A Basic Introduction to Filters—Active, Passive, and Switched-Capacitor

National Semiconductor
Application Note 779
Kerry Lacanette

1.0 INTRODUCTION

Filters of some sort are essential to the operation of most electronic circuits. It is therefore in the interest of anyone involved in electronic circuit design to have the ability to develop filter circuits capable of meeting a given set of specifications. Unfortunately, many in the electronics field are uncomfortable with the subject, whether due to a lack of familiarity with it, or a reluctance to grapple with the mathematics involved in a complex filter design.

This Application Note is intended to serve as a very basic introduction to some of the fundamental concepts and terms associated with filters. It will not turn a novice into a filter designer, but it can serve as a starting point for those wishing to learn more about filter design.

1.1 Filters and Signals: What Does a Filter Do?

In circuit theory, a filter is an electrical network that alters the amplitude and/or phase characteristics of a signal with respect to frequency. Ideally, a filter will not add new frequencies to the input signal, nor will it change the component frequencies of that signal, but it will change the relative amplitudes of the various frequency components and/or their phase relationships. Filters are often used in electronic systems to emphasize signals in certain frequency ranges and reject signals in other frequency ranges. Such a filter has a **gain** which is dependent on signal frequency. As an example, consider a situation where a useful signal at frequency f_1 has been contaminated with an unwanted signal at f_2. If the contaminated signal is passed through a circuit *(Figure 1)* that has very low gain at f_2 compared to f_1, the undesired signal can be removed, and the useful signal will remain. Note that in the case of this simple example, we are not concerned with the gain of the filter at any frequency other than f_1 and f_2. As long as f_2 is sufficiently attenuated relative to f_1, the performance of this filter will be satisfactory. In general, however, a filter's gain may be specified at several different frequencies, or over a band of frequencies.

Since filters are defined by their frequency-domain effects on signals, it makes sense that the most useful analytical and graphical descriptions of filters also fall into the frequency domain. Thus, curves of gain vs frequency and phase vs frequency are commonly used to illustrate filter characteristics,and the most widely-used mathematical tools are based in the frequency domain.

The frequency-domain behavior of a filter is described mathematically in terms of its **transfer function** or **network function**. This is the ratio of the Laplace transforms of its output and input signals. The voltage transfer function H(s) of a filter can therefore be written as:

$$H(s) = \frac{V_{OUT}(s)}{V_{IN}(s)} \qquad (1)$$

where $V_{IN}(s)$ and $V_{OUT}(s)$ are the input and output signal voltages and s is the complex frequency variable.

The transfer function defines the filter's response to any arbitrary input signal, but we are most often concerned with its effect on continuous sine waves. Especially important is the magnitude of the transfer function as a function of frequency, which indicates the effect of the filter on the amplitudes of sinusoidal signals at various frequencies. Knowing the transfer function magnitude (or gain) at each frequency allows us to determine how well the filter can distinguish between signals at different frequencies. The transfer function magnitude versus frequency is called the **amplitude response** or sometimes, especially in audio applications, the **frequency response**.

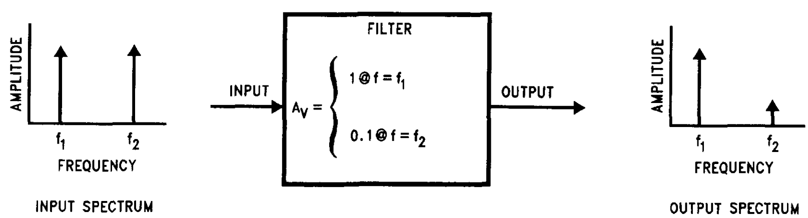
Similarly, the **phase response** of the filter gives the amount of **phase shift** introduced in sinusoidal signals as a function of frequency. Since a change in phase of a signal also represents a change in time, the phase characteristics of a filter become especially important when dealing with complex signals where the time relationships between signal components at different frequencies are critical.

By replacing the variable s in (1) with $j\omega$, where j is equal to $\sqrt{-1}$, and ω is the radian frequency ($2\pi f$), we can find the filter's effect on the magnitude and phase of the input signal. The magnitude is found by taking the absolute value of (1):

$$|H(j\omega)| = \left|\frac{V_{OUT}(j\omega)}{V_{IN}(j\omega)}\right| \qquad (2)$$

and the phase is:

$$\arg H(j\omega) = \arg \frac{V_{OUT}(j\omega)}{V_{IN}(j\omega)} \qquad (3)$$

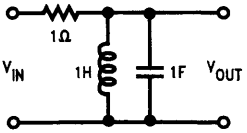


TL/H/11221–1

FIGURE 1. Using a Filter to Reduce the Effect of an Undesired Signal at Frequency f_2, while Retaining Desired Signal at Frequency f_1

As an example, the network of *Figure 2* has the transfer function:

$$H(s) = \frac{s}{s^2 + s + 1} \quad (4)$$



TL/H/11221-2

FIGURE 2. Filter Network of Example

This is a 2nd order system. The **order** of a filter is the highest power of the variable s in its transfer function. The order of a filter is usually equal to the total number of capacitors and inductors in the circuit. (A capacitor built by combining two or more individual capacitors is still one capacitor.) Higher-order filters will obviously be more expensive to build, since they use more components, and they will also be more complicated to design. However, higher-order filters can more effectively discriminate between signals at different frequencies.

Before actually calculating the amplitude response of the network, we can see that at very low frequencies (small values of s), the numerator becomes very small, as do the first two terms of the denominator. Thus, as s approaches zero, the numerator approaches zero, the denominator approaches one, and H(s) approaches zero. Similarly, as the input frequency approaches infinity, H(s) also becomes progressively smaller, because the denominator increases with the square of frequency while the numerator increases linearly with frequency. Therefore, H(s) will have its maximum value at some frequency between zero and infinity, and will decrease at frequencies above and below the peak.

To find the magnitude of the transfer function, replace s with jω to yield:

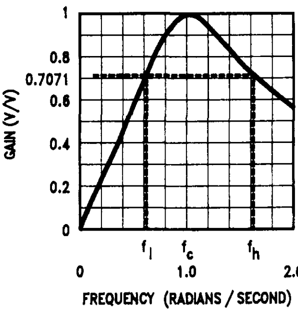
$$A(\omega) = |H(s)| = \left|\frac{j\omega}{-\omega^2 + j\omega + 1}\right| = \frac{\omega}{\sqrt{\omega^2 + (1 - \omega^2)^2}} \quad (5)$$

The phase is:

$$\theta(\omega) = \arg H(s) = 90^\circ - \tan^{-1}\frac{\omega^2}{(1 - \omega^2)} \quad (6)$$

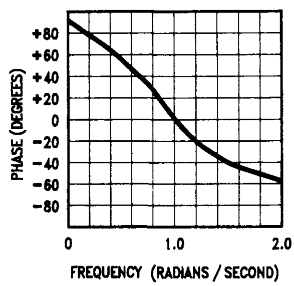
The above relations are expressed in terms of the radian frequency ω, in units of radians/second. A sinusoid will complete one full cycle in 2π radians. Plots of magnitude and phase versus radian frequency are shown in *Figure 3*. When we are more interested in knowing the amplitude and phase response of a filter in units of Hz (cycles per second), we convert from radian frequency using $\omega = 2\pi f$, where f is the frequency in Hz. The variables f and ω are used more or less interchangeably, depending upon which is more appropriate or convenient for a given situation.

Figure 3(a) shows that, as we predicted, the magnitude of the transfer function has a maximum value at a specific frequency (ω_0) between 0 and infinity, and falls off on either side of that frequency. A filter with this general shape is known as a **band-pass** filter because it passes signals falling within a relatively narrow band of frequencies and attenuates signals outside of that band. The range of frequencies passed by a filter is known as the filter's **passband**. Since the amplitude response curve of this filter is fairly smooth, there are no obvious boundaries for the passband. Often, the passband limits will be defined by system requirements. A system may require, for example, that the gain variation between 400 Hz and 1.5 kHz be less than 1 dB. This specification would effectively define the passband as 400 Hz to 1.5 kHz. In other cases though, we may be presented with a transfer function with no passband limits specified. In this case, and in any other case with no explicit passband limits, the passband limits are usually assumed to be the frequencies where the gain has dropped by 3 decibels (to $\sqrt{2}/2$ or 0.707 of its maximum voltage gain). These frequencies are therefore called the **−3 dB frequencies** or the **cutoff frequencies**. However, if a passband gain variation (i.e., 1 dB) is specified, the cutoff frequencies will be the frequencies at which the maximum gain variation specification is exceeded.



TL/H/11221-3

(a)



TL/H/11221-5

(b)

FIGURE 3. Amplitude (a) and phase (b) response curves for example filter. Linear frequency and gain scales.

The precise shape of a band-pass filter's amplitude response curve will depend on the particular network, but any 2nd order band-pass response will have a peak value at the filter's **center frequency**. The center frequency is equal to the geometric mean of the −3 dB frequencies:

$$f_c = \sqrt{f_l\, f_h} \quad (8)$$

where f_c is the center frequency

f_l is the lower −3 dB frequency

f_h is the higher −3 dB frequency

Another quantity used to describe the performance of a filter is the filter's "**Q**". This is a measure of the "sharpness" of the amplitude response. The Q of a band-pass filter is the ratio of the center frequency to the difference between the

−3 dB frequencies (also known as the **−3 dB bandwidth**). Therefore:

$$Q = \frac{f_c}{f_h - f_l} \quad (9)$$

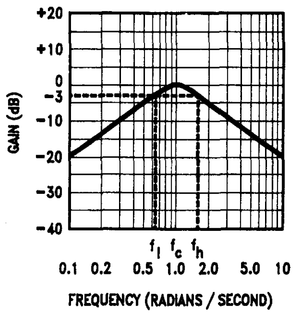
When evaluating the performance of a filter, we are usually interested in its performance over **ratios** of frequencies. Thus we might want to know how much attenuation occurs at twice the center frequency and at half the center frequency. (In the case of the 2nd-order bandpass above, the attenuation would be the same at both points). It is also usually desirable to have amplitude and phase response curves that cover a wide range of frequencies. It is difficult to obtain a useful response curve with a linear frequency scale if the desire is to observe gain and phase over wide frequency ratios. For example, if $f_0 = 1$ kHz, and we wish to look at response to 10 kHz, the amplitude response peak will be close to the left-hand side of the frequency scale. Thus, it would be very difficult to observe the gain at 100 Hz, since this would represent only 1% of the frequency axis. A logarithmic frequency scale is very useful in such cases, as it gives equal weight to equal **ratios** of frequencies.

Since the range of amplitudes may also be large, the amplitude scale is usually expressed in decibels ($20\log|H(j\omega)|$). *Figure 4* shows the curves of *Figure 3* with logarithmic frequency scales and a decibel amplitude scale. Note the improved symmetry in the curves of *Figure 4* relative to those of *Figure 3*.

