Deliyannis, Theodore L. et al "The Approximation Problem"
Continuous-Time Active Filter Design
Boca Raton: CRC Press LLC,1999

Chapter 2

The Approximation Problem

2.1 Introduction

Solution of the approximation problem is a major step in the design procedure of a filter and is equally important in the design of both analog and digital filters. It is through the solution of this problem that the filter designer determines the filter function, the response which satisfies the specifications. Of course, the function obtained this way will satisfy the specifications only approximately and not exactly. However, if the specifications are set within the limitations of the LLF networks, the network realizing the approximating function will fulfil the requirements and thus be suitable for the task for which it is designed.

In this chapter, based on the contents of Chapter 1, we review first the characteristics of the *permitted* functions, and we formulate the approximation problem. Next, we present briefly the best known and most popular functions used in the solution of the approximation problem for the required filter response in the frequency domain. Then, since these functions are lowpass, we introduce suitable frequency transformations in order to obtain highpass, bandpass, or bandstop filters according to the requirements. Finally, we discuss the transformation of elements and the scaling of impedance level.

2.2 Filter Specifications and Permitted Functions

The knowledge gathered from the analysis of LLF networks, in the more general concept of the term *analysis*, which was explained in Section 1.4, can help in the search for the most suitable filter function to meet particular specifications. The results of this analysis impose three important constraints on the *permitted* LLF network functions. These have to be causal, rational, and stable.

Before proceeding to explain how to determine the filter function to meet a set of specifications, we review these constraints briefly.

2.2.1 Causality

In general, causality refers to the fact that there can be no result without cause. In the case of interest here, a causal network will not respond before an excitation has been applied to its terminals. Thus, the unit impulse response is zero for time $t < 0$. The response in Fig. 2.1(a) is not causal; therefore, it cannot be realized. On the other hand, that in Fig. 2.1(b) is causal, therefore realizable. Thus, the ideal lowpass filter is unrealizable, because its impulse response is noncausal.

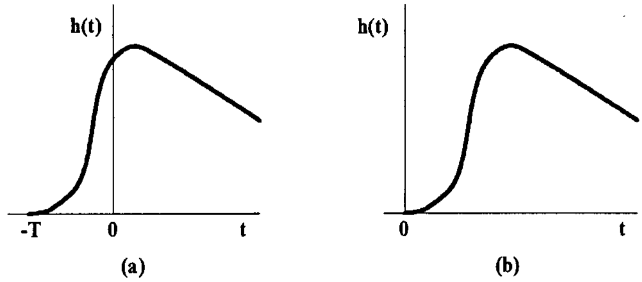

FIGURE 2.1
(a) Noncausal and (b) causal response.

In the frequency domain, causality is determined by means of the Paley-Wiener criterion [1]. Consider the impulse response $h(t)$, which possesses a Fourier Transform $H(j\omega)$ for which

$$\int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega < \infty \tag{2.1}$$

For $H(j\omega)$ to be causal, the criterion is the following:

$$\int_{-\infty}^{\infty} \frac{|\log|H(j\omega)||}{1+\omega^2} d\omega < \infty \tag{2.2}$$

Some consequences of this criterion are the following:

1. The magnitude function $|H(j\omega)|$ cannot be zero for a finite frequency band. However, it can be zero at a finite number of distinct frequencies.
2. The magnitude $|H(j\omega)|$ cannot decrease faster than exponentially.
3. Because of this constraint, the ideal filters are unrealizable.

2.2.2 Rational Functions

The LLF network functions are rational, i.e., ratios of two finite polynomials of the Laplace transform variable s. Therefore, it is not possible to realize the function e^{-sT} by such a network, because this function cannot be expressed in the form of a rational function.

2.2.3 Stability

The response of a stable network is bounded if the excitation is bounded. This means that, if $h(t)$ is the impulse response of the network, then

$$\int_{0}^{\infty} h(t)dt < \infty \tag{2.3}$$

and $\lim\ h(t) \to 0$ when $t \to \infty$.

In the frequency domain, stability implies that

1. the network function $H(s)$ does not have poles in the RH of the s-plane,
2. any poles on the $j\omega$-axis are simple, and
3. the degree of the numerator polynomial cannot be higher than the degree of the denominator by more than one.

However, for a filter to be useful, its function $H(s)$ has to be strictly stable, i.e., all its poles must be located in the LH of the s-plane excluding the $j\omega$-axis (poles at zero and infinity are considered to be located on the $j\omega$-axis).

2.3 Formulation of the Approximation Problem

In practice, the specifications of the filter may be given in terms of the cutoff frequency (or frequencies) ω_c, the maximum allowable deviation (error) A_{max} in the passband, the stop-band edges (frequencies), and the minimum attenuation A_{min} in the stopband. In the case of equalizers, it may be possible that the required frequency response is specified more closely. In general, from the specifications, we will be able to draw a frequency response magnitude plot, which will correspond to a prespecified curve. For example, for a lowpass filter, this diagram will be of the form shown in Fig. 2.2. The required response will have to lie between the limits set by the diagram.

Theoretically, the approximation problem is stated as follows:

1. Time domain:
   The impulse response $h(t)$ has to be approximated. An approximating function $h^*(t)$ is selected such that some error ε is minimal, where

$$\varepsilon = \int_0^{\infty} [h(t) - h^*(t)]^2 dt \tag{2.4}$$

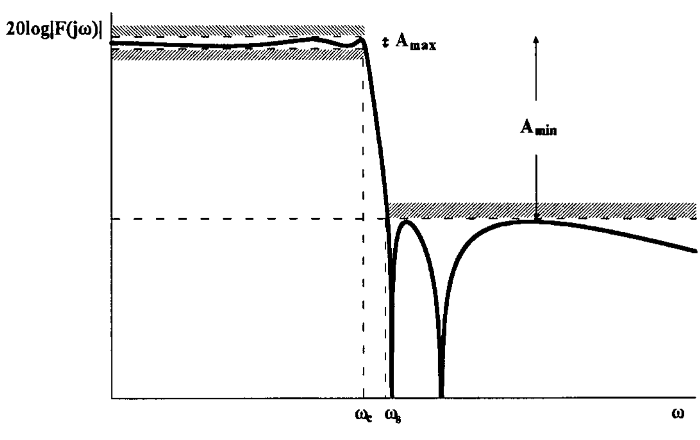

FIGURE 2.2
A possible magnitude response which satisfies the specifications of a lowpass filter.

