of the built-in periodicity to handle a fundamental and all its harmonics. In general, we either want to enhance this complex waveform, or we want to reject it, so we are interested in one of the two responses shown in Fig. 9-10. (Here we are assuming that the damping D is Butterworth or larger, since no ripple is seen in the responses.) The pole/zero positions corresponding to these two responses are also seen. Thus we see that the enhancement case will result from a low-pass prototype while the rejection case will result from a high-pass prototype. In both cases, we are looking for poles that are inside the circle, and relatively close to z=+1. Some additional case of interest are covered in the problems at the end of the chapter.

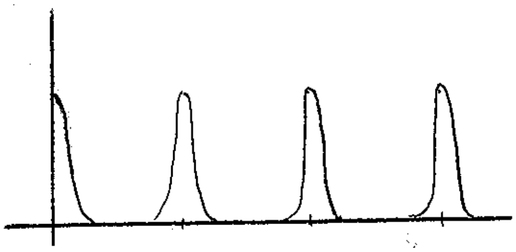


Fig. 9-10a Enhancement

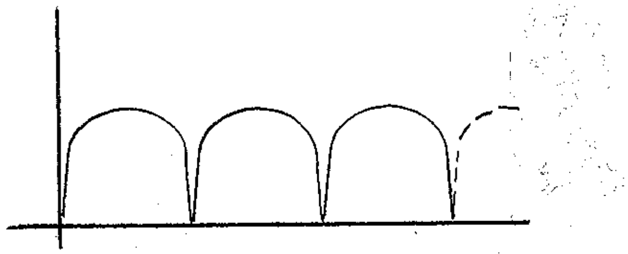


Fig. 9-10b Rejection

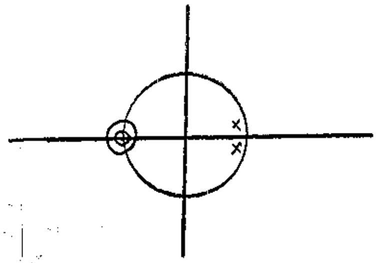


Fig. 9-10c Pole/zero for enhancement

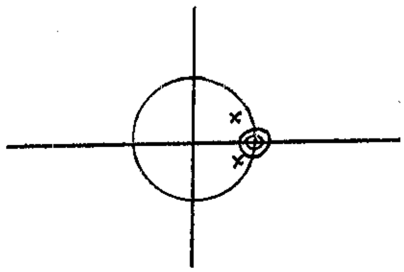


Fig. 9-10d Pole/zero for rejection

We are now left to consider the parameter F. This is needed in equation (9-17) to put in the units, if for nothing else. When used with digital filters, it is common practice to fix F at $2f_s$, but that is often a matter of convenience, or in considering of sampling problems, which we are not considering here. F can be seen as a design parameter that moves the poles to the right (larger F) or to the left (smaller F), but always along a curve such that a ripple corresponding to the selected value of D is achieved. If F = 1, the poles are on the imaginary axis. Values of F greater than 1 place poles to the right, while values of F less than 1 place the poles to the left.

CHAPTER 10

ANALOG ADAPTIVE FILTERING

10-1 THE NEED FOR ADAPTIVE OR SELF-ADJUSTING FILTERS:

We are familiar with the theory and design of filters which have fixed parameters (certainly by this chapter we know about fixed analog filters, and probably know about fixed digital filters as well). Such filters find wide application in cases where the filtering task to be performed is known ahead of time, and is expected to remain unchanged during use. Such filters have fixed parameters by virtue of the fact that they are constructed with fixed resistors and capacitors (analog filters) or with fixed multiplier coefficients and clock rate (if we also consider digital filters).

Since it is the fixed nature of certain of the filter's elements that make the filter itself fixed, we can obtain variable filters by arranging for these elements to vary. In the case of analog filtering, this is usually accomplished with variable resistors to control the filter's time constants. (Variable capacitors, while theoretically as useful as resistors for this purpose, are usually not practical.) The variable resistors may be manually controlled: potentiometers or manually controlled switches for combinations of fixed resistors. However, electronically-controlled resistors, such as those obtained with a transconductance multiplier (or similar), as we saw in Chapter 8, offer filters that can be controlled remotely, with great speed and accuracy, by computer or other control mechanism rather than by hand, and in an arbitrary number or with an arbitrary number of control elements involved.

Another point about variable filters should be made, and that is that these filters are, pretty much by definition, not time-invariant. Linear time-invariance is often one of our assumptions, leading to our most useful procedures. Obviously, a filter that is manually set (thus variable) and put in fixed service is pretty much the same in status as a filter that is "soldered-in" fixed, and this is basically a case of a programmable filter. In such a case, the time variability is not a consideration. At the other extreme, a voltage-controlled filter being controlled by a waveform of frequency comparable to its characteristic frequency is perfectly capable of producing detectable, non-harmonic, modulation sidebands. In such a case, our usual ideas about frequency response and the like, which we find so useful in the case of fixed filters, are not applicable.

Between these cases of filters which are programmable, but otherwise relatively fixed, and those with variations producing modulation effects of significance, we can sometimes find a useful "quasi-stationary" region where, from a theoretical and performance standpoint, fixed filter techniques are still useful. As a guide, a time constant rule-of-thumb can be applied: we want the change in characteristic frequency to take place on a time scale that is long compared with one cycle of that characteristic frequency.

Such slow frequency adjustments are what is found in many cases. Manually adjustable filters where the frequency is reset for different parts of an experiment would be an obvious example. Another example would be an automatic but sufficiently slow sweep found in some bandpass frequency analyzers. We also have cases where a filter should change its performance on a slow time scale in response to some slowly changing external condition.

For example, we might be trying to filter out some frequency that is nominally 400 Hz, which appears as noise on an aircraft's communication system. However, due to changes of engine load, this frequency, determined by onboard generators, is subject to some variation. In such a case, a notch filter might be trimmed manually when the interference becomes a problem. In a corresponding ground-based case, the 60 Hz power lines might be more precisely regulated in frequency, but an analog notch filter might still be subject to drift due to external temperature variations or other such changes, and this would need to be trimmed up from time to time as well.

These and other similar situations indicate the <u>need for a variable or tunable filter</u>, but they also indicate the <u>desirability of a self-tuning or adaptive filter</u>. That is, we would like to <u>not</u> have to adjust the filter manually, but rather have it control and adjust itself, according to some desired performance criterion. This automatic change could be in response to changing input conditions, which could be detected, or it might be in response to the filter's own evaluation of its own performance level at its output, or to both. This leads us to the topic of adaptive or self-adjusting filters.

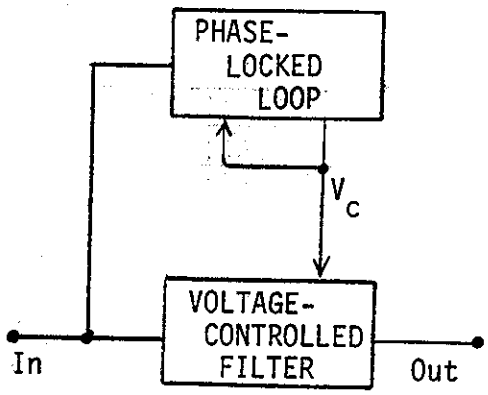
<u>10-2</u> <u>BASICS OF ADAPTIVE FILTERING</u>:

We will be using the term "adaptive filter" in a fairly general sense to include all variable filters that are capable of adjusting their own parameters, in response to signal conditions, so as to better perform their intended functions. However, at the same time it should be realized that by tradition, the term "adaptive filter" has been used in a more restricted sense to describe only an adaptive filter of the linear combiner type (an FIR digital filter), or its operating mode called the "LMS algorithm", or only the linear combiner portion of the structure. (At the same time, the almost identical analog counterpart to this digital FIR one, the so-called "Correlation Cancellation Loop", or CCL, had been largely ignored.) While this particular adaptive filter and the surrounding theory of this digital point of view are extremely interesting and useful, we prefer to use the term more openly, and correspondingly, to suggest a broader range of possible solutions to self-tuning problems in signal processing. The specific approach chosen will then depend on the application and the available resources.

In the examples suggested above - that of cancelling power supply "hum" - we could take a variety of approaches. Some sort of tunable notch filter comes to mind first, and we need to consider how we would recognize that the notch were not correctly positioned. Obviously this would be the recognition that the unwanted signal were coming through, but this only tells us that the notch is not properly trimmed; it in itself does not tell us which way to move the notch position (up or down) to reduce the level of the undesired signal.