1.2 The Basic Filter Types

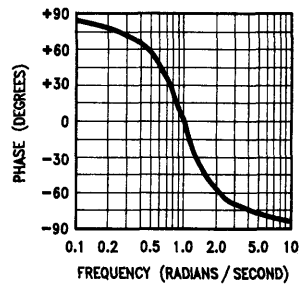
Bandpass

There are five basic filter types (bandpass, notch, low-pass, high-pass, and all-pass). The filter used in the example in the previous section was a bandpass. The number of possible bandpass response characteristics is infinite, but they all share the same basic form. Several examples of bandpass amplitude response curves are shown in *Figure 5*. The curve in *5(a)* is what might be called an "ideal" bandpass response, with absolutely constant gain within the passband, zero gain outside the passband, and an abrupt boundary between the two. This response characteristic is impossible to realize in practice, but it can be approximated to varying degrees of accuracy by real filters. Curves (b) through (f) are examples of a few bandpass amplitude response curves that approximate the ideal curves with varying degrees of accuracy. Note that while some bandpass responses are very smooth, other have **ripple** (gain variations in their passbands. Other have ripple in their **stopbands** as well. The stopband is the range of frequencies over which unwanted signals are attenuated. Bandpass filters have two stopbands, one above and one below the passband.



(a)

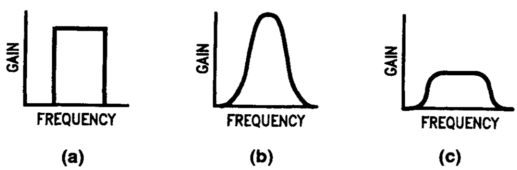
TL/H/11221-4



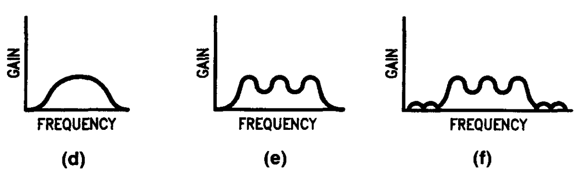
(b)

TL/H/11221-6

FIGURE 4. Amplitude (a) and phase (b) response curves for example bandpass filter. Note symmetry of curves with log frequency and gain scales.



TL/H/11221-7



TL/H/11221-8

FIGURE 5. Examples of Bandpass Filter Amplitude Response

Just as it is difficult to determine by observation exactly where the passband ends, the boundary of the stopband is also seldom obvious. Consequently, the frequency at which a stopband begins is usually defined by the requirements of a given system—for example, a system specification might require that the signal must be attenuated at least 35 dB at 1.5 kHz. This would define the beginning of a stopband at 1.5 kHz.

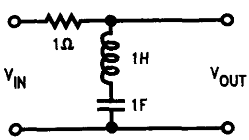
The rate of change of attenuation between the passband and the stopband also differs from one filter to the next. The slope of the curve in this region depends strongly on the order of the filter, with higher-order filters having steeper cutoff slopes. The attenuation slope is usually expressed in dB/octave (an octave is a factor of 2 in frequency) or dB/decade (a decade is a factor of 10 in frequency).

Bandpass filters are used in electronic systems to separate a signal at one frequency or within a band of frequencies from signals at other frequencies. In 1.1 an example was given of a filter whose purpose was to pass a desired signal at frequency f_1, while attenuating as much as possible an unwanted signal at frequency f_2. This function could be performed by an appropriate bandpass filter with center frequency f_1. Such a filter could also reject unwanted signals at other frequencies outside of the passband, so it could be useful in situations where the signal of interest has been contaminated by signals at a number of different frequencies.

Notch or Band-Reject

A filter with effectively the opposite function of the bandpass is the **band-reject** or **notch** filter. As an example, the components in the network of *Figure 3* can be rearranged to form the notch filter of *Figure 6*, which has the transfer function

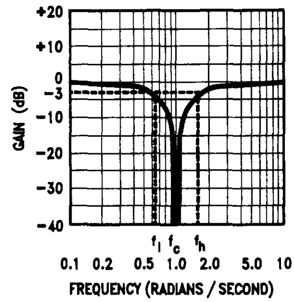
$$H_N(s) = \frac{V_{OUT}}{V_{IN}} = \frac{s^2 + 1}{s^2 + s + 1} \qquad (10)$$



TL/H/11221-9

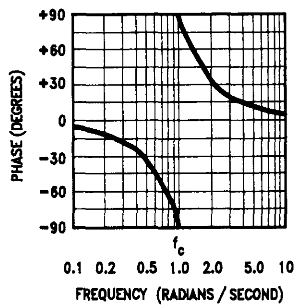
FIGURE 6. Example of a Simple Notch Filter

The amplitude and phase curves for this circuit are shown in *Figure 7*. As can be seen from the curves, the quantities f_c, f_l, and f_h used to describe the behavior of the band-pass filter are also appropriate for the notch filter. A number of notch filter amplitude response curves are shown in *Figure 8*. As in *Figure 5*, curve (a) shows an "ideal" notch response, while the other curves show various approximations to the ideal characteristic.



TL/H/11221-10

(a)



TL/H/11221-11

(b)

FIGURE 7. Amplitude (a) and Phase (b) Response Curves for Example Notch Filter

Notch filters are used to remove an unwanted frequency from a signal, while affecting all other frequencies as little as possible. An example of the use of a notch flter is with an audio program that has been contaminated by 60 Hz power-line hum. A notch filter with a center frequency of 60 Hz can remove the hum while having little effect on the audio signals.

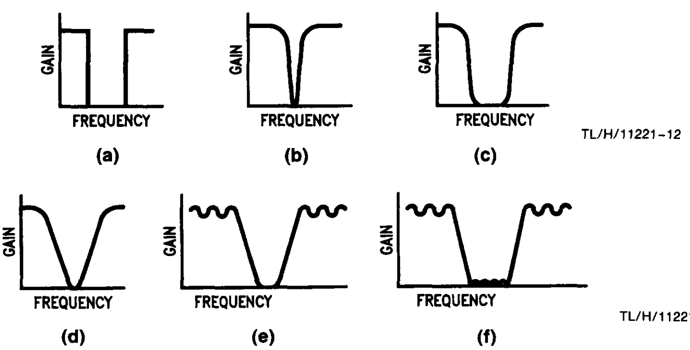
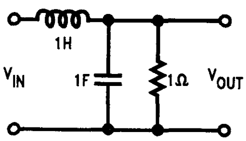


FIGURE 8. Examples of Notch Filter Amplitude Responses

Low-Pass

A third filter type is the **low-pass**. A low-pass filter passes low frequency signals, and rejects signals at frequencies above the filter's cutoff frequency. If the components of our example circuit are rearranged as in *Figure 9*, the resultant transfer function is:

$$H_{LP}(s) = \frac{V_{OUT}}{V_{IN}} = \frac{1}{s^2 + s + 1} \tag{11}$$



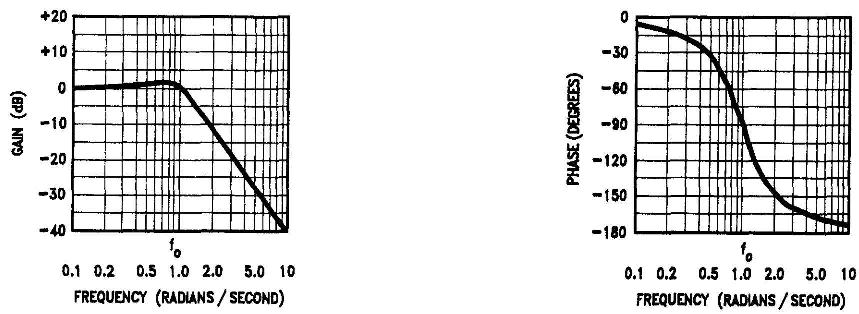
TL/H/11221-14

FIGURE 9. Example of a Simple Low-Pass Filter

It is easy to see by inspection that this transfer function has more gain at low frequencies than at high frequencies. As ω approaches 0, H_{LP} approaches 1; as ω approaches infinity, H_{LP} approaches 0.

Amplitude and phase response curves are shown in *Figure 10*, with an assortment of possible amplitude reponse curves in *Figure 11*. Note that the various approximations to the unrealizable ideal low-pass amplitude characteristics take different forms, some being monotonic (always having a negative slope), and others having ripple in the passband and/or stopband.

Low-pass filters are used whenever high frequency components must be removed from a signal. An example might be in a light-sensing instrument using a photodiode. If light levels are low, the output of the photodiode could be very small, allowing it to be partially obscured by the noise of the sensor and its amplifier, whose spectrum can extend to very high frequencies. If a low-pass filter is placed at the output of the amplifier, and if its cutoff frequency is high enough to allow the desired signal frequencies to pass, the overall noise level can be reduced.



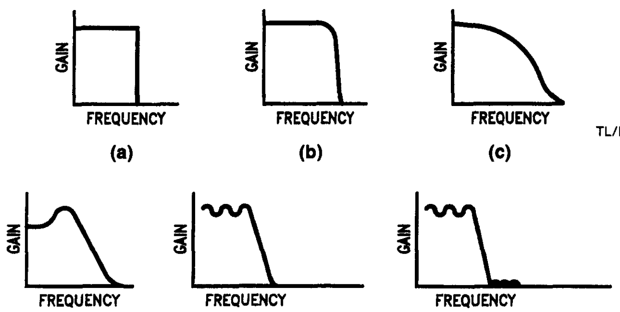
TL/H/11221-15

(a)

TL/H/11221-16

(b)

FIGURE 10. Amplitude (a) and Phase (b) Response Curves for Example Low-Pass Filter



(a) (b) (c)

TL/H/11221-17

(d) (e) (f)

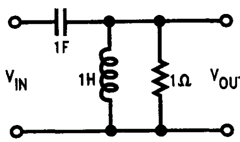
TL/H/11221-18

FIGURE 11. Examples of Low-Pass Filter Amplitude Response Curves

High-Pass

The opposite of the low-pass is the **high-pass** filter, which rejects signals below its cutoff frequency. A high-pass filter can be made by rearranging the components of our example network as in *Figure 12*. The transfer function for this filter is:

$$H_{HP}(s) = \frac{V_{OUT}}{V_{IN}} = \frac{s^2}{s^2 + s + 1} \qquad (12)$$

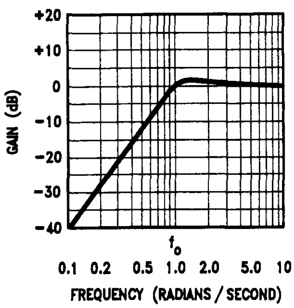


TL/H/11221-19

FIGURE 12. Example of Simple High-Pass Filter

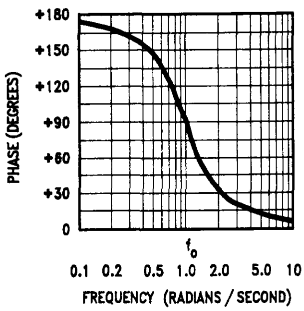
and the amplitude and phase curves are found in *Figure 13*. Note that the amplitude response of the high-pass is a "mirror image" of the low-pass response. Further examples of high-pass filter responses are shown in *Figure 14*, with the "ideal" response in (a) and various approximations to the ideal shown in (b) through (f).

High-pass filters are used in applications requiring the rejection of low-frequency signals. One such application is in high-fidelity loudspeaker systems. Music contains significant energy in the frequency range from around 100 Hz to 2 kHz, but high-frequency drivers (tweeters) can be damaged if low-frequency audio signals of sufficient energy appear at their input terminals. A high-pass filter between the broadband audio signal and the tweeter input terminals will prevent low-frequency program material from reaching the tweeter. In conjunction with a low-pass filter for the low-frequency driver (and possibly other filters for other drivers), the high-pass filter is part of what is known as a "crossover network".