Of course, $h^*(t)$ should be such that its Laplace transform $H(s)$ is a realizable function by a LLF network.

2. Frequency domain:
   In the frequency, domain we often work in terms of lowpass functions, because they are simpler, and because the highpass, bandpass, and bandstop responses can be obtained from lowpass responses by means of suitable frequency transformations.

   Our problem here is to find a function $F(s)$ the magnitude and/or phase response of which approximates the prespecified curve according to a predetermined criterion.

The approximation problem has been solved mathematically in various ways. In the case of magnitude approximation, the best known and most popular lowpass functions are the following: the Butterworth or maximally flat, the Chebyshev (Tschebycheff) or equiripple, the monotonic or Papoulis, and the Cauer or elliptic function filters. Of course, with the use of a computer, one may create one's own approximating functions, particularly in the cases of arbitrary responses of filters and equalizers. In such cases, techniques employing linear segments, curve fitting, pole-zero placements, etc., have proved very useful in solving the approximation problem.

In the case of delay approximation, best known functions are the Bessel-Thomson filters, the Padè approximates, both maximally flat, and those of Chebyshev type.

In what follows, we introduce briefly the best known and practically useful approximations to the ideal lowpass filter and to the ideal delay.

2.4 Approximation of the Ideal Lowpass Filter

In practical filter design, the amplitude response is more often specified than the phase response. The amplitude response of the ideal lowpass filter with normalized cutoff frequency at $\omega_c = 1$ is shown in Fig. 2.3. As has already been explained in Section 2.2., this ideal amplitude response cannot be expressed as a rational function of s. It is thus unrealizable. If we accept a small error in the passband and a non-zero transition band, we may seek a rational function $F(s)$, the magnitude of which will approximate the ideal response as closely as possible. A suitable magnitude function can be of the form

$$|F(j\omega)| = M(\omega) = \frac{1}{[1+\varepsilon^2 w(\omega^2)]^{1/2}} \tag{2.5}$$

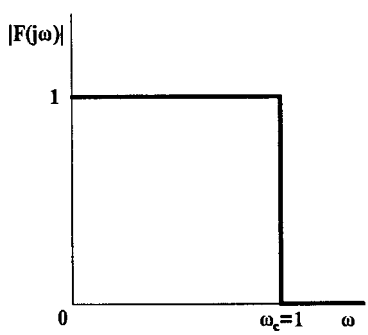

FIGURE 2.3
Ideal lowpass filter amplitude response.

where ε is a constant between zero and one ($0 < \varepsilon \leq 1$), according to the accepted passband error, and $w(\omega^2)$ is a function of ω^2 such that

$$0 \leq w(\omega^2) \leq 1 \qquad 0 \leq \omega \leq 1$$

and which increases very fast with increasing ω, for $\omega > 1$, remaining much greater than one outside the passband.

In general, the numerator of $M(\omega)$ may be a constant other than unity, which will influence the gain (or attenuation) at $\omega = 0$ (at dc).

In the following, we review the most popular functions $w(\omega^2)$ and the corresponding $F(s)$, the magnitude of which approximate the amplitude response of the ideal lowpass filter.

2.4.1 Butterworth or Maximally Flat Approximation

If we let $\varepsilon = 1$ and

$$w(\omega^2) = \omega^{2n}$$

in Eq. (2.5), n being a positive real integer, we will get the following amplitude function:

$$M(\omega) = \frac{1}{[1+\omega^{2n}]^{1/2}} \tag{2.6}$$

It can be seen that

$$M(0) = 1$$

while $M(\omega)$ decreases monotonically with increasing ω.

At $\omega_c = 1$,

$$M(1) = \frac{1}{\sqrt{2}} = 0.707$$

or

$$20\log M(1) = -10\log 2 = -3.01$$

In other words, at $\omega_c = 1$, the amplitude is 3 dB below its value at dc. This is the cutoff frequency of the filter. Clearly, this is independent of n, the order of the filter function, which in fact determines how close to the ideal is the approximating function $M(\omega)$, i.e., how successful the approximation is.

Equation (2.6) for different n gives the amplitude response of the various Butterworth filter functions. The Butterworth approximation is also called the maximally flat approximation, because the first $2n - 1$ derivatives of $M(\omega)$, the maximum number in Eq. (2.2), are zero at $\omega = 0$. The error in the passband is zero at $\omega = 0$ and maximal (3 dB) at cutoff. Between $\omega = 0$ and $\omega = 1$, the error takes intermediate values increasing monotonically from the zero value with increasing ω. For values of $\omega >> 1$, $M(\omega)$ behaves approximately as

$$M(\omega) = \frac{1}{\omega^n}$$

i.e., it changes asymptotically as

$$20\log M(\omega) = -20n\ \log\omega \tag{2.7}$$

or, in other words, it falls off by $6n$ dB/octave ($20n$ dB/decade).

We now seek to obtain the network function $F(s)$ whose magnitude with $s = j\omega$ is $M(\omega)$. We proceed as follows. Observing that

$$M^2(\omega) = |F(j\omega)|^2 = F(j\omega)F(-j\omega) = \frac{1}{1+\omega^{2n}} \tag{2.8}$$

we may write

$$F(s)F(-s)|_{s=j\omega} = \left.\frac{1}{1+(-1)^n s^{2n}}\right|_{s=j\omega}$$

By a process known as analytic continuation, it turns out that we may remove the $s = j\omega$ constraint and write

$$F(s)F(-s) = \frac{1}{1+(-1)^n s^{2n}} \tag{2.9}$$

Now define a function $P(s^2)$ such that

$$P(s^2) = F(s)F(-s) \tag{2.10}$$

when we will also have that

$$M^2(\omega) = P(-\omega^2)$$

Thus, knowing $P(-\omega^2)$ through $M^2(\omega)$, we can obtain $P(s^2)$ by setting s^2 for $-\omega^2$ in Eq. (2.8). Then, expressing $P(s^2)$ in the form of Eq. (2.10), we observe that the poles of $F(s)$ are symmetrical to those of $F(-s)$ about the $j\omega$-axis. Since $F(s)$ has to be a stable function, we identify its poles as those of $P(s^2)$ with negative real part.

The poles of $P(s^2)$ are the roots of the equation

$$1 + (-1)^n s^{2n} = 0 \tag{2.11}$$

It can be shown that the solution of Eq. (2.11) is the following:

$$s_k = \sigma_k + j\omega_k = -\sin\left(\frac{2k-1}{2n}\pi\right) + j\cos\left(\frac{2k-1}{2n}\pi\right) \tag{2.12}$$

for $k = 1, 2, \ldots, 2n$.