One simple approach to self-tuning would be to have a Phased-Locked Loop (PLL), with properly determined capture and hold properties, lock

Fig. 10-1 PLL Tracking



on to the undesired signal, and the feedback voltage in the PLL could in turn be used to tune a voltage-controlled filter (VCF), as seen in Fig. 10-1. A second approach would be to use a Side-Tracking Filter (STF), which is a generalization of a comb-filter technique. As discussed above, we often do not know if a filter's frequency should be increased or decreased in order to perform better in a given case. The idea of the STF is to have two filters, in addition to the main one, above and below (on the sides) which are evaluating the possibilities of changes in the respective directions. A feedback mechanism than adjusts the center or main filter in response to these findings. We will look at STF's in more detail a bit later.

The PLL approach, and the STF approach offer two useful analog techniques for self-tuning filters. The third analog technique that we want to have in our "bag of tricks" is the CCL (Correlation-Cancellation Loop), which is the analog counterpart to the digital adaptive filter (or LMS algorithm). We will later spend a good deal of time on the CCL structure. First however we will take a brief look at the digital adaptive filter, in order to better relate to the CCL when it come up, and for a better understanding of how adaptive filters work.

Fig. 10-2 shows an adaptive filter structure of the digital or LMS type. We can just think of the z^{-1} boxes as delay lines, in which case it is clear that each of the taps available on the line represent only different phases of the reference signal that is shown. We assume that the input is an information-bearing signal such as speech or music, and that added to it is a large "hum" component due to the AC power lines. Such a signal can result from poor grounding practices, for example. In the figure, the hum is the larger sinusoidal-like component while the speech or music is the smaller random-like component. The approach seen in Fig. 10-2 is to take advantage of a reference signal. In this case, we assume that the hum is caused by the power supply lines, and that we have separate access to the power supply lines, for reference purposes. We can think of the reference as a signal that gives us some information on the hum component - at least the correct frequency, and possibly more. In the example, we are further assuming that the waveforms of the reference and of the hum component are both sinusoidal.

The basic idea here is to take the reference and subtract it off, cancelling the hum. This would be easy if the waveform, the frequency,

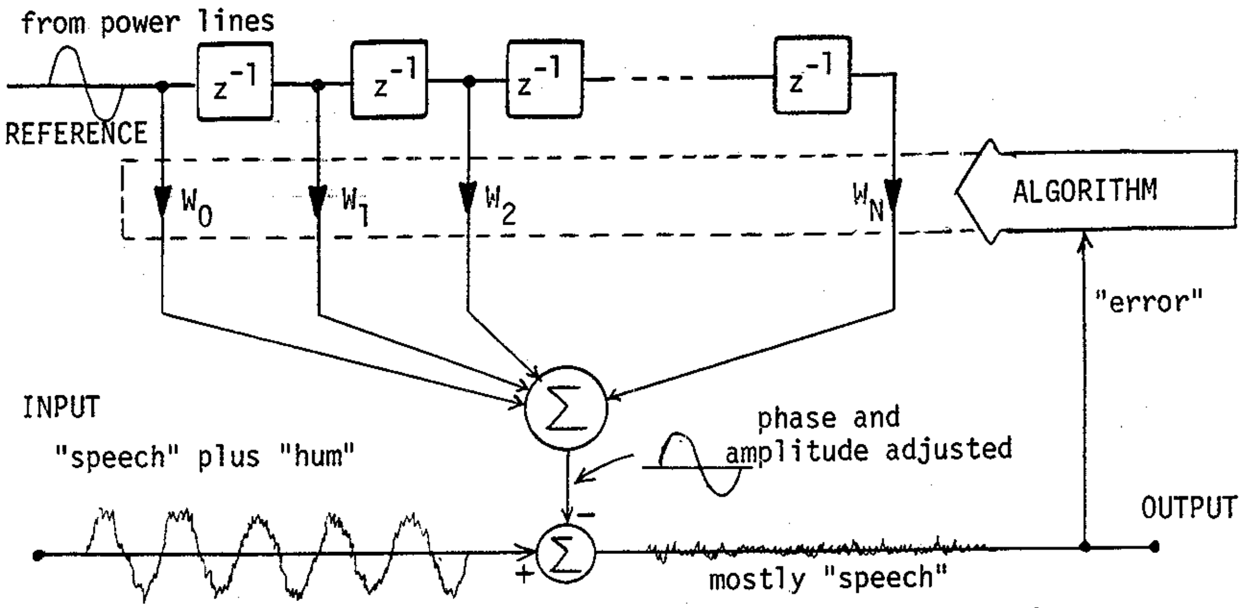


Fig. 10-2 An adaptive filter using a linear combiner (FIR filter) to adjust phase and amplitude of a reference input to cancel the "hum" component from a desired "speech" signal.

the amplitude, and the phase were all the same. Clearly this is more than we can expect. However, it is the purpose of the adaptive filter to adaptively find the correct phase and amplitude (and possibly even more) to achieve the correct cancellation. With our assumption that it is only the phase and the amplitude that are unknown, we can see that we can choose from the variety of phases presented by the delay lines, and adjust the tap weights W_k for the amplitude. Moreover, simple geometric constructions convince us that all we need is two different phases, and we can find amplitudes (tap weights) such that any amplitude and phase condition can be met. Thus, for our example, we would really only need two taps of the N taps shown.

Of course, to have the necessary adjustment ability available is important, but that is only part of the solution. We need to have an automatic adjustment procedure in place, and this we have conveniently ignored by hiding it in the box called "algorithm". We can better understand what is in the box, and how the adaptive filter can work, after we have studied the CCL. However, we can indicate here that it is a matter of looking at the output (called the "error" here - an unfortunate but traditional nomenclature), and seeing if it is correlated with a given tap's version of the reference signal. If it is, then we know two things. First, we have not gotten rid of all the hum, since there is still some in the output. Secondly, the particular tap in question is capable of providing a contribution to improving the situation. Further, we shall see that an algorithm can be chosen so that the correction is in the right direction at all times. We shall return to this a bit later.

10-3 THE SIDE-TRACKING FILTER (STF):

The Side-Tracking Filter (STF) is intended to be self-tracking through the operation of two filtering channels on either side of a main channel. Essentially we are looking for these channels to examine the possibility that the main channel should be moved in their direction. The principle is most useful in the case where a change of output amplitude level in the main channel is not in itself indicative of the direction in which the filter should be retuned for better performance. These cases are mainly those of bandpass and notch responses which have equal amplitude points on either side of a center frequency. For the most part, we are looking to apply STF ideas to case where an interference is relatively strong and relatively (but not absolutely) stationary.

Fig 10-3a shows the STF principle applied to the bandpass case. This would be useful in cases where we have a single sinusoidal component to be tracked and enhanced. The STF here is composed of three voltage-controlled bandpass filters. The center filter is the main processing channel. Above and below we have two side filters that are tuned a bit above and a bit below the main channel. The outputs of the two side filters are not used except that their amplitudes are detected by the magnitude circuits (usually full-wave rectifiers). These outputs are in turn fed to the inputs of a differential integrator, the output of which is the control voltage for all three VCF's. (Here we