TL/H/11221-20

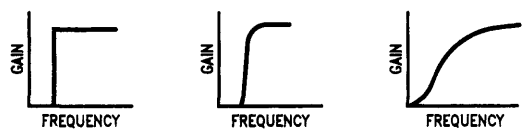
(a)



TL/H/11221-21

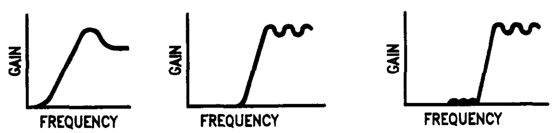
(b)

FIGURE 13. Amplitude (a) and Phase (b) Response Curves for Example High-Pass Filter



TL/H/11221-22

(a) (b) (c)



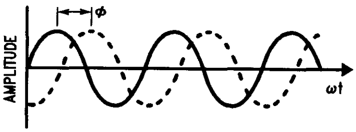
TL/H/11221-23

(d) (e) (f)

FIGURE 14. Examples of High-Pass Filter Amplitude Response Curves

All-Pass or Phase-Shift

The fifth and final filter response type has no effect on the amplitude of the signal at different frequencies. Instead, its function is to change the phase of the signal without affecting its amplitude. This type of filter is called an **all-pass** or **phase-shift** filter. The effect of a shift in phase is illustrated in *Figure 15*. Two sinusoidal waveforms, one drawn in dashed lines, the other a solid line, are shown. The curves are identical except that the peaks and zero crossings of the dashed curve occur at later times than those of the solid curve. Thus, we can say that the dashed curve has undergone a **time delay** relative to the solid curve.



TL/H/11221-24

FIGURE 15. Two sinusoidal waveforms with phase difference θ. Note that this is equivalent to a time delay $\frac{\theta}{\omega}$.

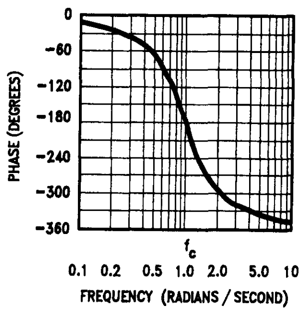
Since we are dealing here with periodic waveforms, time and phase can be interchanged—the time delay can also be interpreted as a **phase shift** of the dashed curve relative to the solid curve. The phase shift here is equal to θ radians. The relation between time delay and phase shift is $T_D = \theta/2\pi\omega$, so if phase shift is constant with frequency, time delay will decrease as frequency increases.

All-pass filters are typically used to introduce phase shifts into signals in order to cancel or partially cancel any unwanted phase shifts previously imposed upon the signals by other circuitry or transmission media.

Figure 16 shows a curve of phase vs frequency for an all-pass filter with the transfer function

$$H_{AP}(s) = \frac{s^2 - s + 1}{s^2 + s + 1}$$

The absolute value of the gain is equal to unity at all frequencies, but the phase changes as a function of frequency.



TL/H/11221-25

FIGURE 16. Phase Response Curve for Second-Order All-Pass Filter of Example

Let's take another look at the transfer function equations and response curves presented so far. First note that all of the transfer functions share the same denominator. Also note that all of the numerators are made up of terms found in the denominator: the high-pass numerator is the first term (s^2) in the denominator, the bandpass numerator is the second term (s), the low-pass numerator is the third term (1), and the notch numerator is the sum of the denominator's first and third terms ($s^2 + 1$). The numerator for the all-pass transfer function is a little different in that it includes all of the denominator terms, but one of the terms has a negative sign.

Second-order filters are characterized by four basic properties: the filter **type** (high-pass, bandpass, etc.), the **passband gain** (all the filters discussed so far have unity gain in the passband, but in general filters can be built with any gain), the **center frequency** (one radian per second in the above examples), and the filter **Q**. Q was mentioned earlier in connection with bandpass and notch filters, but in second-order filters it is also a useful quantity for describing the behavior of the other types as well. The Q of a second-order filter of a given type will determine the **relative shape** of the amplitude response. Q can be found from the denominator of the transfer function if the denominator is written in the form:

$$D(s) = s^2 + \frac{\omega_O}{Q} s + \omega_O^2.$$

As was noted in the case of the bandpass and notch functions, Q relates to the "sharpness" of the amplitude response curve. As Q increases, so does the sharpness of the response. Low-pass and high-pass filters exhibit "peaks" in their response curves when Q becomes large. *Figure 17* shows amplitude response curves for second-order bandpass, notch, low-pass, high-pass and all-pass filters with various values of Q.

There is a great deal of symmetry inherent in the transfer functions we've considered here, which is evident when the amplitude response curves are plotted on a logarithmic frequency scale. For instance, bandpass and notch amplitude resonse curves are symmetrical about f_O (with log frequency scales). This means that their gains at $2f_O$ will be the same as their gains at $f_O/2$, their gains at $10f_O$ will be the same as their gains at $f_O/10$, and so on.

The low-pass and high-pass amplitude response curves also exhibit symmetry, but with each other rather than with themselves. They are effectively mirror images of each other about f_O. Thus, the high-pass gain at $2f_O$ will equal the low-pass gain at $f_O/2$ and so on. The similarities between the various filter functions prove to be quite helpful when designing complex filters. Most filter designs begin by defining the filter as though it were a low-pass, developing a low-pass "prototype" and then converting it to bandpass, high-pass or whatever type is required after the low-pass characteristics have been determined.

As the curves for the different filter types imply, the number of possible filter response curves that can be generated is infinite. The differences between different filter responses within one filter type (e.g., low-pass) can include, among others, characteristic frequencies, filter order, roll-off slope, and flatness of the passband and stopband regions. The transfer function ultimately chosen for a given application will often be the result of a tradeoff between the above characteristics.

1.3 Elementary Filter Mathematics

In 1.1 and 1.2, a few simple passive filters were described and their transfer functions were shown. Since the filters were only 2nd-order networks, the expressions associated with them weren't very difficult to derive or analyze. When the filter in question becomes more complicated than a simple 2nd-order network, however, it helps to have a general

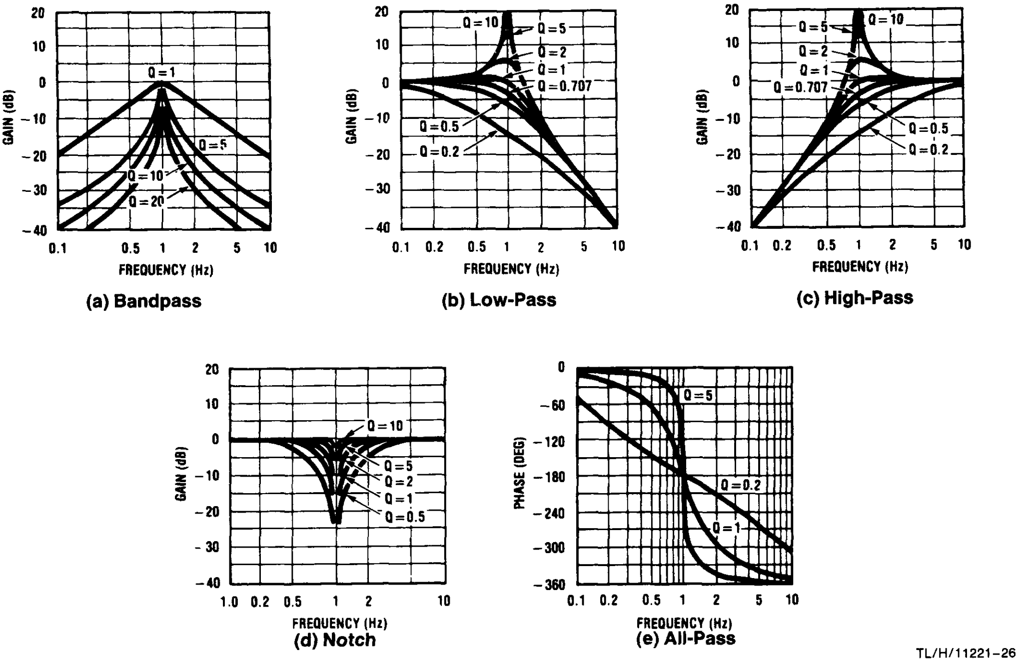


FIGURE 17. Responses of various 2nd-order filters as a function of Q. Gains and center frequencies are normalized to unity.

mathematical method of describing its characteristics. This allows us to use standard terms in describing filter characteristics, and also simplifies the application of computers to filter design problems.

The transfer functions we will be dealing with consist of a numerator divided by a denominator, each of which is a function of s, so they have the form:

$$H(s) = \frac{N(s)}{D(s)} \qquad (13)$$

Thus, for the 2nd-order bandpass example described in (4),

$$H_{BP}(s) = \frac{s}{s^2 + s + 1},$$

we would have $N(s) = s$, and $D(s) = s^2 + s + 1$.

The numerator and denominator can always be written as polynomials in s, as in the example above. To be completely general, a transfer function for an nth-order network, (one with "n" capacitors and inductors), can be written as below.

$$H(s) = H_0 \frac{s^n + b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \ldots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + a_{n-2}s^{n-2} + \ldots + a_1 s + a_0} \qquad (14)$$

This appears complicated, but it means simply that a filter's transfer function can be mathematically described by a numerator divided by a denominator, with the numerator and denominator made up of a number of terms, each consisting of a constant multiplied by the variable "s" to some power. The a_i and b_i terms are the constants, and their subscripts correspond to the order of the "s" term each is associated with. Therefore, a_1 is multiplied by s, a_2 is multiplied by s^2, and so on. Any filter transfer function (including the 2nd-order bandpass of the example) will have the general form of (14), with the values of the coefficients a_i and b_i depending on the particular filter.

The values of the coefficients completely determine the characteristics of the filter. As an example of the effect of changing just one coefficient, refer again to *Figure 17*, which shows the amplitude and phase response for 2nd-order bandpass filters with different values of Q. The Q of a 2nd-order bandpass is changed simply by changing the coefficient a_1, so the curves reflect the influence of that coefficient on the filter response.

Note that if the coefficients are known, we don't even have to write the whole transfer function, because the expression can be reconstructed from the coefficients. In fact, in the interest of brevity, many filters are described in filter design tables solely in terms of their coefficients. Using this aproach, the 2nd-order bandpass of *Figure 1* could be sufficiently specified by "$a_0 = a_1 = a_2 = b_1 = 1$", with all other coefficients equal to zero.