The n poles of $F(s)$ are those obtained from Eq. (2.12) with $k = 1, 2, \ldots, n$. All poles have magnitude equal to unity and lie on the circumference of the unit circle equally spaced.

As an example, consider the case for $n = 4$. We have

$$M^2(\omega) = \frac{1}{1 + \omega^8}$$

Then

$$P(s^2) = \frac{1}{1 + s^8}$$

The poles of $P(s^2)$, are found from Eq. (2.12) for $k = 1, 2, \ldots, 8$. These lie on the circumference of the unit circle as shown in Fig. 2.4.

Of these, the first four ($k = 1, 2, 3$, and 4) will be assigned to $F(s)$, since they have to lie on the LH of the s-plane. They are as follows:

$$s_1 = -0.3827 + j0.9239$$

$$s_2 = -0.9239 + j0.3827$$

$$s_3 = -0.9239 - j0.3827$$

$$s_4 = -0.3827 - j0.9239$$

Therefore, the fourth-order Butterworth lowpass function will be

$$F(s) = \frac{1}{(s - s_1)(s - s_2)(s - s_3)(s - s_4)}$$

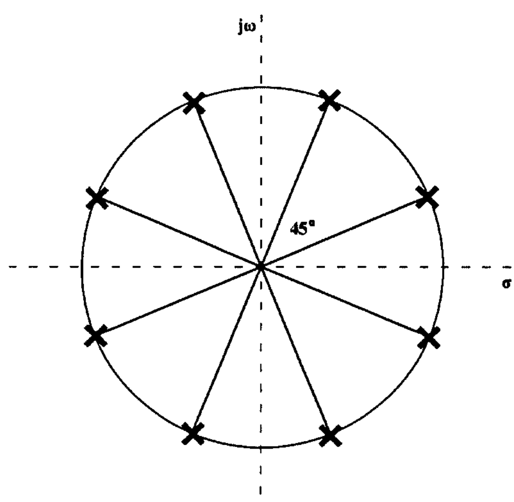

FIGURE 2.4
Poles of the Butterworth filter for n = 4.

Grouping complex conjugate pole terms, it turns out that

$$F(s) = \frac{1}{(s^2 + 0.7654s + 1)(s^2 + 1.8478s + 1)}$$

or multiplying out in full,

$$F(s) = \frac{1}{s^4 + 2.613s^3 + 3.414s^2 + 2.613s + 1}$$

The denominator of this last expression is known as a Butterworth polynomial. The first ten Butterworth polynomials in two forms are given on Table A.1 at the end of the book for easy reference. Finally, in Fig. 2.5 the magnitude response of the third-order Butterworth filter is shown together with the responses of two other filters we are dealing with next for comparison.

2.4.2 Chebyshev or Equiripple Approximation

In this case, Eq. (2.5) takes the following form:

$$|F(j\omega)|^2 = M^2(\omega) = \frac{1}{1 + \varepsilon^2 C_n^2(\omega)} \tag{2.13}$$

Here again, $0 < \varepsilon \leq 1$, and $C_n(\omega)$ is a Chebyshev polynomial of degree n having the following form:

$$\begin{aligned} C_n(\omega) &= \cos(n \cos^{-1}\omega) \qquad 0 \leq |\omega| \leq 1 \\ &= \cosh(n \cosh^{-1}\omega) \qquad 1 \leq |\omega| \end{aligned} \tag{2.14}$$

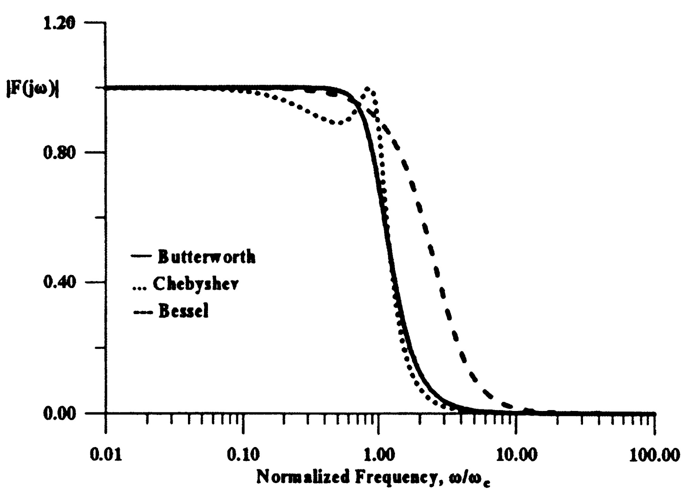

FIGURE 2.5
Magnitude responses of the Butterworth third-order lowpass filter and the corresponding Chebyshev (1 dB ripple) and Bessel filters.

Clearly, $C_n(\omega)$ varies between +1 and –1 in the passband ($0 \le \omega \le 1$), while its absolute value increases rapidly with ω above ω = 1. Consequently, $M(\omega)$ varies between 1 and $(1 + \varepsilon^2)^{-1/2}$ in the passband having an oscillatory or ripple error of

$$20\log(1+\varepsilon^2)^{1/2} = 10\log(1+\varepsilon^2) \ \text{dB}$$

Thus, the accepted error in the passband determines the value of ε.

The Chebyshev polynomials can be obtained by the recursion formula

$$C_{n+1}(\omega) = 2\omega C_n(\omega) - C_{n-1}(\omega) \tag{2.15}$$

with

$$C_0(\omega) = 1$$

$$C_1(\omega) = \omega$$

A plot of the Chebyshev polynomials with n = 1, 2, and 3 is given in Fig. 2.6.