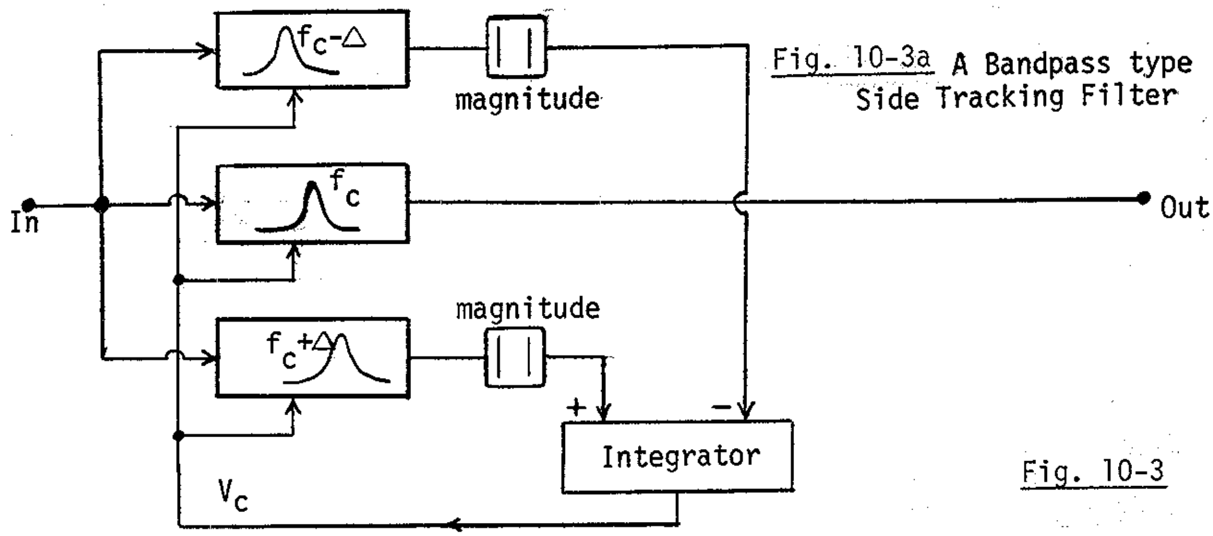


Fig. 10-3a A Bandpass type Side Tracking Filter

Fig. 10-3

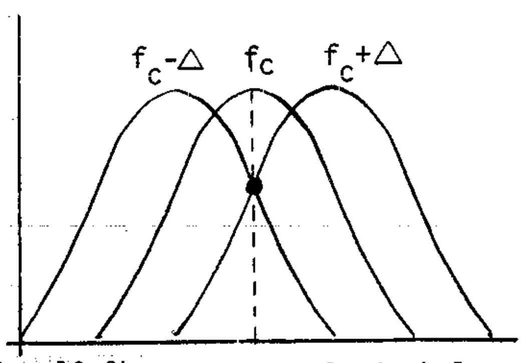


Fig. 10-3b Side channels in balance used to hold center channel in proper position.

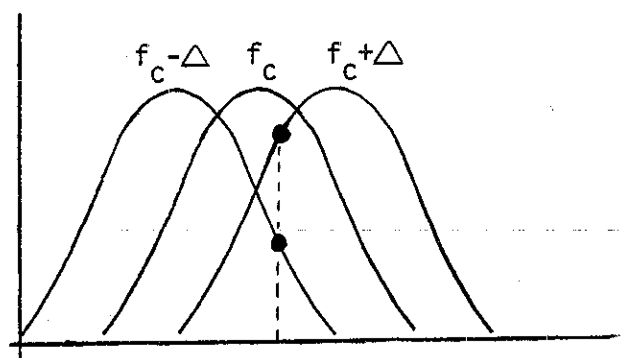


Fig. 10-3c If input moves upward as shown, side channels are no longer in balance, and differential integrator ramps upward.

are assuming that the frequency offset of the side channels has been set by a voltage offset not shown, or perhaps by using slightly different capacitors in the three channels.) When the control voltage V_c changes, all three filters shift, but maintain relative positions.

Fig. 10-3b shows the case which illustrate how the side channels lock the center channel in place. In this case, the amplitudes of the two side channels are exactly the same, and the differential integrator has a net input of zero, and thus its output V_c stands still. Next suppose that the input sinusoidal moves up in frequency slightly, as in Fig. 10-3c. Now the two side channels will be out of balance, with more amplitude in the upper on, and less in the lower one. This will cause the differential integrator to ramp upward, moving all three channels upward until the center is on the new frequency, and the side channels are balanced again, in a manner similar to Fig. 10-3b. Note that if the input frequency had drifted downward instead of upward, exactly the opposite thing would have happened, and balance would have been restored at a lower V_c. The basic feedback operation here is not at all unfamiliar, being like PLL's and other negative feedback devices.

We can see that the circuit is capable of capturing as well as tracking. If an input signal appears roughly within Δ of the current center frequency, the filter can capture and lock in the same manner in which it responded to a change of input frequency. With this in mind, we can see that in some cases there would be an advantage to keeping the side filters of relatively low Q and somewhat further from the center channel, if we desire a wider capture range. At the same time, the center channel need not have this same lower Q. In fact, the center filter need not even be bandpass, but could be notch, or even low-pass or high-pass if the situation dictates this sort of need. Fig. 10-3a thus is representative of a fairly general idea for tracking filters. There is a wide flexibility with respect to the nature of the center channel, and even a good deal with respect to the side channels which could be bandpass, notch, or even a combination of high-pass and low-pass, which can be illustrated by a simplification discussed immediately below.

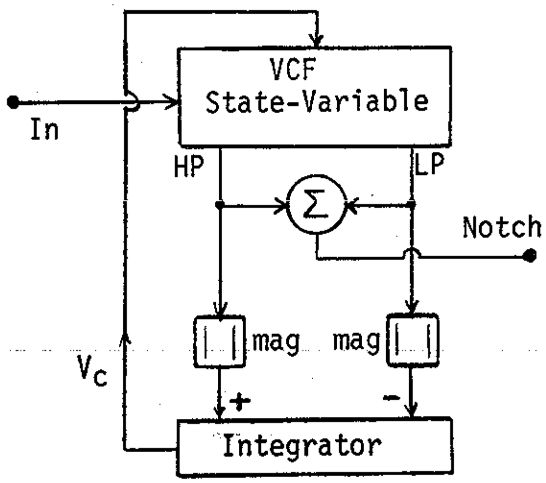


Fig. 10-4a A single VCF state variable filter may be used to do its own side-tracking under certain circumstances.

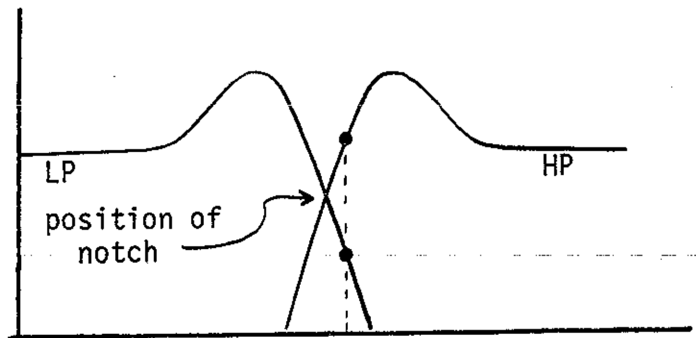


Fig. 10-4b An out of balance condition that would cause the frequency of the VCF of Fig. 3d to move up so that the overlap point matches the new input frequency.

Fig 10-4a shows the side tracking idea extended and simplified so that only one VCF is needed. Here we show a state-variable VCF (which would probably have been our choice for a VCF above anyway) which has a low-pass and a high-pass output as well as the bandpass. In Fig. 10-4a, we show yet a third response - a notch, formed by summing the low-pass and high-pass - as the output, but any of the output could be used. Here the capture and locking mechanism is indicated in Fig. 10-4b, which corresponds to Fig. 10-3c for the bandpass case. In the specific instance shown, the input frequency is a bit above the center frequency of the state-variable filter, and there is more amplitude in the high-pass than in the low-pass, This will cause the differential integrator to ramp upward. Note that here we don't have quite the freedom we did in the three VCF case in that the side filters must have the same Q as the center, which in many cases is not a problem.

10-4 CORRELATION-CANCELLATION-LOOPS (CCL'S)

A CCL is a configuration of two multipliers, an integrator, and a summer as shown in Fig. 10-5a. The CCL in itself is a simple adaptive filter, and it can also be used as an element in a more complex adaptive filter structure, such as serving as the "algorithm" of Fig. 10-2. Because it functions in close analogy with the so-called "LMS algorithm" of the digital adaptive filter, once we understand how the CCL works we can better appreciate the LMS algorithm.

Fig. 10-5a and Fig. 10-5b indicate how the CCL is used to cancel the "hum" component from an input signal containing a mixture of speech and hum. Note that we assume here that we have a reference to the hum available, and we shall also assume in this case that the reference and the hum are in phase with each other. Later we can look at a more general case.

In Fig. 10-5a, we are assuming that the integrator has been reset so that its output is initially zero. This in turn blocks the reference signal from the summer since there is a zero voltage on the right input of the lower multiplier. Therefore, the output of the filter is the same as the input, since nothing else is fed to the summer. However, note that the output is being fed back and is being multiplied by the reference at the top multiplier. Since the output is the same as the input for the moment, and since the input is in phase with the reference, and both are sinusoidals or close to being sinusoidal (the "speech" being the smaller random-like component on the input), the output of the upper multiplier is much like a Sin^2 function, which is only positive, as shown. Next we consider releasing the integrator from this reset condition.

The Sin^2 component at the input of the integrator causes the integrator to ramp positive, which in turn causes the lower multiplier to start passing some of the reference signal. Since the two are in phase, this subtraction results in the sinusoidal component at the output being reduced. Continuing back around the loop, we see that the Sin^2 component is now reduced, which in turn slows the ramping of the multiplier. Eventually the steady-state of Fig. 10-5b is achieved. Here the integrator has ramped to some value V_c such that an amount of reference is passed through the lower multiplier which exactly cancels the sinusoidal component in the input. The output is now just the "speech" as shown. Following back around the loop, we see that this is multiplied by the reference, with the product still fed to the integrator. However, the speech and the reference are not correlated over any significant period of time, and in the product, the positive and negative portions pretty much average to zero. The integrator output may be fluctuating slightly, but as long as the integrator time constant is long enough, the output stands still at V_c for all practical purposes.