Another way of writing a filter's transfer function is to factor the polynomials in the numerator and denominator so that they take the form:

$$H(s) = H_0 \frac{(s - z_0)(s - z_1)(s - z_2) \ldots (s - z_n)}{(s - p_0)(s - p_1)(s - p_2) \ldots (s - p_n)} \qquad (15)$$

The roots of the numerator, $z_0, z_1, z_2, \ldots z_n$ are known as **zeros**, and the roots of the denominator, $p_0, p_1, \ldots p_n$ are called **poles**. z_i and p_i *are in general complex numbers, i.e.,* $R + jI$, where R is the real part, $j = \sqrt{-1}$, and I is the imaginary part. All of the poles and zeros will be either real roots (with no imaginary part) or complex conjugate pairs. A complex conjugate pair consists of two roots, each of which has a real part and an imaginary part. The imaginary parts of the two members of a complex conjugate pair will have opposite signs and the reals parts will be equal. For example, the 2nd-order bandpass network function of (4) can be factored to give:

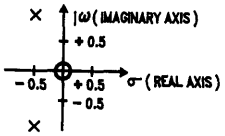
$$H(s) = \frac{s}{\left(s + 0.5 + j\frac{\sqrt{3}}{2}\right)\left(s + 0.5 - j\frac{\sqrt{3}}{2}\right)} \quad (16)$$

The factored form of a network function can be depicted graphically in a **pole-zero diagram**. *Figure 18* is the pole-zero diagram for equation (4). The diagram shows the zero at the origin and the two poles, one at

$$s = -0.5 - j\sqrt{3}/2,$$

and one at

$$s = -0.5 + j\sqrt{3}/2.$$



TL/H/11221-27

FIGURE 18. Pole-Zero Diagram for the Filter in *Figure 2*

The pole-zero diagram can be helpful to filter designers as an aid in visually obtaining some insight into a network's characteristics. A pole anywhere to the right of the imaginary axis indicates instability. If the pole is located on the positive real axis, the network output will be an increasing exponential function. A positive pole not located on the real axis will give an exponentially increasing sinusoidal output. We obviously want to avoid filter designs with poles in the right half-plane!

Stable networks will have their poles located on or to the left of the imaginary axis. Poles on the imaginary axis indicate an undamped sinusoidal output (in other words, a sine-wave oscillator), while poles on the left real axis indicate damped exponential response, and complex poles in the negative half-plane indicate damped sinusoidal response. The last two cases are the ones in which we will have the most interest, as they occur repeatedly in practical filter designs.

Another way to arrange the terms in the network function expression is to recognize that each complex conjugate pair is simply the factored form of a second-order polynomial. By multiplying the complex conjugate pairs out, we can get rid of the complex numbers and put the transfer function into a form that essentially consists of a number of 2nd-order transfer functions multiplied together, possibly with some first-order terms as well. We can thus think of the complex filter as being made up of several 2nd-order and first-order filters connected in series. The transfer function thus takes the form:

$$H(s) = H_0 \frac{(s^2 + b_{11}s + b_{10})(s^2 + b_{21}s + b_{20})\ldots}{(s^2 + a_{11}s + a_{10})(s^2 + a_{21}s + a_{20})\ldots} \quad (17)$$

This form is particularly useful when you need to design a complex active or switched-capacitor filter. The general approach for designing these kinds of filters is to cascade second-order filters to produce a higher-order overall response. By writing the transfer function as the product of second-order polynomials, we have it in a form that directly corresponds to a cascade of second-order filters. For example, the fourth-order low-pass filter transfer function

$$H_{LP}(s) = \frac{1}{(s^2 + 1.5s + 1)(s^2 + 1.2s + 1)} \quad (18)$$

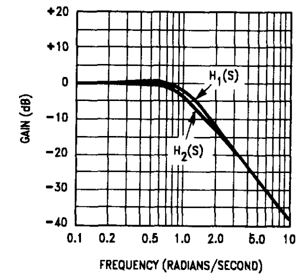
can be built by cascading two second-order filters with the transfer functions

(19)

and

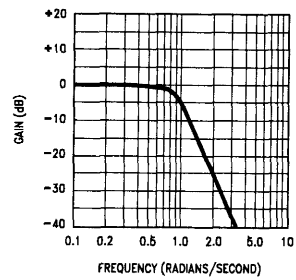
$$H_2(s) = \frac{1}{(s^2 + 1.2s + 1)} \quad (20)$$

This is illustrated in *Figure 19*, which shows the two 2nd-order amplitude responses together with the combined 4th-order response.



TL/H/11221-28

(a)



TL/H/11221-29

(b)

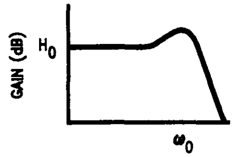
FIGURE 19. Two Second-Order Low-Pass Filters (a) can be Cascaded to Build a Fourth-Order Filter (b).

Instead of the coefficients a_0, a_1, etc., second-order filters can also be described in terms of parameters that relate to observable quantities. These are the filter gain H_0, the characteristics radian frequency ω_O, and the filter Q. For the general second-order low-pass filter transfer function we have:

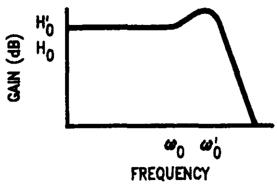
$$H(s) = \frac{H_0 a_0}{(s^2 + a_1 s + a_0)} = \frac{H_0 \omega_0^2}{\left(s^2 + \frac{\omega_0}{Q}s + \omega_0^2\right)} \quad (21)$$

which yields: $\omega^2_0 = a_0$, and $Q = \omega_0/a_1 = \sqrt{a_0}/a_1$.

The effects of H_0 and ω_0 on the amplitude response are straightforward: H_0 is the gain scale factor and ω_0 is the frequency scale factor. Changing one of these parameters will alter the amplitude or frequency scale on an amplitude response curve, but the shape, as shown in *Figure 20*, will remain the same. The basic shape of the curve is determined by the filter's Q, which is determined by the denominator of the transfer function.



(a) TL/H/11221-30



(b) TL/H/11221-31

FIGURE 20. Effect of changing H_0 and ω_0. Note that, when log frequency and gain scales are used, a change in gain or center frequency has no effect on the shape of the response curve. Curve shape is determined by Q.

1.4 Filter Approximations

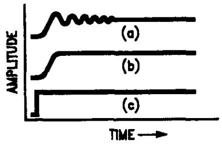
In Section 1.2 we saw several examples of amplitude response curves for various filter types. These always included an "ideal" curve with a rectangular shape, indicating that the boundary between the passband and the stopband was abrupt and that the rolloff slope was infinitely steep. This type of response would be ideal because it would allow us to completely separate signals at different frequencies from one another. Unfortunately, such an amplitude response curve is not physically realizable. We will have to settle for the best **approximation** that will still meet our requirements for a given application. Deciding on the best approximation involves making a compromise between various properties of the filter's transfer function. The important properties are listed below.

Filter Order. The order of a filter is important for several reasons. It is directly related to the number of components in the filter, and therefore to its cost, its physical size, and the complexity of the design task. Therefore, higher-order filters are more expensive, take up more space, and are more difficult to design. The primary advantage of a higher-order filter is that it will have a steeper rolloff slope than a similar lower-order filter.

Ultimate Rolloff Rate. Usually expressed as the amount of attenuation in dB for a given ratio of frequencies. The most common units are "dB/octave" and "dB/decade". While the ultimate rolloff rate will be 20 dB/decade for every filter pole in the case of a low-pass or high-pass filter and 20 dB/decade for every pair of poles for a bandpass filter, some filters will have steeper attenuation slopes near the cutoff frequency than others of the same order.

Attenuation Rate Near the Cutoff Frequency. If a filter is intended to reject a signal very close in frequency to a signal that must be passed, a sharp cutoff characteristic is desirable between those two frequencies. Note that this steep slope may not continue to frequency extremes.

Transient Response. Curves of amplitude response show how a filter reacts to steady-state sinusoidal input signals. Since a real filter will have far more complex signals applied to its input terminals, it is often of interest to know how it will behave under transient conditions. An input signal consisting of a step function provides a good indication of this. *Figure 21* shows the responses of two low-pass filters to a step input. Curve (b) has a smooth reaction to the input step, while curve (a) exhibits some **ringing**. As a rule of thumb, filters will sharper cutoff characteristics or higher Q will have more pronounced ringing.



TL/H/11221-32

FIGURE 21. Step response of two different filters. Curve (a) shows significant "ringing", while curve (b) shows none. The input signal is shown in curve (c).

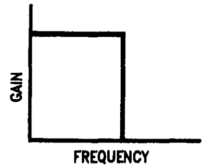
Monotonicity. A filter has a monotonic amplitude response if its gain slope never changes sign—in other words, if the gain always increases with increasing frequency or always decreases with increasing frequency. Obviously, this can happen only in the case of a low-pass or high-pass filter. A bandpass or notch filter can be monotonic on either side of the center frequency, however. *Figures 11(b)* and *(c)* and *14(b)* and *(c)* are examples of monotonic transfer functions.

Passband Ripple. If a filter is not monotonic within its passband, the transfer function within the passband will exhibit one or more "bumps". These bumps are known as "ripple". Some systems don't necessarily require monotonicity, but do require that the passband ripple be limited to some maximum value (usually 1 dB or less). Examples of passband ripple can be found in *Figures 5(e)* and *(f)*, *8(f)*, *11(e)* and *(f)*, and *14(e)* and *(f)*. Although bandpass and notch filters do not have monotonic transfer functions, they can be free of ripple within their passbands.

Stopband Ripple. Some filter responses also have ripple in the stopbands. Examples are shown in *Figure 5(f)*, *8(g)*, *11(f)*, and *14(f)*. We are normally unconcerned about the amount of ripple in the stopband, as long as the signal to be rejected is sufficiently attenuated.

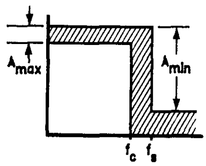
Given that the "ideal" filter amplitude response curves are not physically realizable, we must choose an acceptable approximation to the ideal response. The word "acceptable" may have different meanings in different situations.

The acceptability of a filter design will depend on many interrelated factors, including the amplitude response characteristics, transient response, the physical size of the circuit and the cost of implementing the design. The "ideal" low-pass amplitude response is shown again in *Figure 22(a)*. If we are willing to accept some deviations from this ideal in order to build a practical filter, we might end up with a curve like the one in *Figure 22(b)*, which allows ripple in the passband, a finite attenuation rate, and stopband gain greater than zero. Four parameters are of concern in the figure:



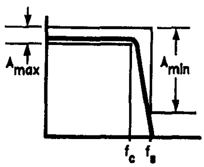
TL/H/11221-33

(a) "ideal" Low-Pass Filter Response



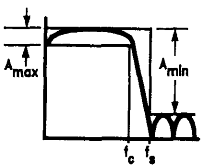
TL/H/11221-34

(b) Amplitude Response Limits for a Practical Low-Pass Filter



TL/H/11221-35

(c) Example of an Amplitude Response Curve Falling with the Limits Set by f_c, f_s, A_{min}, and A_{max}



TL/H/11221-36

(d) Another Amplitude Response Falling within the Desired Limits

FIGURE 22

$\mathbf{A_{max}}$ is the maximum allowable change in gain within the passband. This quantity is also often called the maximum passband ripple, but the word "ripple" implies non-monotonic behavior, while A_{max} can obviously apply to monotonic response curves as well.

$\mathbf{A_{min}}$ is the minimum allowable attenuation (referred to the maximum passband gain) within the stopband.

$\mathbf{f_c}$ is the cutoff frequency or passband limit.

$\mathbf{f_s}$ is the frequency at which the stopband begins.

If we can define our filter requirements in terms of these parameters, we will be able to design an acceptable filter using standard "cookbook" design methods. It should be apparent that an unlimited number of different amplitude response curves could fit within the boundaries determined by these parameters, as illustrated in *Figure 22(c) and (d)*. Filters with acceptable amplitude response curves may differ in terms of such characteristics as transient response, passband and stopband flatness, and complexity. How does one choose the best filter from the infinity of possible transfer functions?