Therefore, at dc (ω = 0), we will have

$$M(0) = \begin{cases} 1 & n \text{ odd} \\ (1+\varepsilon^2)^{-1/2} & n \text{ even} \end{cases}$$

Outside the passband and for ω >> 1, $M(\omega)$ behaves approximately like $(\varepsilon\, 2^{n-1}\omega^n)^{-1}$, i.e., the attenuation for ω >> 1 will be

$$\begin{aligned} 20\log(\varepsilon\, 2^{n-1}\omega^n) &= 20\log\varepsilon + 20\log 2^{n-1} + 20\log\omega^n \\ &= 20\log\varepsilon + 6(n-1) + 20n\ \log\omega \quad \text{dB} \end{aligned} \tag{2.16}$$

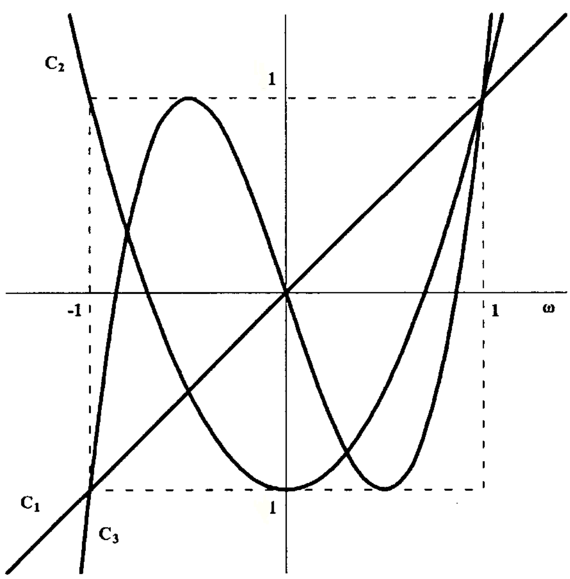

FIGURE 2.6
Plot of Chebyshev polynomials of degrees n = 1, 2, and 3.

compared to $20\,n \log \omega$ in the corresponding Butterworth function. Thus, for $\varepsilon > 1$, the Chebyshev approximation has an advantage of $20 \log \varepsilon + 6(n - 1)$ dB over the Butterworth approximation. However when $\varepsilon < 1$, this advantage becomes less significant, since then $\log \varepsilon$ will be negative.

By a similar procedure to the Butterworth case, the poles of the Chebyshev filter functions can be shown to be as follows:

$$s_k = \sigma_k \pm j\omega_k$$

where

$$\sigma_k = \sinh \beta_k \cdot \sin\left(\frac{2k-1}{2n}\pi\right) \tag{2.17}$$

$$\omega_k = \cosh \beta_k \cos\left(\frac{2k-1}{2n}\pi\right) \tag{2.18}$$

with

$$\beta_k = \frac{1}{n}\sinh^{-1}\frac{1}{\varepsilon} \qquad k = 1, 2, \ldots, 2n$$

These poles lie on an ellipse defined by the following equation:

$$\frac{\sigma_k^2}{\sinh^2 \beta_k} + \frac{\omega_k^2}{\cosh^2 \beta_k} = 1 \tag{2.19}$$

The major semi-axis of the ellipse falls on the $j\omega$-axis, its length being $\pm \cosh \beta_k$, whereas the length of the minor semiaxis is $\pm \sinh \beta_k$.

The points of intersection of the ellipse and the $j\omega$-axis define the –3 dB frequencies (half-power frequencies), which are thus equal to $\pm \cosh \beta_k$. The corresponding Butterworth frequencies are always $\omega_c = \pm 1$.

We may normalize the poles of the Chebyshev functions in order to have the half-power frequencies (–3 dB frequencies) appearing at $\omega_c = 1$, by dividing s_k by $\cosh \beta_k$. Then, the normalized poles s_k will be as follows:

$$s'_k = \frac{s_k}{\cosh \beta_k} = \sigma'_k \pm j\omega'_k$$

with

$$\sigma'_k = \tanh \beta_k \sin\left(\frac{2k-1}{2n}\pi\right) \tag{2.20}$$

$$\omega'_k = \cos\left(\frac{2k-1}{2n}\pi\right) \tag{2.21}$$

Comparing s'_k with the corresponding poles of the Butterworth functions, it can be seen that they have the same imaginary parts, whereas their real parts differ by the factor $\tanh\beta_k$. The relative locations on the s-plane of the Butterworth and normalized Chebyshev function poles are given in Fig. 2.7 for $n = 3$. For $\varepsilon = 0$, when $\beta_k = \infty$ and $\tanh\beta_k = 1$, the poles of the Butterworth and Chebyshev functions coincide.

The coefficients of the Chebyshev filter functions, as well as their poles, have been tabulated for various ripple values i.e., 0.1, 0.5, 1, …, 3 dB. A sample of such a tabulation is given on Table A.2 in Appendix A. This table does not give the normalized Chebyshev filter functions. In all these cases, the upper edge of the passband ripple occurs at $\omega_c = 1$.

The amplitude responses for the 1 dB ripple and the 3 dB ripple third-order Chebyshev lowpass functions

$$F_{1dB}(s) = \frac{0.491}{s^3 + 0.988s^2 + 1.238s + 0.491}$$

and

$$F_{3dB}(s) = \frac{0.2506}{s^3 + 0.597s^2 + 0.928s + 0.2506}$$

are shown in Fig. 2.8 for comparison with the response of the corresponding Butterworth filter. It can be seen that the Chebyshev filters have equiripple response in the passband and fall off monotonically outside it.

2.4.3 Inverse Chebyshev Approximation

The Chebyshev polynomials are also used to obtain the so-called Inverse Chebyshev filter functions, the magnitude of which is given as follows:

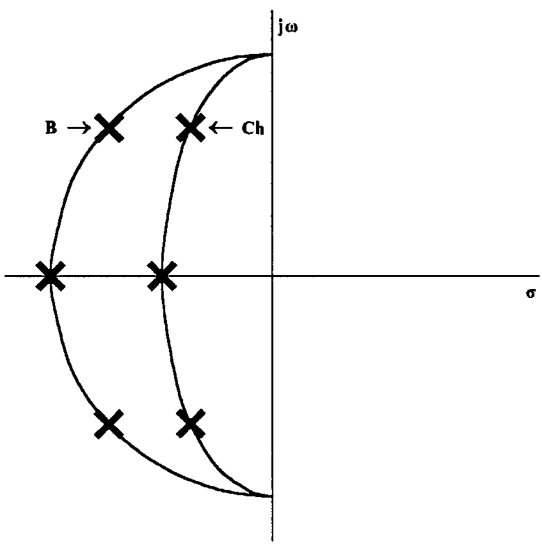

FIGURE 2.7
Relative positions of Butterworth and normalized Chebyshev poles.