Note that the cancellation is now locked in by a negative feedback mechanism of the type we have seen many times before. If for example, the output of the integrator fluctuates up, then the sinusoidal at the input is over-cancelled, and a small negative sinusoidal appears at the output. This in turn leads to a negative Sin^2 term at the integrator input, which causes the integrator to ramp back down. A similar argument of course applies to a negative fluctuation of the integrator's output.

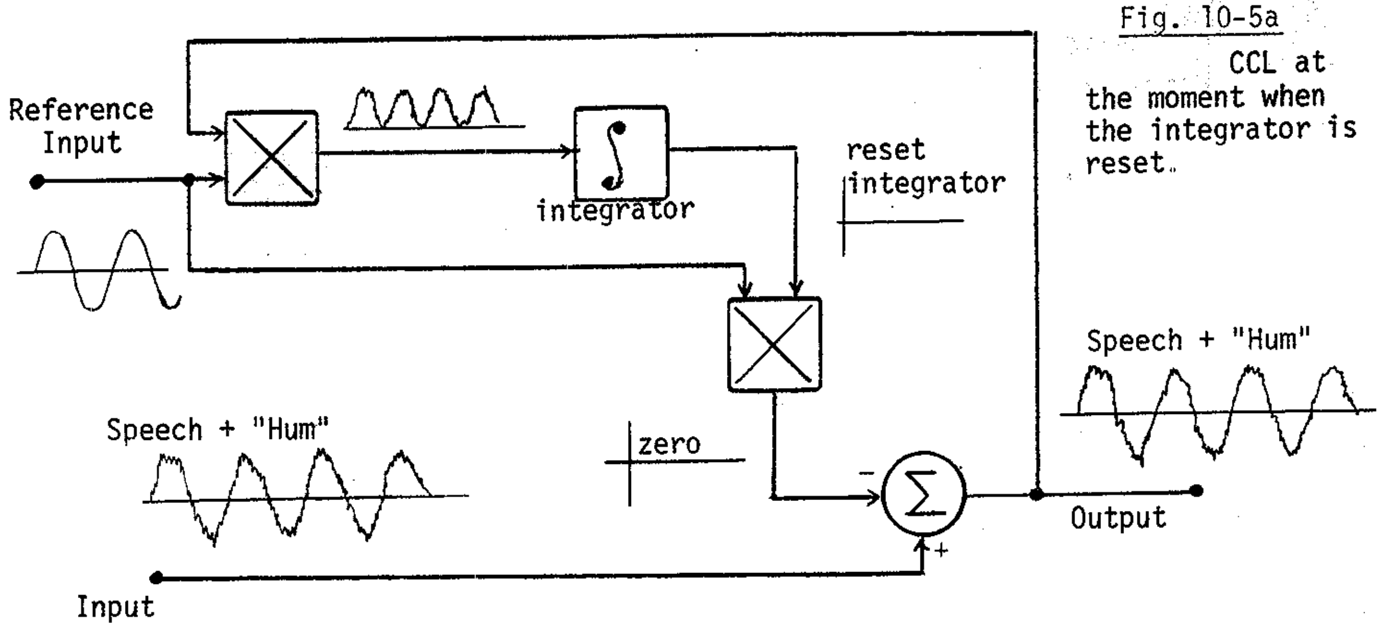


Fig. 10-5a

CCL at the moment when the integrator is reset.

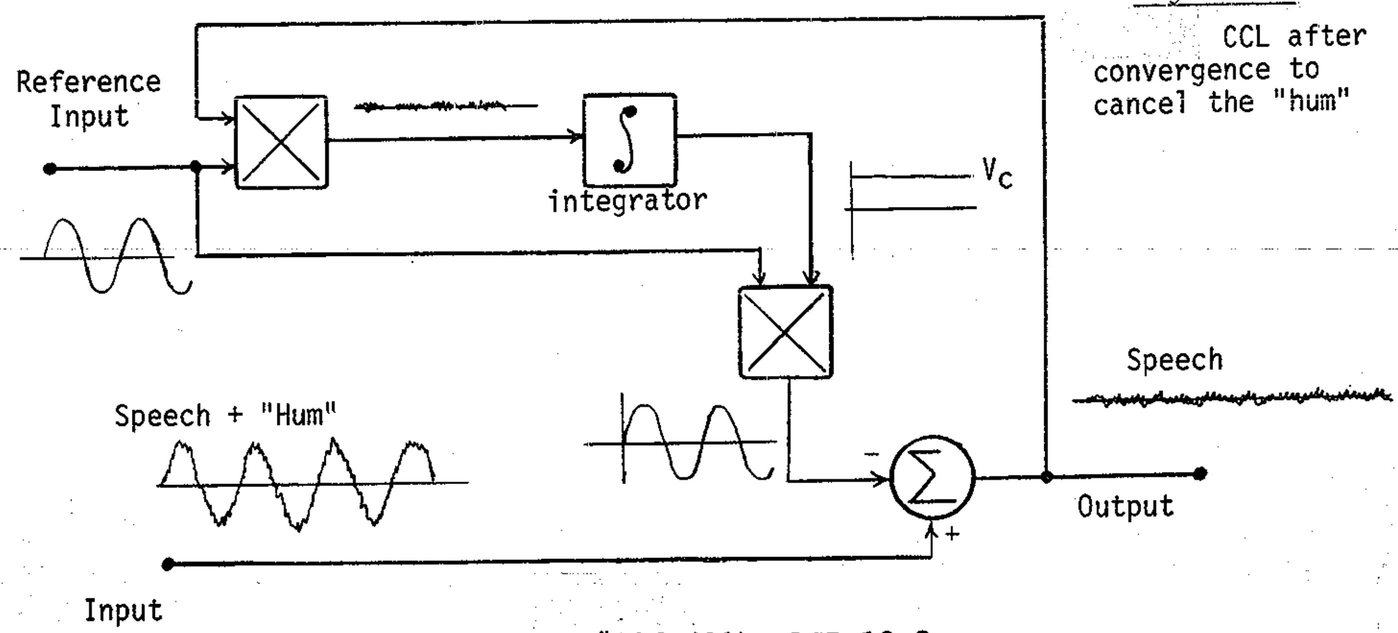


Fig. 10-5b

CCL after convergence to cancel the "hum"

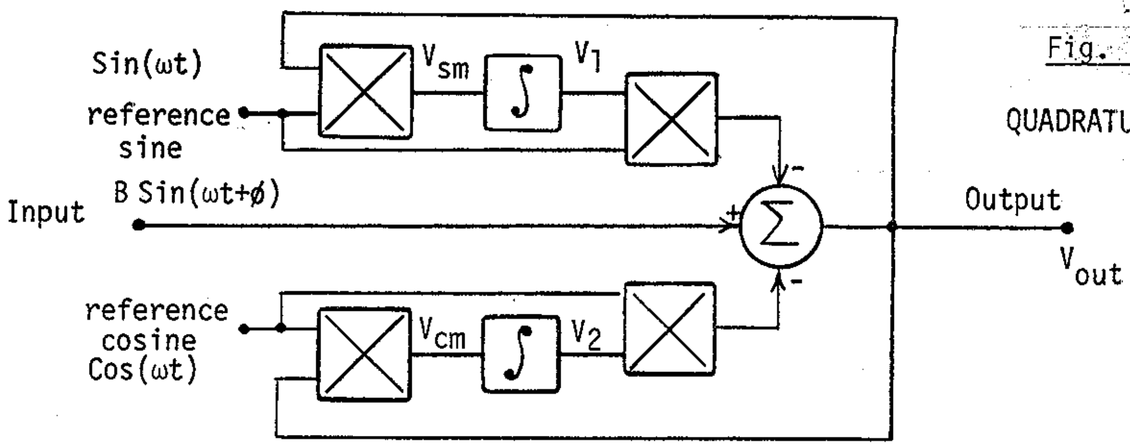


Fig. 10-6

QUADRATURE CCL's

Fig. 10-6 shows a case where we now have two CCL's instead of just the one. This is significant first because it will show us how to handle the case of an arbitrary phase shift between the input sinusoidal and the reference sinusoidal, and because it is the first step in going from the CCL to a full adaptive filter. Note that the summer of the CCL's is now a common summer, which we can compare to Fig. 10-2, where the common summer is in cascade with the subtracting summer, for an equivalent result.