Fortunately for the circuit designer, a great deal of work has already been done in this area, and a number of standard filter characteristics have already been defined. These usually provide sufficient flexibility to solve the majority of filtering problems.

The "classic" filter functions were developed by mathematicians (most bear their inventors' names), and each was designed to optimize some filter property. The most widely-used of these are discussed below. No attempt is made here to show the mathematical derivations of these functions, as they are covered in detail in numerous texts on filter theory.

Butterworth

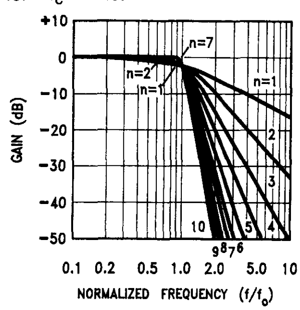
The first, and probably best-known filter approximation is the **Butterworth** or **maximally-flat** response. It exhibits a nearly flat passband with no ripple. The rolloff is smooth and monotonic, with a low-pass or high-pass rolloff rate of 20 dB/decade (6 dB/octave) for every pole. Thus, a 5th-order Butterworth low-pass filter would have an attenuation rate of 100 dB for every factor of ten increase in frequency beyond the cutoff frequency.

The general equation for a Butterworth filter's amplitude response is

$$H(\omega) = \frac{1}{1 + \left(\frac{\omega}{\omega_0}\right)^{2n}} \qquad (22)$$

where n is the order of the filter, and can be any positive whole number (1, 2, 3, . . .), and ω is the -3 dB frequency of the filter.

Figure 23 shows the amplitude response curves for Butterworth low-pass filters of various orders. The frequency scale is normalized to $f/f_{-3\ dB}$ so that all of the curves show 3 dB attenuation for $f/f_C = 1.0$.



TL/H/11221-37

FIGURE 23. Amplitude Response Curves for Butterworth Filters of Various Orders

The coefficients for the denominators of Butterworth filters of various orders are shown in Table 1(a). Table 1(b) shows the denominators factored in terms of second-order polynomials. Again, all of the coefficients correspond to a corner frequency of 1 radian/s (finding the coefficients for a different cutoff frequency will be covered later). As an example,

TABLE 1(a). Butterworth Polynomials

Denominator coefficients for polynomials of the form $s^n + a_{n-1}s^{n-1} + a_{n-2}s^{n-2} + \ldots + a_1s + a_0$.

n	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
1	1									
2	1	1.414								
3	1	2.000	2.000							
4	1	2.613	3.414	2.613						
5	1	3.236	5.236	5.236	3.236					
6	1	3.864	7.464	9.142	7.464	3.864				
7	1	4.494	10.098	14.592	14.592	10.098	4.494			
8	1	5.126	13.137	21.846	25.688	21.846	13.137	5.126		
9	1	5.759	16.582	31.163	41.986	41.986	31.163	16.582	5.759	
10	1	6.392	20.432	42.802	64.882	74.233	64.882	42.802	20.432	6.392

TABLE 1(b). Butterworth Quadratic Factors

n	
1	$(s + 1)$
2	$(s^2 + 1.4142s + 1)$
3	$(s + 1)(s^2 + s + 1)$
4	$(s^2 + 0.7654s + 1)(s^2 + 1.8478s + 1)$
5	$(s + 1)(s^2 + 0.6180s + 1)(s^2 + 1.6180s + 1)$
6	$(s^2 + 0.5176s + 1)(s^2 + 1.4142s + 1)(s^2 + 1.9319)$
7	$(s + 1)(s^2 + 0.4450s + 1)(s^2 + 1.2470s + 1)(s^2 + 1.8019s + 1)$
8	$(s^2 + 0.3902s + 1)(s^2 + 1.1111s + 1)(s^2 + 1.6629s + 1)(s^2 + 1.9616s + 1)$
9	$(s + 1)(s^2 + 0.3473s + 1)(s^2 + 1.0000s + 1)(s^2 + 1.5321s + 1)(s^2 + 1.8794s + 1)$
10	$(s^2 + 0.3129s + 1)(s^2 + 0.9080s + 1)(s^2 + 1.4142s + 1)(s^2 + 1.7820s + 1)(s^2 + 1.9754s + 1)$

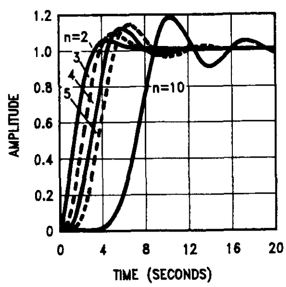
the tables show that a fifth-order Butterworth low-pass filter's transfer function can be written:

$$H(s) = \frac{1}{s^5 + 3.236s^4 + 5.236s^3 + 5.236s^2 + 3.236s + 1} \quad (22)$$

$$= \frac{1}{(s + 1)(s^2 + 0.6180s + 1)(s^2 + 1.6180s + 1)}$$

This is the product of one first-order and two second-order transfer functions. Note that neither of the second-order transfer functions alone is a Butterworth transfer function, but that they both have the same center frequency.

Figure 24 shows the step response of Butterworth low-pass filters of various orders. Note that the amplitude and duration of the ringing increases as n increases.

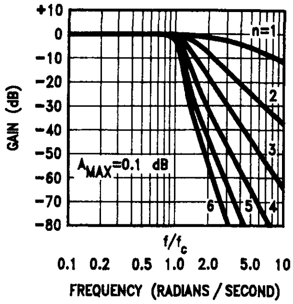


TL/H/11221-38

FIGURE 24. Step responses for Butterworth low-pass filters. In each case $\omega_0 = 1$ and the step amplitude is 1.0.

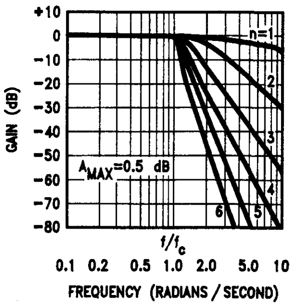
Chebyshev

Another approximation to the ideal filter is the **Chebyshev** or **equal ripple** response. As the latter name implies, this sort of filter will have ripple in the passband amplitude response. The amount of passband ripple is one of the parameters used in specifying a Chebyshev filter. The Chebyschev characteristic has a steeper rolloff near the cutoff frequency when compared to the Butterworth, but at the expense of monotonicity in the passband and poorer transient response. A few different Chebyshev filter responses are shown in *Figure 25*. The filter responses in the figure have 0.1 dB and 0.5 dB ripple in the passband, which is small compared to the amplitude scale in *Figure 25(a)* and *(b)*, so it is shown expanded in *Figure 25(c)*.



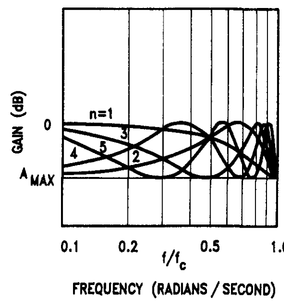
(a)

TL/H/11221-39



(b)

TL/H/11221-40



(c)

TL/H/11221-41

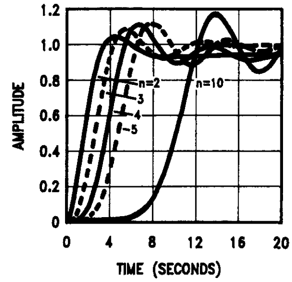
FIGURE 25. Examples of Chebyshev amplitude responses. (a) 0.1 dB ripple (b) 0.5 dB ripple. (c) Expanded view of passband region showing form of response below cutoff frequency.

Note that a Chebyshev filter of order n will have n − 1 peaks or dips in its passband response. Note also that the nominal gain of the filter (unity in the case of the responses in *Figure 25*) is equal to he filter's maximum passband gain. An odd-order Chebyshev will have a dc gain (in the low-pass case) equal to the nominal gain, with "dips" in the amplitude response curve equal to the ripple value. An even-order Chebyshev low-pass will have its dc gain equal to he nominal filter gain **minus** the ripple value; the nominal gain for an even-order Chebyshev occurs at the peaks of the passband ripple. Therefore, if you're designing a fourth-order Chebyshev low-pass filter with 0.5 dB ripple and you want it to have unity gain at dc, you'll have to design for a nominal gain of 0.5 dB.

The cutoff frequency of a Chebyshev filter is not assumed to be the −3 dB frequency as in the case of a Butterworth filter. Instead, the Chebyshev's cutoff frequency is normally the frequency at which the ripple (or A_{max}) specification is exceeded.

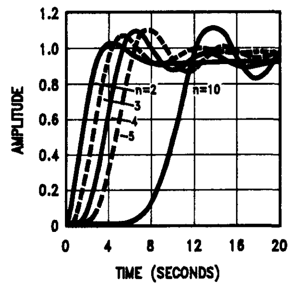
The addition of passband ripple as a parameter makes the specification process for a Chebyshev filter a bit more complicated than for a Butterworth filter, but also increases flexibility.

Figure 26 shows the step response of 0.1 dB and 0.5 dB ripple Chebyshev filters of various orders. As with the Butterworth filters, the higher order filters ring more.



(a) 0.1 dB Ripple

TL/H/11221-42



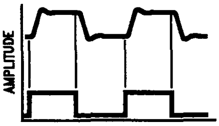
(b) 0.5 dB Ripple

TL/H/11221-43

FIGURE 26. Step responses for Chebyshev low-pass filters. In each case, $\omega_0 = 1$, and the step amplitude is 1.0.

Bessel

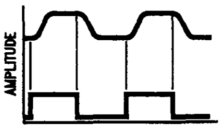
All filters exhibit phase shift that varies with frequency. This is an expected and normal characteristic of filters, but in certain instances it can present problems. If the phase increases linearly with frequency, its effect is simply to delay the output signal by a constant time period. However, if the phase shift is not directly proportional to frequency, components of the input signal at one frequency will appear at the output shifted in phase (or time) with respect to other frequencies. The overall effect is to distort non-sinusoidal waveshapes, as illustrated in *Figure 27* for a square wave passed through a Butterworth low-pass filter. The resulting waveform exhibits ringing and overshoot because the square wave's component frequencies are shifted in time with respect to each other so that the resulting waveform is very different from the input square wave.



TL/H/11221–44

FIGURE 27. Response of a 4th-order Butterworth low-pass (upper curve) to a square wave input (lower curve). The "ringing" in the response shows that the nonlinear phase shift distorts the filtered wave shape.

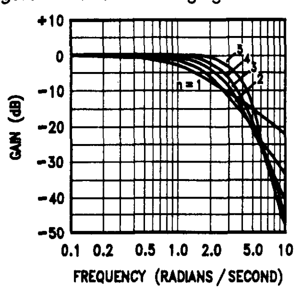
When the avoidance of this phenomenon is important, a **Bessel** or **Thompson** filter may be useful. The Bessel characteristic exhibits approximately linear phase shift with frequency, so its action within the passband simulates a delay line with a low-pass characteristic. The higher the filter order, the more linear the Bessel's phase response. *Figure 28* shows the square-wave response of a Bessel low-pass filter. Note the lack of ringing and overshoot. Except for the "rounding off" of the square wave due to the attenuation of high-frequency harmonics, the waveshape is preserved.