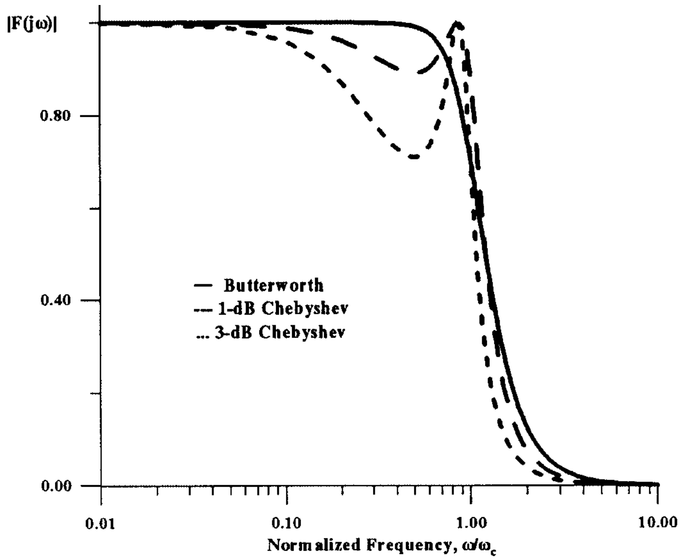

FIGURE 2.8
Comparison of the magnitude response of 1 dB and 3 dB ripple Chebyshev third-order lowpass filters and with the corresponding Butterworth filter.

$$M^2(\omega) = \frac{\varepsilon^2 C_n^2\left(\frac{1}{\omega}\right)}{1 + \varepsilon^2 C_n^2\left(\frac{1}{\omega}\right)} \tag{2.22}$$

The properties of these functions are complementary to those of the Chebyshev functions in the sense that they present maximum flatness in the passband and equiripple behavior in the stopband. Also, their phase response and consequently their group delay is better than that of the Chebyshev filter. In Fig. 2.9, the magnitude response of the third-order Inverse Chebyshev and the corresponding Chebyshev function are shown for comparison.

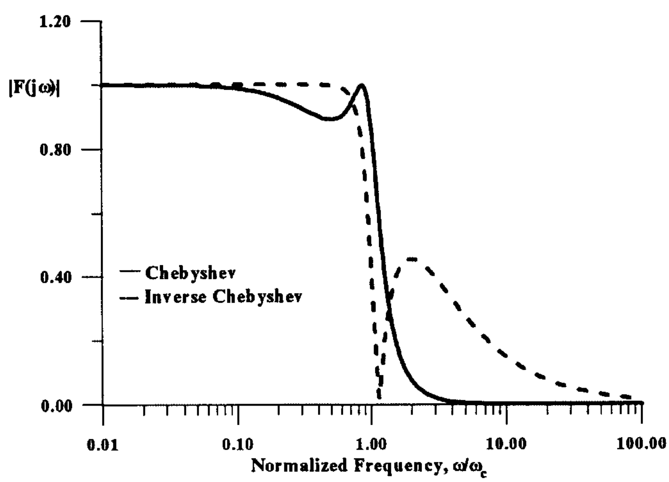

FIGURE 2.9
Magnitude response of the third-order Chebyshev (1 dB ripple) and the corresponding Inverse Chebyshev functions.

2.4.4 Papoulis Approximation

Filter L or Papoulis functions approximating the ideal lowpass filter response are obtained from Eq. (2.5) if we let $\varepsilon = 1$ and

$$w(\omega^2) = L_n(\omega^2) \tag{2.23}$$

with $L_n(\omega^2)$ having the following properties:

a. $L_n(0) = 0$

b. $L_n(1) = 1$

c. $\dfrac{dL_n(\omega^2)}{d\omega} \geq 0$

d. $\left.\dfrac{dL_n(\omega^2)}{d\omega}\right|_{\omega = 1} = \max$

Property c secures monotonicity in the amplitude response, whereas property *d* that the fall off rate at cutoff ($\omega = 1$) is the greatest possible, if monotonicity is assumed.

$L_n(\omega^2)$ polynomials are related to the Legendre $P_k(x)$ polynomials of the first kind. Some of them are given on Table 2.1.

TABLE 2.1

$L_n(\omega^2)$ Polynomials

n	$L_n(\omega^2)$
2	ω^4
3	$3\omega^6 - 3\omega^4 + \omega^2$
4	$6\omega^8 - 8\omega^6 + 3\omega^4$
5	$20\omega^{10} - 40\omega^8 + 28\omega^6 - 8\omega^4 + \omega^2$

The corresponding filters are known as Legendre, Class L, or Papoulis filters [2]. Their poles are found by the procedure that was followed in the case of Butterworth filters.

The main characteristics of these filters are the following:

- Their amplitude response is monotonic.
- The falloff rate at cutoff is the greatest, assuming monotonicity.
- All of their zeros are at infinity.

Because L-filters are less sharp than Chebyshev filters, they are not as popular. However, in cases where the ripple in the passband is undesirable and so the use of Chebyshev filters is excluded, they could be preferable to Butterworth because of their steepest slope at cutoff.

2.4.5 Elliptic Function or Cauer Approximation

The filters examined so far have, except for the Inverse Chebyshev, all of their zeros at infinity. However, in some cases, a higher falloff rate is required in the transition band; in other words, a very high attenuation is required very near the cutoff frequency. This requirement mandates the use of elliptic functions in the approximation, thus obtaining the elliptic functions or, simply, elliptic or Cauer filters.

These filters display equiripple behavior both in the passband and the stopband. The typical magnitude response of a third-order elliptic filter is shown in Fig. 2.10, corresponding to the following general filter function $F(s)$.

$$F(s) = \frac{K(s^2 + \omega_o^2)}{(s + \alpha)(s^2 + \beta s + \gamma)} \tag{2.24}$$

The characteristic quantities that determine the elliptic filter specifications are the maximum passband error, given as maximum attenuation A_{max} in the passband, the minimum attenuation A_{min} in the stopband, the frequency ω_s at which the stopband starts, and the passband edge or cutoff frequency ω_c.