Fig. 10-6 shows that we have available a Sine and a Cosine of the frequency we wish to cancel. In general, to cancel any given sinusoidal component we would expect to have to achieve a particular amplitude and a particular phase. However, we can always do this by a linear combination of a Sine component and a Cosine component, as is seen in Fig. 10-7a, and it can also be seen (Fig. 10-7b) that we do not need to have exactly Sine and Cosine components. In fact, any components with relative phases of say 70° to 110° would probably work quite well. Theoretically, any two different phases, even 0° and 1° would work, but this can put a very severe requirements on the amplitude range of the multipliers, as is shown in Fig. 10-7c. Accordingly we would prefer to have a pair of components involved that have a phase difference of something close to 90°, but we must keep in mind that this is a matter of practical convenience, and not required by theory.

The convergence solution for the quadrature CCL can be demonstrated as follows, with the single CCL being a special case with only a Sine reference. According to Fig. 10-6, the output is:

$$V_{out} = B\,\mathrm{Sin}(\omega t+\phi) - V_1\,\mathrm{Sin}(\omega t) - V_2\,\mathrm{Cos}(\omega t) \qquad (10\text{-}1)$$

At the upper left multiplier, we have this output multiplied by Sin(ωt) which gives:

$$V_{sm} = B\,\mathrm{Sin}(\omega t+\phi)\mathrm{Sin}(\omega t) - V_1\,\mathrm{Sin}^2(\omega t) - V_2\,\mathrm{Cos}(\omega t)\mathrm{Sin}(\omega t) \qquad (10\text{-}2)$$

$$= (B/2)\mathrm{Cos}(\phi) - (B/2)\mathrm{Cos}(2\omega t-\phi) - V_1/2 - (V_1/2)\mathrm{Cos}(2\omega t) - (V_2/2)\mathrm{Sin}(2\omega t) \qquad (10\text{-}3)$$

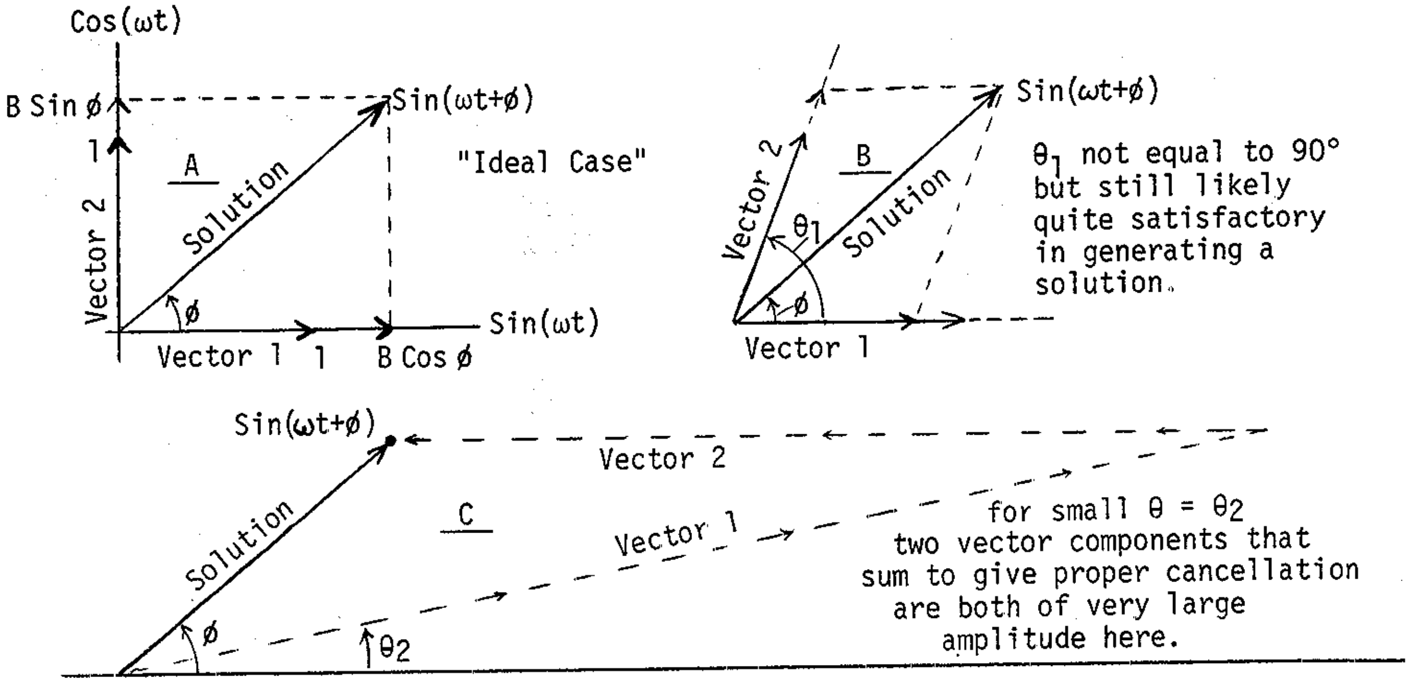


Fig. 10-7 Case A is ideal, using perfect 90° reference, equations (5) and (7). Same vector can be generated with non-90° angle in B and C.

This is the input to the integrator, and by assumption, the CCL has converged, and therefore V_1, the output of the integrator, is a constant. Therefore, any DC terms in the input, V_{sm} must vanish. This gives:

$$V_1 = B \, \mathrm{Cos}(\phi) \qquad (10\text{-}4)$$

In a similar manner, it can be shown that:

$$V_2 = B \, \mathrm{Sin}(\phi) \qquad (10\text{-}5)$$

From the discussion above and from Fig. 10-7 it can be seen that the CCL with two different phases, reasonably close to 90°, is capable of cancelling a sinusoidal component of arbitrary amplitude and phase. If the frequency of the interfering sinusoidal is fairly well known (as would be the case for power line hum, for example), a simple first-order all-pass (phase shifter) set for a 90° frequency at the nominal frequency, should be more than adequate. (Note however that even a broad band 90° network will not cancel two different frequencies with only the two tap case seen here - except in very special circumstances).

Fig 10-8 and Fig. 10-9 show some practical implementations of CCL's. Fig 10-8 is the single loop, corresponding to Fig. 10-5, while Fig. 10-9 is a double loop corresponding to the quadrature reference system of Fig. 10-6. Both circuits use the inexpensive transconductance multiplier along with an op-amp to form a four-quadrant multiplier. Parts cost for either circuit is very low, under $10 for Fig. 10-9.

In Fig. 10-8, the locations of the multipliers are indicated by their X and Y inputs and their Z outputs. The multiplier is actually configured as Z = -XY/5 as a matter of convenience here. The remainder of the circuit is standard analog circuitry, with the summer and