TL/H/11221–45

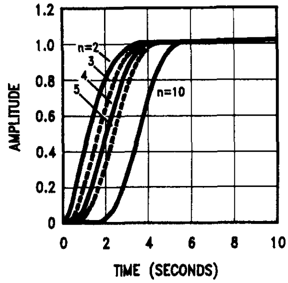
FIGURE 28. Response of a 4th-order Bessel low-pass (upper curve) to a square wave input (lower curve). Note the lack of ringing in the response. Except for the "rounding of the corners" due to the reduction of high frequency components, the response is a relatively undistorted version of the input square wave.

The amplitude response of the Bessel filter is monotonic and smooth, but the Bessel filter's cutoff characteristic is quite gradual compared to either the Butterworth or Chebyshev as can be seen from the Bessel low-pass amplitude response curves in *Figure 29*. Bessel step responses are plotted in *Figure 30* for orders ranging from 2 to 10.



TL/H/11221–46

FIGURE 29. Amplitude response curves for Bessel filters of various orders. The nominal delay of each filter is 1 second.



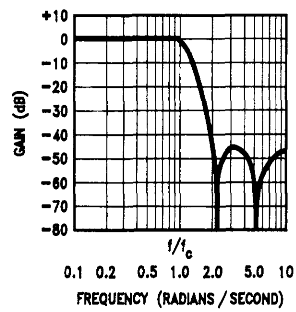
TL/H/11221–47

FIGURE 30. Step responses for Bessel low-pass filters. In each case, $\omega_0 = 1$ and the input step amplitude is 1.0.

Elliptic

The cutoff slope of an **elliptic** filter is steeper than that of a Butterworth, Chebyshev, or Bessel, but the amplitude response has ripple in both the passband and the stopband, and the phase response is very non-linear. However, if the primary concern is to pass frequencies falling within a certain frequency band and reject frequencies outside that band, regardless of phase shifts or ringing, the elliptic response will perform that function with the lowest-order filter.

The elliptic function gives a sharp cutoff by adding notches in the stopband. These cause the transfer function to drop to zero at one or more frequencies in the stopband. Ripple is also introduced in the passband (see *Figure 31*). An elliptic filter function can be specified by three parameters (again excluding gain and cutoff frequency): passband ripple, stopband attenuation, and filter order n. Because of the greater complexity of the elliptic filter, determination of coefficients is normally done with the aid of a computer.



TL/H/11221–48

FIGURE 31. Example of a elliptic low-pass amplitude response. This particular filter is 4th-order with A_{max} = 0.5 dB and $f_s/f_c = 2$. The passband ripple is similar in form to the Chebyshev ripple shown in *Figure 25(c)*.

1.5 Frequency Normalization and Denormalization

Filter coefficients that appear in tables such as Table 1 are **normalized** for cutoff frequencies of 1 radian per second, or $\omega_O = 1$. Therefore, if these coefficients are used to generate a filter transfer function, the cutoff (or center) frequency of the transfer function will be at $\omega = 1$. This is a convenient way to standardize filter coefficients and transfer functions. If this were not done, we would need to produce a different set of coefficients for every possible center frequency. Instead, we use coefficients that are normalized for $\omega_O = 1$ because it is simple to rescale the frequency behavior of a 1 r.p.s. filter. In order to **denormalize** a transfer function we merely replace each "s" term in the transfer function with s/ω_O, where ω_O is the desired cutoff frequency. Thus the second-order Butterworth low-pass function

$$H(s) = \frac{1}{(s^2 + 2s + 1)} \qquad (23)$$

could be denormalized to have a cutoff frequency of 1000 Hz by replacing s with $s/2000\pi$ as below:

$$H(s) = \frac{1}{\frac{s^2}{4 \times 10^6\pi^2} + \frac{\sqrt{2}s}{2000\pi} + 1}$$

$$= \frac{4 \times 10^6\pi^2}{s^2 + 2828.4\pi s + 4 \times 10^6\pi^2}$$

$$= \frac{3.948 \times 10^7}{s^2 + 8885.8s + 3.948 \times 10^7}$$

If it is necessary to normalize a transfer function, the opposite procedure can be performed by replacing each "s" in the transfer function with $\omega_O s$.

APPROACHES TO IMPLEMENTING FILTERS: ACTIVE, PASSIVE, AND SWITCHED-CAPACITOR

2.1 Passive Filters

The filters used for the earlier examples were all made up of passive components: resistors, capacitors, and inductors, so they are referred to as **passive filters**. A passive filter is simply a filter that uses no amplifying elements (transistors, operational amplifiers, etc.). In this respect, it is the simplest (in terms of the number of necessary components) implementation of a given transfer function. Passive filters have other advantages as well. Because they have no active components, passive filters require no power supplies. Since they are not restricted by the bandwidth limitations of op amps, they can work well at very high frequencies. They can be used in applications involving larger current or voltage levels than can be handled by active devices. Passive filters also generate little nosie when compared with circuits using active gain elements. The noise that they produce is simply the thermal noise from the resistive components, and, with careful design, the amplitude of this noise can be very low.

Passive filters have some important disadvantages in certain applications, however. Since they use no active elements, they cannot provide signal gain. Input impedances can be lower than desirable, and output impedances can be higher the optimum for some applications, so buffer amplifiers may be needed. Inductors are necessary for the synthesis of most useful passive filter characteristics, and these can be prohibitively expensive if high accuracy (1% or 2%, for example), small physical size, or large value are required. Standard values of inductors are not very closely spaced, and it is diffcult to find an off-the-shelf unit within 10% of any arbitrary value, so adjustable inductors are often used. Tuning these to the required values is time-consuming and expensive when producing large quantities of filters. Futhermore, complex passive filters (higher than 2nd-order) can be difficult and time-consuming to design.

2.2 Active Filters

Active filters use amplifying elements, especially op amps, with resistors and capacitors in their feedback loops, to synthesize the desired filter characteristics. Active filters can have high input impedance, low output impedance, and virtually any arbitrary gain. They are also usually easier to design than passive filters. Possibly their most important attribute is that they lack inductors, thereby reducing the problems associated with those components. Still, the problems of accuracy and value spacing also affect capacitors, although to a lesser degree. Performance at high frequencies is limited by the gain-bandwidth product of the amplifying elements, but within the amplifier's operating frequency range, the op amp-based active filter can achieve very good accuracy, provided that low-tolerance resistors and capacitors are used. Active filters will generate noise due to the amplifying circuitry, but this can be minimized by the use of low-noise amplifiers and careful circuit design.

Figure 32 shows a few common active filter configurations (There are several other useful designs; these are intended to serve as examples). The second-order Sallen-Key low-pass filter in (a) can be used as a building block for higher-order filters. By cascading two or more of these circuits, filters with orders of four or greater can be built. The two resistors and two capacitors connected to the op amp's non-inverting input and to V_{IN} determine the filter's cutoff frequency and affect the Q; the two resistors connected to the inverting input determine the gain of the filter and also affect the Q. Since the components that determine gain and cutoff frequency also affect Q, the gain and cutoff frequency can't be independently changed.

Figures 32(b) and *32(c)* are multiple-feedback filters using one op amp for each second-order transfer function. Note that each high-pass filter stage in *Figure 32(b)* requires three capacitors to achieve a second-order response. As with the Sallen-Key filter, each component value affects more than one filter characteristic, so filter parameters can't be independently adjusted.

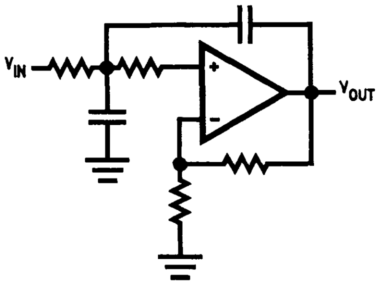
The second-order state-variable filter circuit in *Figure 32(d)* requires more op amps, but provides high-pass, low-pass, and bandpass outputs from a single circuit. By combining the signals from the three outputs, any second-order transfer function can be realized.

When the center frequency is very low compared to the op amp's gain-bandwidth product, the characteristics of active RC filters are primarily dependent on external component tolerances and temperature drifts. For predictable results in critical filter circuits, external components with very good absolute accuracy and very low sensitivity to temperature variations must be used, and these can be expensive.

When the center frequency multiplied by the filter's Q is more than a small fraction of the op amp's gain-bandwidth product, the filter's response will deviate from the ideal transfer function. The degree of deviation depends on the filter topology; some topologies are designed to minimize the effects of limited op amp bandwidth.

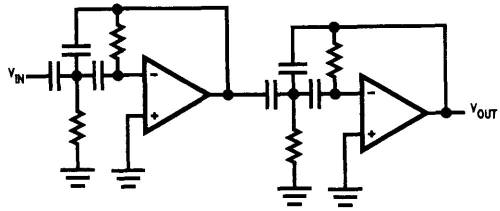
2.3 The Switched-Capacitor Filter

Another type of filter, called the **switched-capacitor filter**, has become widely available in monolithic form during the last few years. The switched-capacitor approach overcomes some of the problems inherent in standard active filters, while adding some interesting new capabilities. Switched-capacitor filters need no external capacitors or inductors, and their cutoff frequencies are set to a typical accuracy of $\pm 0.2\%$ by an external clock frequency. This allows consistent, repeatable filter designs using inexpensive crystal-controlled oscillators, or filters whose cutoff frequencies are variable over a wide range simply by changing the clock frequency. In addition, switched-capacitor filters can have low sensitivity to temperature changes.



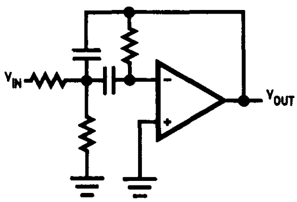
TL/H/11221-49

(a) Sallen-Key 2nd-Order Active Low-Pass Filter



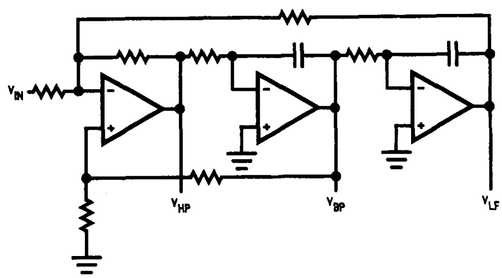
TL/H/11221-50

(b) Multiple-Feedback 4th-Order Active High-Pass Filter. Note that there are more capacitors than poles.



TL/H/11221-51

(c) Multiple-Feedback 2nd-Order Bandpass Filter



TL/H/11221-52

(d) Universal State-Variable 2nd-Order Active Filter

FIGURE 32. Examples of Active Filter Circuits Based on Op Amps, Resistors, and Capacitors

Switched-capacitor filters are clocked, sampled-data systems; the input signal is sampled at a high rate and is processed on a discrete-time, rather than continuous, basis. This is a fundamental difference between switched-capacitor filters and conventional active and passive filters, which are also referred to as "continuous time" filters.