In the case of the elliptic filters, Eq. (2.5) is written in the following form:

$$|F(j\omega)|^2 = \frac{1}{1 + \varepsilon^2 R_n(\omega^2)} \tag{2.25}$$

where R_n, depending on whether n is odd or even, is either

$$R_n(\omega) = \frac{\omega(\omega_1^2 - \omega^2)(\omega_2^2 - \omega^2)\dots(\omega_k^2 - \omega^2)}{(1 - \omega_1^2\omega^2)(1 - \omega_2^2\omega^2)\dots(1 - \omega_k^2\omega^2)} \quad n \text{ odd } (n = 2k + 1) \tag{2.26}$$

or

$$R_n(\omega) = \frac{(\omega_1^2 - \omega^2)(\omega_2^2 - \omega^2)\dots(\omega_k^2 - \omega^2)}{(1 - \omega_1^2\omega^2)(1 - \omega_2^2\omega^2)\dots(1 - \omega_k^2\omega^2)} \quad n \text{ even } (n = 2k) \tag{2.27}$$

It can be seen from Eqs. (2.26) and (2.27) that

$$R_n\left(\frac{1}{\omega}\right) = \frac{1}{R_n(\omega)} \tag{2.28}$$

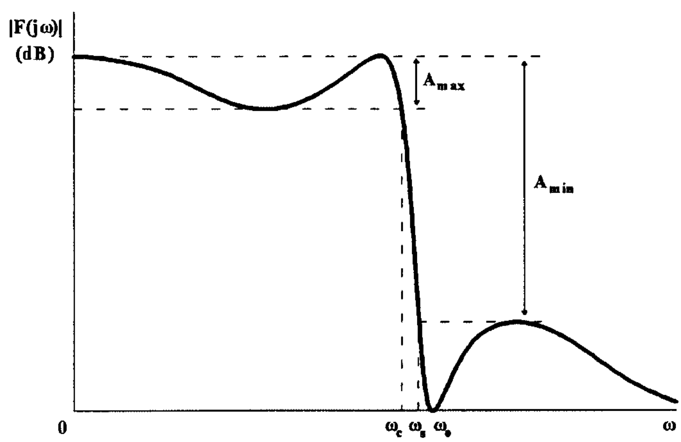

FIGURE 2.10
Typical magnitude response of an elliptic filter of third order.

The meaning of this is that the value of $R_n(\omega)$ at a frequency ω' in the band $0 \le \omega \le 1$ is the reciprocal of its value at the frequency $1/\omega'$ in the $1 < \omega < \infty$ frequency band. Therefore, if the critical frequencies could be found that lead to equiripple behavior in the passband, automatically the function will have equiripple behavior in the stopband also. Since $|F|^2$ is bounded, the poles of $F(s)$ cannot lie on the $j\omega$-axis. Also, since $|F(j\omega)|^2$ cannot be zero inside the passband, its zeros should lie outside the passband. However, the zeros of $|F(j\omega)|^2$ are the poles of $R_n(\omega)$. Therefore, all the poles of $R_n(\omega)$ should be greater than unity. This means that the zeros of $R_n(\omega)$ should all lie in the band $0 \le \omega < 1$.

The poles, zeros, and frequencies ω_s have been tabulated for various combinations of values of A_{max} and A_{min}. In such tables, the ripple in the passband is usually given in terms of the reflection coefficient ρ, which is related to A_{max} as follows:

$$A_{max} = -10\log(1-\rho^2)\ \text{dB} \tag{2.29}$$

It must be stressed that, with A_{max}, A_{min}, ω_c, and ω_s known, the solution of the approximation problem by means of the elliptic filters requires the lowest-order function and, therefore, it can be realized with the lowest cost. For this reason, the elliptic filters are used most often in practice.

Some elliptic filter functions are given in the Appendix A in Table A.3 for a certain value of maximum attenuation A_{max} in the passband. Clearly, for the selection of the suitable elliptic filter, the specifications should include the values of A_{max}, A_{min}, $\Omega_s(\omega_s/\omega_c)$, and n. In contrast, only the filter order n is required in the case of the Butterworth filter, whereas in the case of the Chebyshev filter the values of n and ε (or A_{max}) should be given. It is very important for the reader to know that, given A_{min}, A_{max}, $\Omega_s(\omega_s/\omega_c)$, the required value of n can be quickly determined for the Butterworth, Chebyshev, and Cauer filters from corresponding nomograms [4].

2.4.6 Selecting the Filter from Its Specifications

In the table giving the Butterworth filter functions, it is assumed that the cutoff frequency is normalized to unity, i.e., $\Omega_c = 1$. The suitable filter function can be read off this table, if its order n has been determined from the specifications. Thus, if the desired rate of fall in the transition band is $6N$ dB/octave, because n is an integer, we select $n = N$ if N is an integer; otherwise, the value of n will be equal to the nearest integer greater than N.

However, in some cases the filter specifications may be given differently. Let us suppose, in the more general case, that the filter specifications require the maximum attenuation in the passband to be A_{max} (<3 dB), occurring at ω_p rad/s (not normalized), and that beyond the frequency ω_s (rad/s) the minimum attenuation should be A_s (in dB). In such cases, the determination of the Butterworth filter order n can proceed as follows.

We suppose that the 3 dB (cutoff) frequency is ω_c rad/s, which of course corresponds to the normalized cutoff frequency Ω_c. Then the normalized frequencies Ω_p and Ω_s will be the following:

$$\Omega_p = \frac{\omega_p}{\omega_c} \qquad \Omega_s = \frac{\omega_s}{\omega_c} \tag{2.30}$$

From Eq. (2.6), with $A = -20\log M(\omega)$, we have

$$A_p = 20\log(1+\Omega_p^{2n})^{1/2} = 10\log(1+\Omega_p^{2n}) \tag{2.31}$$

from which, solving for Ω_p^{2n}, we obtain

$$\Omega_p^{2n} = 10^{0.1A_p} - 1 \tag{2.32}$$

Similarly, we will have for Ω_s

$$\Omega_s^{2n} = 10^{0.1A_s} - 1 \tag{2.33}$$

Dividing Eq. (2.33) by Eq. (2.32) gives

$$\left(\frac{\Omega_s}{\Omega_p}\right)^{2n} = \frac{10^{0.1A_s} - 1}{10^{0.1A_p} - 1} \tag{2.34}$$

But, because of Eq. (2.30),

$$\frac{\Omega_s}{\Omega_p} = \frac{\omega_s}{\omega_p} \tag{2.35}$$

Then, substituting ω_s/ω_p for Ω_s/Ω_p in Eq. (2.34), we obtain the following required value of n:

$$n = \frac{1}{2}\log\left(\frac{10^{0.1A_s} - 1}{10^{0.1A_p} - 1}\right)\left[\log\left(\frac{\omega_s}{\omega_p}\right)\right]^{-1} \tag{2.36}$$

It should be pointed out that this value of n will not be necessarily an integer. Then, the order of the required Butterworth filter function will be equal to the nearest integer greater than this value.

Next, we must determine the value of ω_c. Substituting the value of n that was found in Eq. (2.32) [or in Eq. (2.33)], the actual value of Ω_p (or Ω_s) is determined and, using Eq. (2.30), the value of ω_c is obtained. In general, the value of ω_c which is obtained based on the value of Ω_p will be different from that obtained based on Ω_s. However, any one of these values of ω_c will satisfy the filter specifications. The same is true if we use the mean of these two values of ω_c.