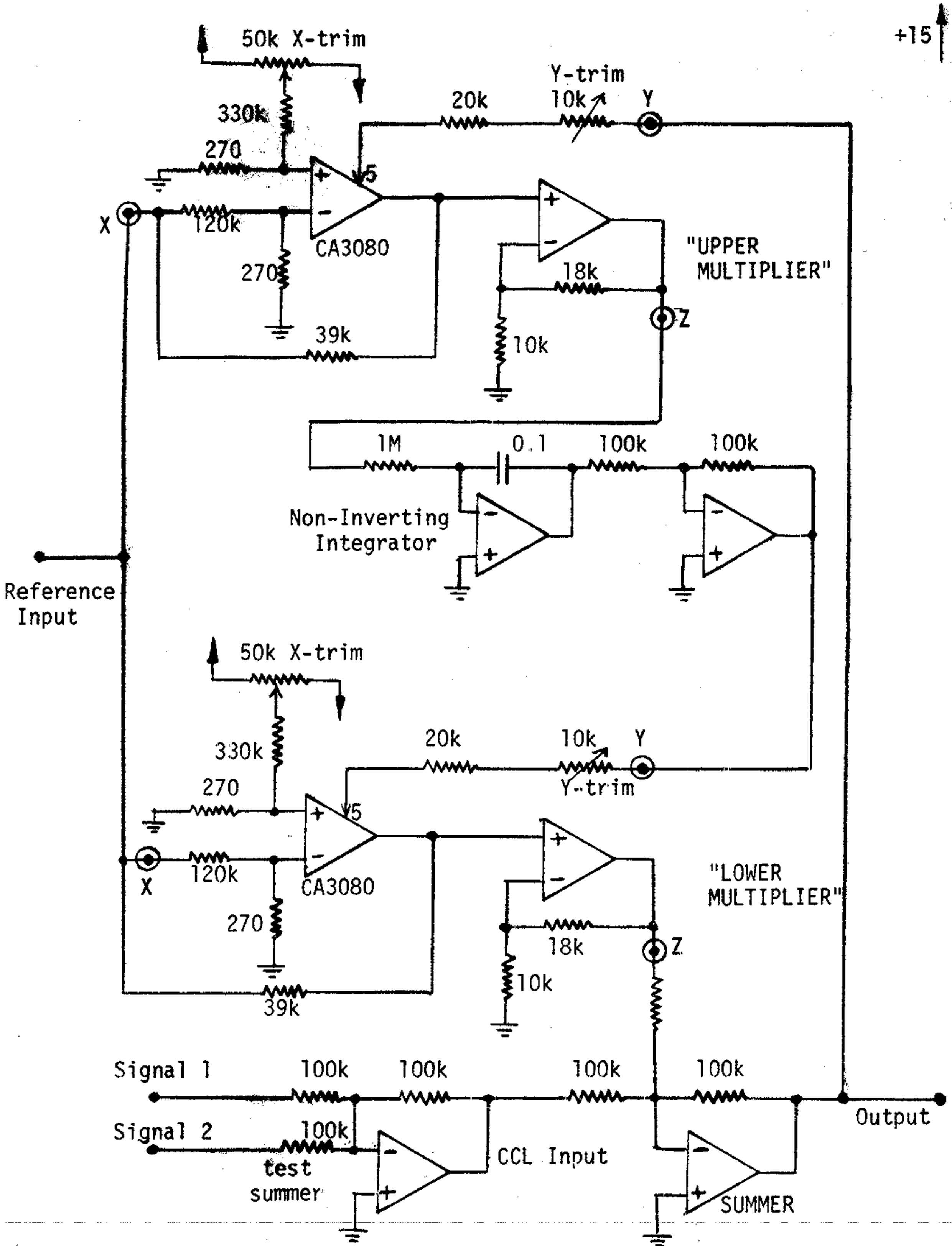


Fig. 10-8

A realization of an analog CCL using an inexpensive transcondictance multiplier (Type CA3080)

integrator as indicated. Just before the actual CCL summer, there is an additional "test summer" shown, the output of which is the actual CCL input. This permits a simple test and demonstration in which two different sinusoidal signals are summed by the test summer, becoming the CCL input, one of the signals being regarded as the "signal" while the other is regarded as the "noise." Whichever of these we wish to cancel is then connected to the reference input as well as to the test summer, and it will disappear from the output. The multipliers may be trimmed by conventional four-quadrant multiplier balancing techniques if desired, but it is usually more productive to adjust the trimmers to optimize the rejection of the unwanted component. In general, the

X-trim will have more effect on the performance than the Y-trim. Also, for demonstration purposes, it may be useful to greatly increase the time constant of the integrator (e.g., make the capacitor 1 microfarad instead of 0.1 microfarad). This makes the convergence time longer, of course, and the cancellation will take place gradually over several seconds (this also depends on the amplitudes of the signals, with faster convergence taking place for larger signals, as the integrator ramps faster). This slower convergence during demonstrations is more convincing as it seems to be psychologically more impressive to see something happen than it is to see that it has happened. Of course, in actual use, the CCL's time constant would be set on a performance basis.

Fig. 10-9 is basically an extension of Fig. 10-8. Here we could add a simple phase shifter to provide appropriate reference signals, and another phase shifter to provide the arbitrary phase shift ϕ between the input and the reference. This does work, and provides a demonstration similar to that of Fig. 10-8. In general, Fig. 10-8 will provide the

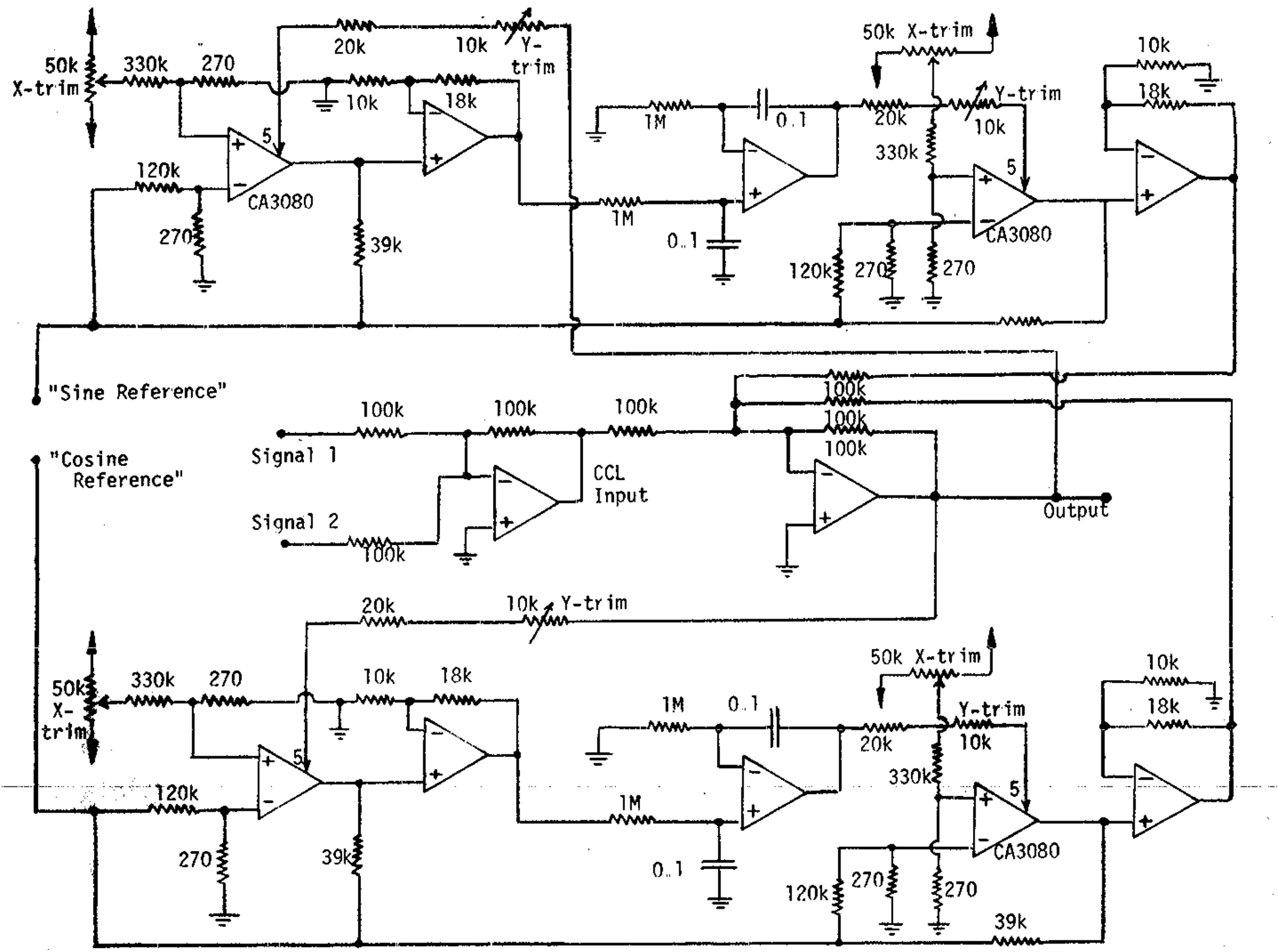


Fig. 10-9 Realization of a Quadrature CCL

most useful approach to practical problems of single component cancellation. In such a case, only one of Signal 1 or Signal 2 would be used, or we would go directly into the CCL input point, feeding in the signal to be cleaned up. The reference signal, and a phase shifter not shown, would then be used to provide the sine and cosine references indicated.

Another interesting demonstration that can be done with the double loop of Fig. 10-9 is to use one sinusoidal signal generator and appropriate 90° phase shifter for the reference, and to provide a second sinusoidal generator to the CCL input, and monitor the output for different input frequencies. That is, we test the CCL set up in this way exactly as we would a filter, measuring its frequency response. When the frequency of the input is exactly the same as the reference, we know that we only have an arbitrary phase difference ϕ, and the CCL should be able to cancel this, according to our discussion above. More interestingly, when the frequencies differ only slightly, we can still get substantial cancellation. This we can understand as the CCL system interpreting this small difference in frequency as a phase difference that it is continually trying to correct (and which it is capable of correcting). As this frequency difference gets larger, the CCL system is less capable of making up the apparent phase error, and cancellation is less complete. This is because of the integrator time constants that determine how fast the weights can change. The longer the time constants, the slower the correction, and the less complete the cancellation. Thus the system configured as described looks somewhat like a notch filter, with the notch position set by the reference oscillator frequency, and with the Q determined by the RC time constant of the integrator, getting higher for larger RC time constant. By the same argument, it is possible to see that the Q also depends on the inverse square of the reference amplitude, since the amplitude affects the integrator charging rate through two multipliers.