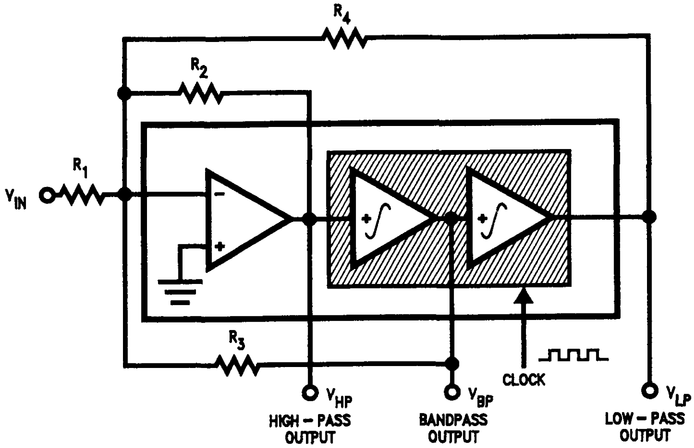
The operation of switched-capacitor filters is based on the ability of on-chip capacitors and MOS switches to simulate resistors. The values of these on-chip capacitors can be closely matched to other capacitors on the IC, resulting in integrated filters whose cutoff frequencies are proportional to, and determined only by, the external clock frequency. Now, these integrated filters are nearly always based on state-variable active filter topologies, so they are also active filters, but normal terminology reserves the name "active filter" for filters built using non-switched, or continuous, active filter techniques. The primary weakness of switched-capacitor filters is that they have more noise at their outputs—both random noise and clock feedthrough—than standard active filter circuits.

National Semiconductor builds several different types of switched-capacitor filters. Three of these, the LMF100, the MF5, and the MF10, can be used to synthesize any of the filter types described in Section 1.2, simply by appropriate choice of a few external resistors. The values and placement of these resistors determine the basic shape of the amplitude and phase response, with the center or cutoff frequency set by the external clock. *Figure 33* shows the filter block of the LMF100 with four external resistors connected to provide low-pass, high-pass, and bandpass outputs. Note that this circuit is similar in form to the universal state-variable filter in *Figure 32(d)*, except that the switched-capacitor filter utilizes **non-inverting** integrators, while the conventional active filter uses **inverting** integrators. Changing the switched-capacitor filter's clock frequency changes the value of the integrator resistors, thereby proportionately changing the filter's center frequency. The LMF100 and MF10 each contain two universal filter blocks, while the MF5 has a single second-order filter.

While the LMF100, MF5, and MF10 are **universal filters**, capable of realizing all of the filter types, the LMF40, LMF60, MF4, and MF6 are configured only as fourth- or sixth-order Butterworth low-pass filters, with no external components necessary other than a clock (to set f_O) and a power supply. *Figures 34* and *35* show typical LMF40 and LMF60 circuits along with their amplitude response curves.

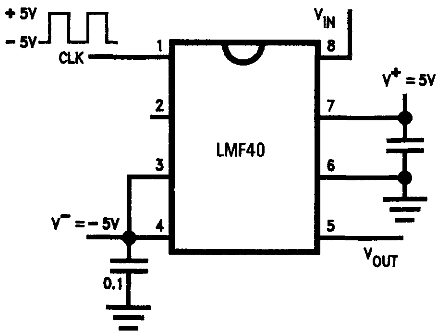
Some switched-capacitor filter products are very specialized. The LMF380 *(Figure 36)* contains three fourth-order Chebyshev bandpass filters with bandwidths and center frequency spacings equal to one-third of an octave. This filter is designed for use with audio and acoustical instrumentation and needs no external components other than a clock. An internal clock oscillator can, with the aid of a crystal and two capacitors, generate the master clock for a whole array of LMF380s in an audio real-time analyzer or other multi-filter instrument.

Other devices, such as the MF8 fourth-order bandpass filter *(Figure 37)* and the LMF90 fourth-order notch filter *(Figure 38)* have specialized functions but may be programmed for a variety of response curves using external resistors in the case of the MF8 or logic inputs in the case of the LMF90.



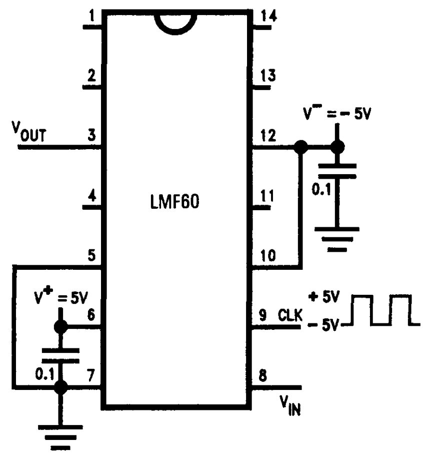
TL/H/11221-53

FIGURE 33. Block diagram of a second-order universal switched-capacitor filter, including external resistors connected to provide High-Pass, Bandpass, and Low-Pass outputs. Notch and All-Pass responses can be obtained with different external resistor connections. The center frequency of this filter is proportional to the clock frequency. Two second-order filters are included on the LMF100 or MF10.



TL/H/11221-54

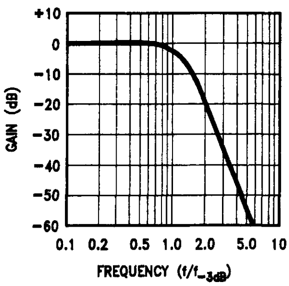
(a)



TL/H/11221-55

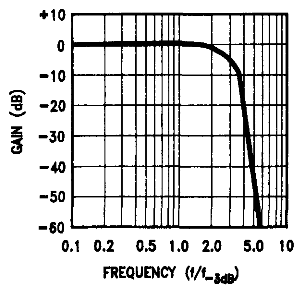
(b)

FIGURE 34. Typical LMF40 and LMF60 application circuits. The circuits shown operate on ±5V power supplies and accept CMOS clock levels. For operation on single supplies or with TTL clock levels, see *Sections 2.3 and 2.4.*



TL/H/11221-56

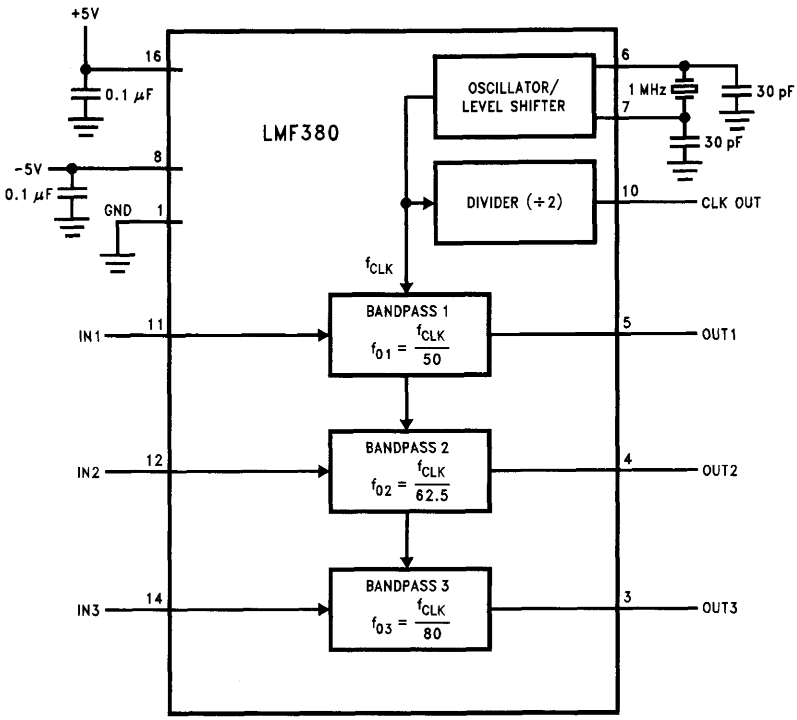
(a) LMF40



TL/H/11221-57

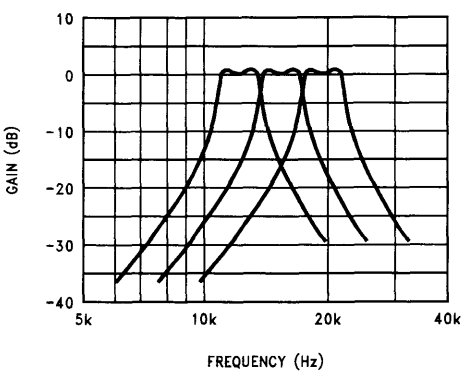
(b) LMF60

FIGURE 35. Typical LMF40 and LMF60 amplitude response curves. The cutoff frequency has been normalized to 1 in each case.



TL/H/11221-58

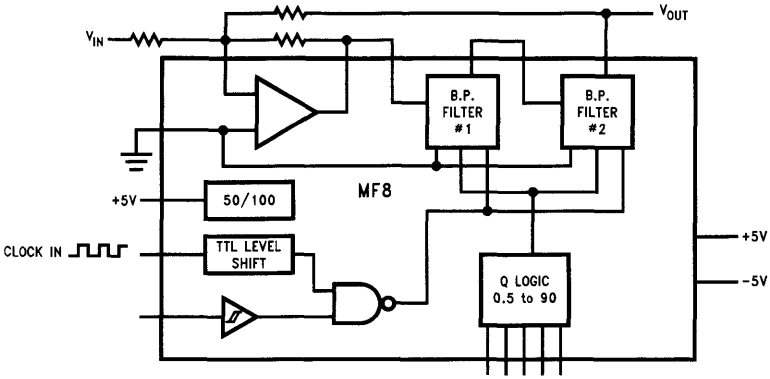
(a)



TL/H/11221-59

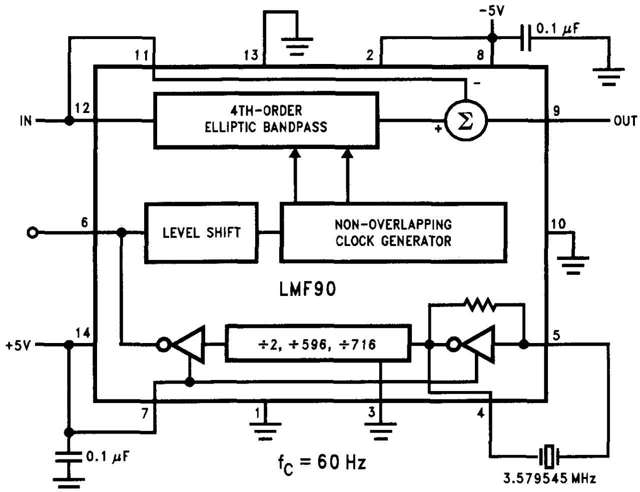
(b)

FIGURE 36. LMF380 one-third octave filter array. (a) Typical application circuit for the top audio octave. The clock is generated with the aid of the external crystal and two 30 pF capacitors. (b) Response curves for the three filters.



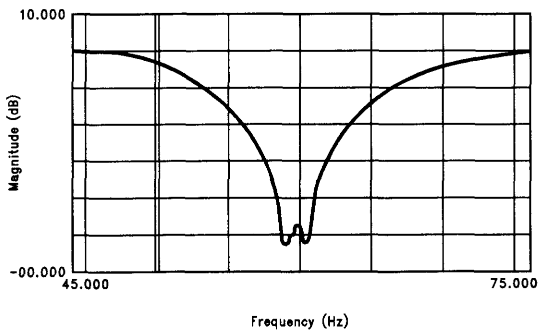
TL/H/11221-60

FIGURE 37. The MF8 is a fourth-order bandpass filter. Three external resistors determine the filter function. A five-bit digital input sets the bandwidth and the clock frequency determines the center frequency.