To demonstrate, this let the filter specifications be the following:

$$f_p = 3 \text{ kHz}, \qquad A_p \le 1 \text{ dB}, \qquad f_s = 6 \text{ kHz}, \qquad A_s \ge 20 \text{ dB}$$

Substituting in Eq. (2.36) gives

$$n = 4.3$$

We select n to be the next integer value, i.e., 5.

Substituting for n in Eqs. (2.32) and (2.33), we find the following values for Ω_p and Ω_s:

$$\Omega_p = 0.87361$$

$$\Omega_s = 1.5833$$

Then, from Eq. (2.30), for $\Omega_p = 0.87361$, we get one value of ω_c.

$$\omega_{c_p} = 2\pi \times 3.434 \text{ krad/s}$$

and for $\Omega_s = 1.5833$ another

$$\omega_{c_s} = 2\pi \times 3.78955 \text{ krad/s}$$

We may choose to consider as the ω_c value the mean of ω_{c_p} and ω_{c_s}, namely

$$\omega_c = \frac{\omega_{c_p} + \omega_{c_s}}{2} = 2\pi \times 3.61148 \text{ krad/s}$$

Either of ω_{c_p}, ω_{c_s}, or ω_c can be used as the required cutoff frequency ω_c. To show this, we calculate the values of A_p and A_s by means of Eqs. (2.32) and (2.33), which correspond to each of these three values of ω_c. Results are given in Table 2.2.

TABLE 2.2

f_c (kHz)	A_p (dB)	A_s (dB)
3.434	1	24.25
3.78955	0.401	20
3.61148	0.63	22.1

Clearly, the specifications are satisfied in all three cases.

Let us now consider the selection of the Chebyshev filter satisfying certain specifications. Since $C_n(1) = 1$ for any integer n, we will have from Eq. (2.13)

$$20\log[1 + \varepsilon^2 C_n^2(1)]^{1/2} = 10\log(1 + \varepsilon^2) = A_p \tag{2.37}$$

from which

$$\varepsilon = \sqrt{10^{0.1A_p} - 1} \tag{2.38}$$

For the previous example, substituting for $A_p (= A_{max}) = 1$ dB Eq. (2.38) gives

$$\varepsilon = 0.505$$

Considering that at Ω_s the filter function behaves approximately as $(\varepsilon 2^{n-1}\Omega^n)^{-1}$ we obtain

$$A_s \cong 20\log\varepsilon + 6(n-1) + 20n\log\Omega \tag{2.39}$$

In the case of the present example, $\Omega = \Omega_s = 2$ since $\omega_c = \omega_p$. Then, Eq. (2.39) gives

$$20 = 20\log 0.505 + 6(n-1) + 20n\log 2$$

from which, solving for n, we get $n = 2.69$. Therefore, the required order of Chebyshev filter will be $n = 3$, which is lower than the required order $n = 5$ of the Butterworth filter.

The case of the Cauer filter is much simpler, since n can be obtained straight from tables, given the specifications ω_p, A_p, ω_s, and A_s, again using ω_p as ω_c. From such tables [3], or the corresponding nomogram [4] in the case of the previous example, we obtain $n = 3$. Since n is also 3 in the Chebyshev case, we prefer to realize the Chebyshev function, since the cost will be lower, as we shall see in later chapters. It should be mentioned, however, that in practical filter design, in which the fall-off rate is higher than that in this example, the order of the Chebyshev filter is always higher than the order of the corresponding Cauer filter, and the most economical filter will be the Cauer filter.

2.4.7 Amplitude Equalization

In some cases, the amplitude response of the practical filter may not match the amplitude response of the function it realizes. This is because the performance of its components is not ideal. To avoid the subsequent distortion of the signal when passing through the filter, it is necessary that its amplitude response be corrected or, in an other word, equalized.

It is obvious that the transfer function of the equalizer cannot be selected from a predetermined set of functions, since it depends uniquely on the individual filter response that requires equalization. Once the equalizer response has been deduced from the difference between the expected response and the "actual" filter response, the latter obtained by simulation of the filter on the computer using non-ideal components, a function approximating the equalizer response should be found. This can be achieved by applying curve-fitting techniques and an optimization program, while care should be taken in order that the resulting function will be realizable (permitted function). Another approach would start with a certain pole and zero placement and use then the optimization program to adjust their locations until the required response is obtained.

A practical approach [6] suitable in the case of passive filters is to use a cascade of simple networks, e.g., the constant-resistance bridge-T network, the amplitude response of which can be relatively easily adjusted. By properly selecting the component values of the sections, the overall response of the cascaded sections can be adjusted to match the required equalizer response.

2.5 Filters with Linear Phase: Delays

As was explained in Section 1.6.2, if

$$\varphi(\omega) \equiv \arg H(j\omega)$$

we define the group delay τ_g as

$$\tau_g \equiv -\frac{d\varphi(\omega)}{d\omega} \tag{2.40}$$

whereas

$$-\frac{\varphi(\omega)}{\omega} = \tau_p \tag{2.41}$$

is the phase delay.

It can be shown [5] that the definition of the group delay has a physical meaning only when (a) the magnitude function varies slowly with frequency, and (b) the phase varies nearly linearly with frequency over the band of interest.

If $H(s)$ is a rational function, the same will be true for the group delay τ_g, while the phase delay τ_p will not be a rational function.

The function

$$H(s) = e^{-sT}$$

has linear phase, since

$$\varphi(\omega) \equiv \arg\ e^{-j\omega T} = -\omega T$$

and represents the pure delay T, since

$$-\frac{d\varphi}{d\omega} = T$$

However, e^{-sT} is not a rational function of s.

Thus, although it can be realized by a lossless transmission line terminated at both ends in its characteristic impedance, it cannot be realized by an LLF network.

We can approximate e^{-sT} though by a rational function having either all its zeros at infinity (polyonimic) or in the RH of the s-plane (non-minimum phase function).

This approximation can be achieved either by way of approximating the linear phase $-\omega T$ [Fig. 2.11(a)] or by way of approximating the group delay T [Fig. 2.11(b)], as was indicated in the case of the ideal lowpass filter. Some useful delay approximation functions are briefly reviewed below.