A final interesting point can be made about the single loop of Fig. 10-5. As long as we assume that the interference signal and the reference signal are exactly the same waveform, then the single loop is capable of cancelling arbitrary waveforms, and is not restricted to sinusoidals. This can be argued simply in a manner similar to that surrounding Fig. 10-5a and Fig. 10-5b. This fact can be very useful in cases where phase shift is very small, so that one loop can be used, or where there is a fixed time delay instead of an actual phase shift. The fixed time delay can be handled with an analog delay line, for example. The more general case, of a general waveform with arbitrary phase shifts in different components, requires a larger number of CCL's and a corresponding larger number of reference phases, typically obtained from a tapped delay line.

10-5 A COMPARISON OF THE CCL AND THE LMS ALGORITHM:

Fig. 10-10a shows a general view of an adaptive filter, with Fig. 10-10b showing an LMS algorithm realization, and Fig. 10-10c showing our familiar CCL. The theory of adaptive filtering, not discussed here, leads to the LMS algorithm equation, which is stated in the form of a tap weight update equation as:

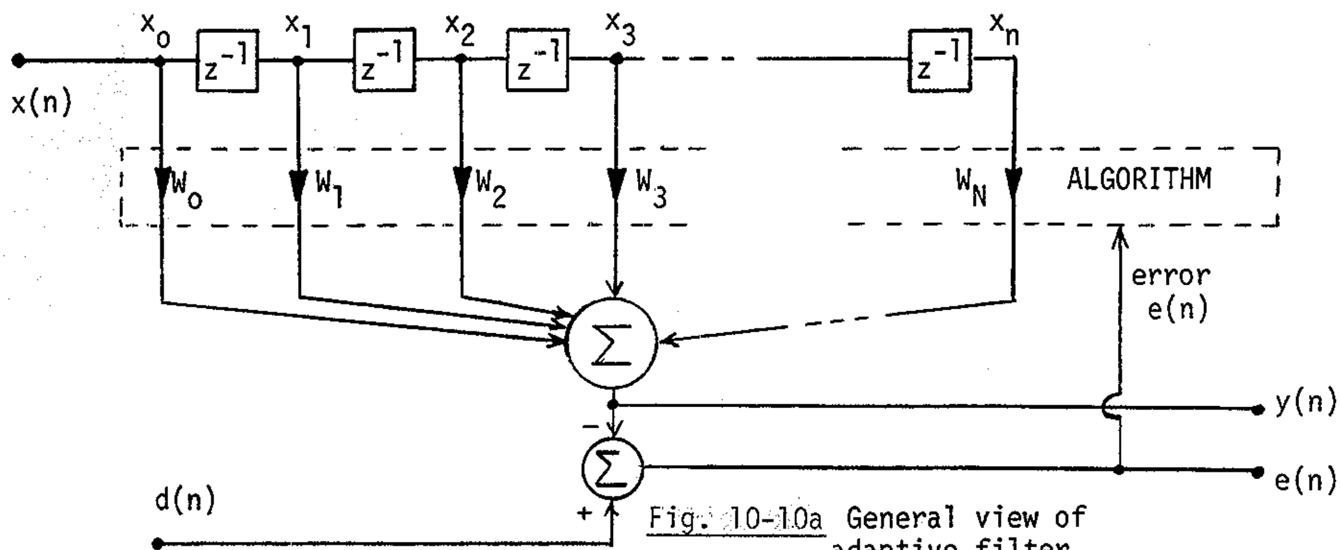


Fig. 10-10a General view of adaptive filter

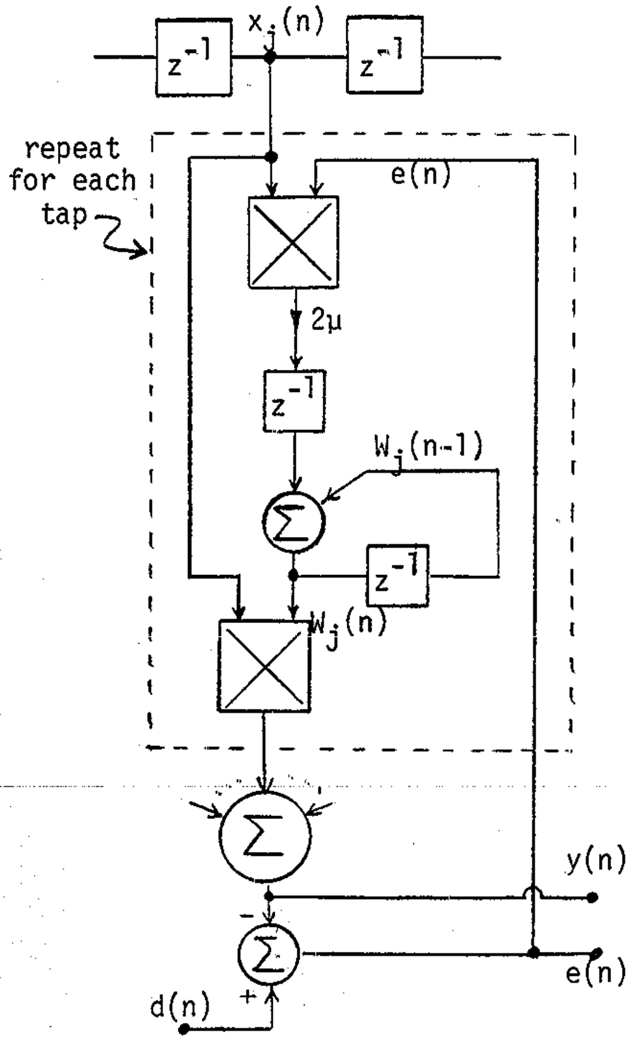


Fig. 10-10b A realization of the LMS algorithm

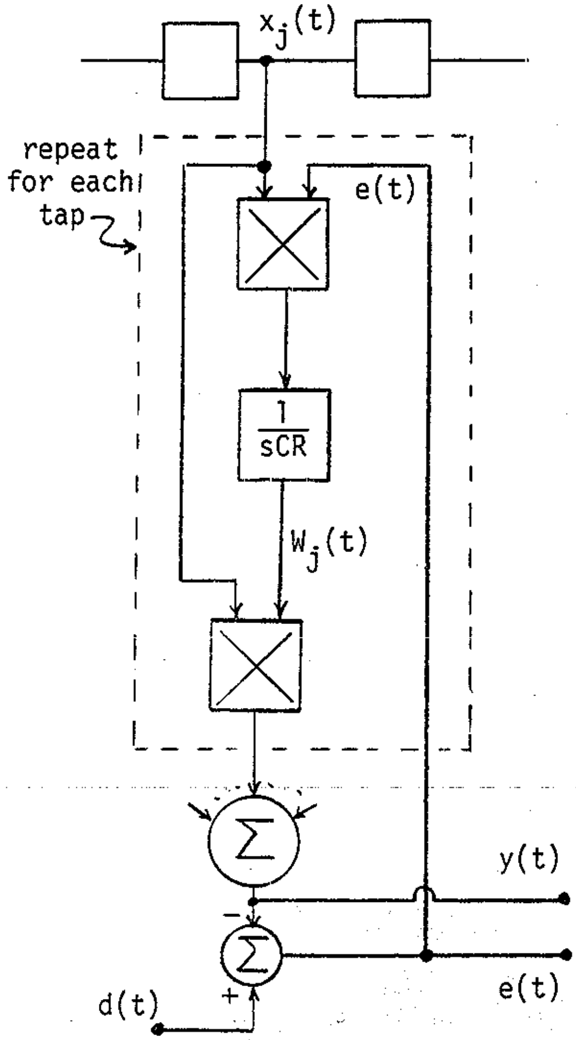


Fig. 10-10c Adjusting the tap weight with a CCL

$$W_j(n+1) = W_j(n) + 2\mu e(n)x_j(n) \qquad (10\text{-}6)$$

which says that the next tap weight, for time instant n+1, is the current tap weight, plus a correction term. Note that this change of tap weight is equivalent to a ramping of the integrator output in the CCL case. What is of interest is the correction term. Here, the parameter μ or 2μ is small (something like 0.001), so the correction term at any one time is small. The only way the tap can change substantially in one direction is for the product $e(n)x_j(n)$ to have the same sign over many time instances. This is to say that $e(n)$ and $x_j(n)$ must be correlated over a substantial amount of time, otherwise the tap weight is not being called upon to contribute to convergence, but is rather being just kicked up and down a bit. Of course, this reminds us of what is happening with the CCL. In fact, it is possible to show that the two systems can be related so that:

$$2\mu = T/RC \qquad (10\text{-}7)$$

ASP 10-15

END OF ANALOG SIGNAL PROCESSING TEXT

Tuning Equations Derived from Passive Sensitivity

(Continued from Page 2)

But, suppose we need to do somewhat better, or perhaps we take great pride in our work and want this particular set (my set) of ten filters to be better tuned. What sort of things can we do that may work well and not cost us too much in money or time?