TL/H/11221-61

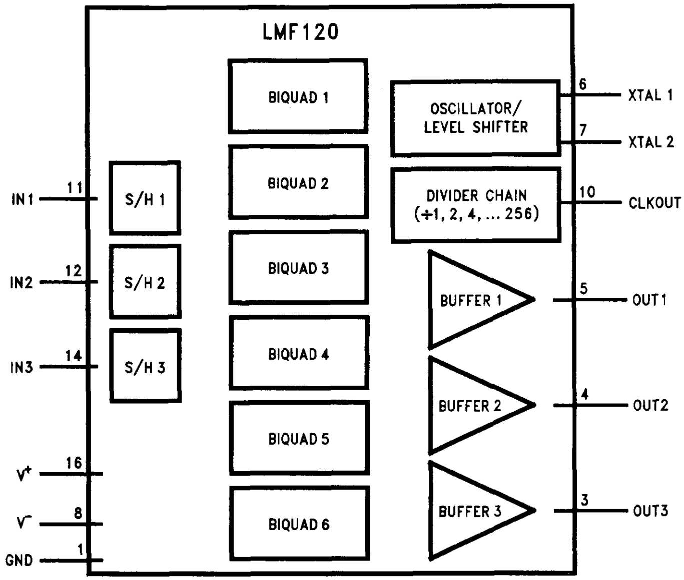
(a)



TL/H/11221-62

(b)

FIGURE 38. LMF90 fourth-order elliptic notch filter. The clock can be generated externally, or internally with the aid of a crystal. Using the circuit as shown in (a), a 60 Hz notch can be built. Connecting pin 3 to V+ yields a 50 Hz notch. By tying pin to ground or V+, the center frequency can be doubled or tripled. The response of the circuit in (a) is shown in (b).



TL/H/11221-63

FIGURE 39. Block diagram of the LMF120 customizable switched-capacitor filter array. The internal circuit blocks can be internally configured to provide up to three filters with a total of 12 poles. Any unused circuitry can be disconnected to reduce power consumption.

Finally, when a standard filter product for a specific application can't be found, it often makes sense to use a cell-based approach and build an application-specific filter. An example is the LMF120, a 12th-order customizable switched-capacitor filter array that can be configured to perform virtually any filtering function with no external components. A block diagram of this device is shown in *Figure 39*. The three input sample-and-hold circuits, six second-order filter blocks, and three output buffers can be interconnected to build from one to three filters, with a total order of twelve.

2.4 Which Approach Is Best—Active, Passive, or Switched-Capacitor?

Each filter technology offers a unique set of advantages and disadvantages that makes it a nearly ideal solution to some filtering problems and completely unacceptable in other applications. Here's a quick look at the most important differences between active, passive, and switched-capacitor filters.

Accuracy: Switched-capacitor filters have the advantage of better accuracy in most cases. Typical center-frequency accuracies are normally on the order of about 0.2% for most switched-capacitor ICs, and worst-case numbers range from 0.4% to 1.5% (assuming, of course, that an accurate clock is provided). In order to achieve this kind of precision using passive or conventional active filter techniques requires the use of either very accurate resistors, capacitors, and sometimes inductors, or trimming of component values to reduce errors. It is possible for active or passive filter designs to achieve better accuracy than switched-capacitor circuits, but additional cost is the penalty. A resistor-programmed switched-capacitor filter circuit can be trimmed to achieve better accuracy when necessary, but again, there is a cost penalty.

Cost: No single technology is a clear winner here. If a single-pole filter is all that is needed, a passive RC network may be an ideal solution. For more complex designs, switched-capacitor filters can be very inexpensive to buy, and take up very little expensive circuit board space. When good accuracy is necessary, the passive components, especially the capacitors, used in the discrete approaches can be quite expensive; this is even more apparent in very compact designs that require surface-mount components. On the other hand, when speed and accuracy are not important concerns, some conventional active filters can be built quite cheaply.

Noise: Passive filters generate very little noise (just the thermal noise of the resistors), and conventional active filters generally have lower noise than switched-capacitor ICs. Switched-capacitor filters use active op amp-based integrators as their basic internal building blocks. The integrating capacitors used in these circuits must be very small in size, so their values must also be very small. The input resistors on these integrators must therefore be large in value in order to achieve useful time constants. Large resistors produce high levels of thermal noise voltage; typical output noise levels from switched-capacitor filters are on the order of 100 μV to 300 μVrms over a 20 kHz bandwidth. It is interesting to note that the integrator input resistors in switched-capacitor filters are made up of switches and capacitors, but they produce thermal noise the same as "real" resistors.

(Some published comparisons of switched-capacitor vs. op amp filter noise levels have used very noisy op amps in the op amp-based designs to show that the switched-capacitor filter noise levels are nearly as good as those of the op amp-based filters. However, filters with noise levels at least 20 dB below those of most switched-capacitor designs can be built using low-cost, low-noise op amps such as the LM833.)

Although switched-capacitor filters tend to have higher noise levels than conventional active filters, they still achieve dynamic ranges on the order of 80 dB to 90 dB—easily quiet enough for most applications, provided that the signal levels applied to the filter are large enough to keep the signals "out of the mud".

Thermal noise isn't the only unwanted quantity that switched-capacitor filters inject into the signal path. Since these are clocked devices, a portion of the clock waveform (on the order of 10 mV p–p) will make its way to the filter's output. In many cases, the clock frequency is high enough compared to the signal frequency that the clock feedthrough can be ignored, or at least filtered with a passive RC network at the output, but there are also applications that cannot tolerate this level of clock noise.

Offset Voltage: Passive filters have no inherent offset voltage. When a filter is built from op amps, resistors and capacitors, its offset voltage will be a simple function of the offset voltages of the op amps and the dc gains of the various filter stages. It's therefore not too difficult to build filters with sub-millivolt offsets using conventional techniques. Switched-capacitor filters have far larger offsets, usually ranging from a few millivolts to about 100 mV; there are some filters available with offsets over 1V! Obviously, switched-capacitor filters are inappropriate for applications requiring dc precision unless external circuitry is used to correct their offsets.

Frequency Range: A single switched-capacitor filter can cover a center frequency range from 0.1 Hz or less to 100 kHz or more. A passive circuit or an op amp/resistor/capacitor circuit can be designed to operate at very low frequencies, but it will require some very large, and probably expensive, reactive components. A fast operational amplifier is necessary if a conventional active filter is to work properly at 100 kHz or higher frequencies.

Tunability: Although a conventional active or passive filter can be designed to have virtually any center frequency that a switched-capacitor filter can have, it is very difficult to vary that center frequency without changing the values of several components. A switched-capacitor filter's center (or cutoff) frequency is proportional to a clock frequency and can therefore be easily varied over a range of 5 to 6 decades with no change in external circuitry. This can be an important advantage in applications that require multiple center frequencies.

Component Count/Circuit Board Area: The switched-capacitor approach wins easily in this category. The dedicated, single-function monolithic filters use no external components other than a clock, even for multipole transfer functions, while passive filters need a capacitor or inductor per pole, and conventional active approaches normally require at least one op amp, two resistors, and two capacitors per second-order filter. Resistor-programmable switched-capacitor devices generally need four resistors per second-order filter, but these usually take up less space than the components needed for the alternative approaches.

Aliasing: Switched-capacitor filters are sampled-data devices, and will therefore be susceptible to aliasing when the input signal contains frequencies higher than one-half the clock frequency. Whether this makes a difference in a particular application depends on the application itself. Most switched-capacitor filters have clock-to-center-frequency ratios of 50:1 or 100:1, so the frequencies at which aliasing begins to occur are 25 or 50 times the center frequencies. When there are no signals with appreciable amplitudes at frequencies higher than one-half the clock frequency, aliasing will not be a problem. In a low-pass or bandpass application, the presence of signals at frequencies nearly as high as the clock rate will often be acceptable because although these signals are aliased, they are reflected into the filter's stopband and are therefore attenuated by the filter.

When aliasing is a problem, it can sometimes be fixed by adding a simple, passive RC low-pass filter ahead of the switched-capacitor filter to remove some of the unwanted high-frequency signals. This is generally effective when the switched-capacitor filter is performing a low-pass or bandpass function, but it may not be practical with high-pass or notch filters because the passive anti-aliasing filter will reduce the passband width of the overall filter response.

Design Effort: Depending on system requirements, either type of filter can have an advantage in this category, but switched-capacitor filters are generally much easier to design. The easiest-to-use devices, such as the LMF40, require nothing more than a clock of the appropriate frequency. A very complex device like the LMF120 requires little more design effort than simply defining the desired performance characteristics. The more difficult design work is done by the manufacturer (with the aid of some specialized software). Even the universal, resistor-programmable filters like the LMF100 are relatively easy to design with. The procedure is made even more user-friendly by the availability of filter software from a number of vendors that will aid in the design of LMF100-type filters. National Semiconductor provides one such filter software package free of charge. The program allows the user to specify the filter's desired performance in terms of cutoff frequency, a passband ripple, stopband attenuation, etc., and then determines the required characteristics of the second-order sections that will be used to build the filter. It also computes the values of the external resistors and produces amplitude and phase vs. frequency data.

Where does it make sense to use a switched-capacitor filter and where would you be better off with a continuous filter? Let's look at a few types of applications:

Tone Detection (Communications, FAXs, Modems, Biomedical Instrumentation, Acoustical Instrumentation, ATE, etc.): Switched-capacitor filters are almost always the best choice here by virtue of their accurate center frequencies and small board space requirements.

Noise Rejection (Line-Frequency Notches for Biomedical Instrumentation and ATE, Low-Pass Noise Filtering for General Instrumentation, Anti-Alias Filtering for Data Acquisition Systems, etc.): All of these applications can be handled well in most cases by either switched-capacitor or conventional active filters. Switched-capacitor filters can run into trouble if the signal bandwidths are high enough relative to the center or cutoff frequencies to cause aliasing, or if the system requires dc precision. Aliasing problems can often be fixed easily with an external resistor and capacitor, but if dc precision is needed, it is usually best to go to a conventional active filter built with precision op amps.

Controllable, Variable Frequency Filtering (Spectrum Analysis, Multiple-Function Filters, Software-Controlled Signal Processors, etc.): Switched-capacitor filters excel in applications that require multiple center frequencies because their center frequencies are clock-controlled. Moreover, a single filter can cover a center frequency range of 5 decades. Adjusting the cutoff frequency of a continuous filter is much more difficult and requires either analog switches (suitable for a small number of center frequencies), voltage-controlled amplifiers (poor center frequency accuracy) or DACs (good accuracy over a very limited control range).

Audio Signal Processing (Tone Controls and Other Equalization, All-Pass Filtering, Active Crossover Networks, etc.): Switched-capacitor filters are usually too noisy for "high-fidelity" audio applications. With a typical dynamic range of about 80 dB to 90 dB, a switched-capacitor filter will usuallly give 60 dB to 70 dB signal-to-noise ratio (assuming 20 dB of headroom). Also, since audio filters usually need to handle three decades of signal frequencies at the same time, there is a possibility of aliasing problems. Continuous filters are a better choice for general audio use, although many communications systems have bandwidths and S/N ratios that are compatible with switched capacitor filters, and these systems can take advantage of the tunability and small size of monolithic filters.