The first thing is obviously to consider using variable resistors ("trim pots") to adjust about a certain range. Keep in mind that a good quality trim pot costs several dollars, a cheap one perhaps 30 cents, and a fixed resistor only 3 cents. So good quality trimmers may be out of the question because of cost, and lesser quality ones may be subject to drift and contamination. Another problem with trimmers is that there is always a temptation to adjust them one more time, or to worry if perhaps we accidentally jarred one. All these considerations make a suggestion of using a just-one-added-resistor-soldered-in tune-up quite attractive. We just need to be able to determine what resistor to use without undue trial and error. Let's look at an example.

Suppose our cutoff frequency is determined by four passive components as:

$$f_c = 1/2\pi(R_1R_2C_1C_2)^{1/2} \qquad (1)$$

as is typical of many configurations (e.g., Sallen-Key). We can easily calculate:

$$S^{f_c}_{R_1} = (R_1/f_c)\ \partial f_c/\partial R_1 = -1/2 \qquad (2)$$

This is a particularly simple case, and makes a good example. By replacing the partial derivatives with deltas, we obtain:

$$\Delta R_1 = -2\ (R_1/f_c)\Delta f_c \qquad (3)$$

Now suppose that we have intentionally set R_1 slightly lower than nominal, so we expect the frequency to be initially slightly high. At the same time, we arrange on our circuit boards to make R_1 as a series combination of two resistors, the second of which is initially just a wire. Now we measure the cutoff frequency, and we find it to be, for example, 1032 Hz. That is, Δf_c = +32 Hz. Suppose the initial value of R_1 is nominally 10,000 (we don't know for sure what it is unless we measure it, but expect it to be within advertised tolerances). In a sense, by measuring the cutoff, we have made an overall measurement of all four frequency-determining components. (Any and all of them contribute to the error in general.) But, <u>if</u> the error were totally due to R_1, what error in R_1 would account for the observed frequency error? The answer is:

$$\Delta R_1 = -2(10000/1000)\ 32 = -640 \text{ ohms} \qquad (4)$$

This means that the frequency error observed is as though R_1 were wrong by -640 ohms. This means that R_1 is 640 ohms too small. Thus we would clip out the zero ohm wire and install a 640 ohm correction. Note that in equation (4) we have plugged in nominal values for R_1 and for f_c, but things would be little changed if we tried other possibly better values. The level of correction is at the 6% level of the component (640 ohms in 10000 ohms). If we had used the measured value of 1032 for f_c, the correction resistor would be 620 ohms. In either case here (640 or 620 ohms) we would have chosen the closest nominal 5% resistor, which would have been 620 ohms. Not unreasonably, we expect this correction to put us close to 5% of 6% or 0.3%, and the frequency to within half this error because of the sensitivity having a magnitude of 1/2.

This can usually give astoundingly good results. Note that it involves the use of two resistors, where a large one is effectively measured, and a small correction then chosen by formula, but not measured. Obviously, the process is easily iterated if we like. But after all, we are not using much more than the fact that a -3% frequency correction required a +6% change in a frequency determining resistor, as indicated by the sensitivity value of -1/2.

This first example was fairly straightforward. We knew how to calculate the sensitivity and interpret the result as a tuning formula. More complicated cases may be as easy only in theory. Consider as a second example the popular "Deliyannis" bandpass filter. [See details in Application Note 145, "Analysis of the Deliyannis Filter," Sept. 4, 1979. This filter also appears as Fig. 5-8 of Chapter 5 of Analog Signal Processing, pg 8 of EN#194.] This filter has a center frequency given by

$$f_0 = 1/2\pi\sqrt{B}RC \tag{5}$$

where B is a ratio of two resistors. This is similar to the first example. The filter's Q is given by:

$$Q = (1-a)\sqrt{B} \ / \ [2(1-a) - aB] \tag{6}$$

and the gain at the center frequency (the "peak gain") is given by:

$$g = B \ / \ [\ 2(1-a) - aB] \tag{7}$$

Here the ratio a is the fraction of the output fed back to the (+) input of the op-amp, and is just the result of another resistor ratio.

It is evident that the manipulation of the center frequency is dependent on B but not on a. So we might consider adjusting the frequency first. Further, it is difficult to accurately measure the Q, since the Deliyannis filter is generally used for its ability to achieve very high Q's. So we may well opt for adjusting the peak gain, with the idea that when we get this right, the Q may well come along. Equations (6) and (7) suggest this. (Or, we may well be mainly concerned with achieving the correct peak gain.) Thus, we suppose B is fixed and we need to adjust a. What is the sensitivity of g to a? Well, we could do this analytically by taking the usual partial derivative:

$$S_a^g = (a/g)\ \partial g/\partial a = a(2+B) \ / \ [2 - a(2+B)] \tag{8}$$

but we can also cheat. We really need $\Delta g/\Delta a$ for our tuning. All we really need to do is put in our nominal value of B, wiggle the ratio a ever so slightly, and see how much g changes. This we do not by a circuit measurement, but simply by using equation (7).

Note that the ratio a, for a positive, non-infinite Q, must be less than 2/(B-2). So if B=16 for example, a would need to be less than 0.1429. And, we would already have a nominal value of a as part of our design calculations. Perhaps, for example, a was supposed to be 0.1. (Overall, this means we designed for Q=18.) So what is the sensitivity of g to a about the nominal design? Well, suppose we try two values of a: a=0.1 and a=0.1001. We get g values of 80 and 80.726. So Δa is 0.0001 and Δg is 0.726. The sensitivity of g to a is right around 9.075. This is not good from the point of view of getting it right without tuning, but not all that bad when we consider that we need for g to depend on a or we would have no way of doing the tuning. Note that while we got this without taking derivatives, plugging a=0.1 and B=16 into equation (7)(from taking derivatives) does give us exactly 9.

What does this really mean? Suppose g needs be 80 but is only 50, as could well occur. Then a would have to be changed by (a/g)(Δg/9) or about 0.004. Well, a would be determined by a ratio of resistors and the individual resistors would typically be on the order to 10k to 100k. So making an adjustment of a on the order of 0.4% would involve an additional resistor added to the lower leg of the divider on the order of 40 to 400 ohms. This is something we can handle. It is worth noting that if we actually do solder in such an adjustment, that while the soldering heat does not change the small additional resistor much, the heat can often wonder into the other resistor of the divider and throw off the results for perhaps 10 seconds before the resistors all cool down. It can be amusing to watch the performance converge - as we encourage it by blowing on it. Because of our calculations, the results can come around quite well, and with little or no additional trial-and-error.

Finally, suppose that in a particular case we want a=0.1 and B=16. We choose to use two standard 5% resistors, 10k and 91k, to achieve the ratio a. Suppose the actual values we end up with are 10.1k and 92.1k, so that a=10.1/(92.1+10.1)=0.0988. We do not know these actual values, but according to equations (6) and (7) we expect to measure Q=16.3 and g=72.4 (not g=80). The actual sensitivity, using equation (8), is S=8.0, but we don't know this, and use the nominal S=9. Using equation (8) we get:

$$\Delta a = (a/g)(\Delta g/9) = (0.1/72.4)(-7.6/9) = -0.001166 \qquad (9)$$

If we make this correction a = 0.0988+0.001166 = 0.09999. Using this new value of a, we calculate that we would observe Q=17.99 and g=79.9, very close to nominal. This we would achieve with a series correction to the 10k resistor. This value, R_x, would be $R_x/\Delta a$ = 10k/0.1 or right around 117 ohms. Thus, from our easily measured Δg we calculate Δa based on <u>observed</u> Δg, g, and <u>nominal</u> values for a and S. From Δa we calculate R_x based on nominal values of a and one of the resistors. Note that Q as well as g does come up near normal as a result.

Electronotes, Vol. 20, No. 196, December 2000
Published by B. Hutchins, 1016 Hanshaw Rd., Ithaca, NY 14850
(607)-257-8010