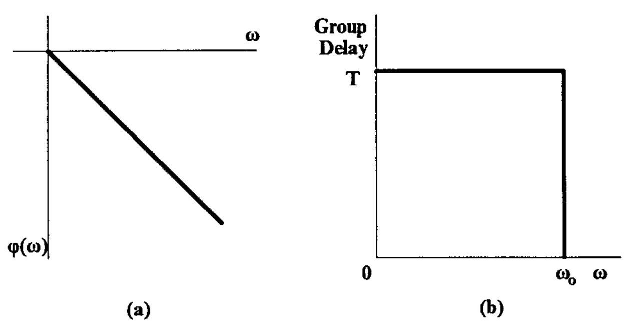

FIGURE 2.11
(a) Ideal linear phase and (b) group delay response.

2.5.1 Bessel-Thomson Delay Approximation

The approximation of the constant group delay T normalized to unity can be obtained by a procedure similar to that followed in the case of magnitude response. Let us select the approximation function $F(s)$ to be polyonimic, i.e., of the form

$$F(s) = \frac{K}{D_n(s)} \tag{2.42}$$

where K is a constant, and $D_n(s)$ a polynomial with positive constant coefficients of order n.

From $F(s)$ we obtain the phase function

$$\varphi(\omega) = \arg F(j\omega)$$

Next, we perform a Taylor expansion of $\varphi(\omega)$ about $\omega = 0$ and take the derivative with respect to ω, which we equate to the negative of the group delay $T = 1$. Equating then coefficients of equal powers of ω, we obtain a number of equations equal to the desired order of approximation n. Clearly, only the constant term of $-d\varphi/d\omega$ is equated to 1. All the other coefficients are set equal to zero. Solution of this set of equations will give the values of the n coefficients of $D_n(s)$. The value of K in Eq. (2.42) is equal to the constant term of $D_n(s)$, normalizing thus the magnitude of $F(j\omega)$ at $\omega = 0$ to unity.

As an example, consider the case $n = 2$. Let

$$D_2(s) = s^2 + \alpha s + \beta$$

Then,

$$\varphi(\omega) = \arg F(j\omega) = -\tan^{-1}\frac{\alpha\omega}{\beta - \omega^2}$$

Taylor's expansion of $\varphi(\omega)$, assuming that $y(\omega) = |\alpha\omega/(\beta - \omega^2)| < 1$, is

$$\varphi(\omega) = -\left[\frac{\alpha\omega}{\beta - \omega^2} - \frac{\alpha^3\omega^3}{3(\beta - \omega^2)^3} + \dots\right]$$

Taking the derivative of $\varphi(\omega)$ with respect to ω of the first two terms in the series, and ignoring the rest for ω such that $\omega << 1$ gives

$$-\frac{d\varphi(\omega)}{d\omega} = \frac{\alpha\beta + \alpha\omega^2}{(\beta - \omega^2)^2} - \frac{1}{3}\left[\frac{3\alpha^3\omega^2(\beta - \omega^2)^3 + 6\alpha^3\omega^4(\beta - \omega^2)^2}{(\beta - \omega^2)^6}\right]$$

or

$$-\frac{d\varphi(\omega)}{d\omega} = \frac{(\alpha\beta + \alpha\omega^2)(\beta - \omega^2)^2 - \alpha^3\omega^2(\beta - \omega^2) - 2\alpha^3\omega^4}{(\beta - \omega^2)^4}$$

Equating this to 1 and multiplying through by $(\beta - \omega^2)^4$ gives the following equation:

$$(\alpha\beta + \alpha\omega^2)(\beta - \omega^2)^2 - (\alpha^3\omega^2(\beta - \omega^2) - 2\alpha^3\omega^4) = (\beta - \omega^2)^4$$

Now we equate the constant term on the one side of this equation to the constant term on the other side and obtain

$$\alpha\beta^3 = \beta^4$$

or

$$\alpha = \beta$$

We do the same for the coefficients of ω^2 and get

$$\alpha\beta^2 + \alpha^3\beta = 4\beta^3$$

But, since $\alpha = \beta$, this equation gives $\beta = 3$. Therefore, the second-order delay function will be the following:

$$F_2(s) = \frac{3}{s^2 + 3s + 3}$$

Following this approach, it is found [5] that the polynomials $D_n(s)$ are related to the Bessel polynomials $G_n(s)$ of degree n by the following relationship:

$$D_n(s) = G_n\left(\frac{1}{s}\right)s^n \qquad (2.43)$$

These Bessel polynomials are defined as follows:

$$G_n\left(\frac{1}{s}\right) \equiv \sum_{k=0}^{n} \frac{(n+k)!}{(n-k)!k!(2s)^k} \qquad (2.44)$$

It can be shown that all of $D_n(s)$ zeros are located in the LH of the s-plane, and there exists at most one zero on the negative real semi-axis.

The first two polynomials and the recursion formula for obtaining $D_n(s)$ of any degree n are as follows:

$$D_o = 1$$

$$D_1(s) = s + 1$$

$$D_n(s) = (2n - 1)D_{n-1}(s) + s^2 D_{n-2}(s) \qquad (2.45)$$

The first 5 $D_n(s)$ polynomials and their roots are given in Table 2.3.

The delay functions $F(s)$ obtained this way are called Bessel or Thomson filters, and they approximate the ideal delay according to the maximally flat criterion. Their amplitude

TABLE 2.3

$D_n(s)$ Polynomials and Their Roots

n	$D_n(s)$	Roots of $D_n(s)$
1	$s + 1$	-1
2	$s^2 + 3s + 3$	$-1.5 \pm j0.867$
3	$s^3 + 6s^2 + 15s + 15$	-2.322, $-1.839 \pm j1.754$
4	$s^4 + 10s^3 + 45s^2 + 105s + 105$	$-2.896 \pm j0.867$, $-2.104 \pm j2.657$
5	$s^5 + 15s^4 + 105s^3 + 420s^2 + 945s + 945$	-3.647, $-3.352 \pm j1.743$, $-2.325 \pm j3.571$

response is lowpass with a cutoff frequency depending on the value of n and given by the following approximate formula (for $n \geq 3$):

$$\omega_{3dB} = \sqrt{(2n-1)\ln 2} \tag{2.46}$$

This can be easily seen in Fig. 2.12(a), showing the magnitude response of the first three Bessel filters of odd orders.

The corresponding phase response plots are shown in Fig. 2.12(b). It can be seen that the bandwidth with nearly linear phase also increases with increasing n.

In Fig. 2.5 the amplitude, and in Fig. 2.13 the phase response, of the third-order Bessel filter are shown along with the corresponding responses of the Butterworth and the 1 dB ripple Chebyshev filters of the same order. It can be seen that, from the selectivity point of

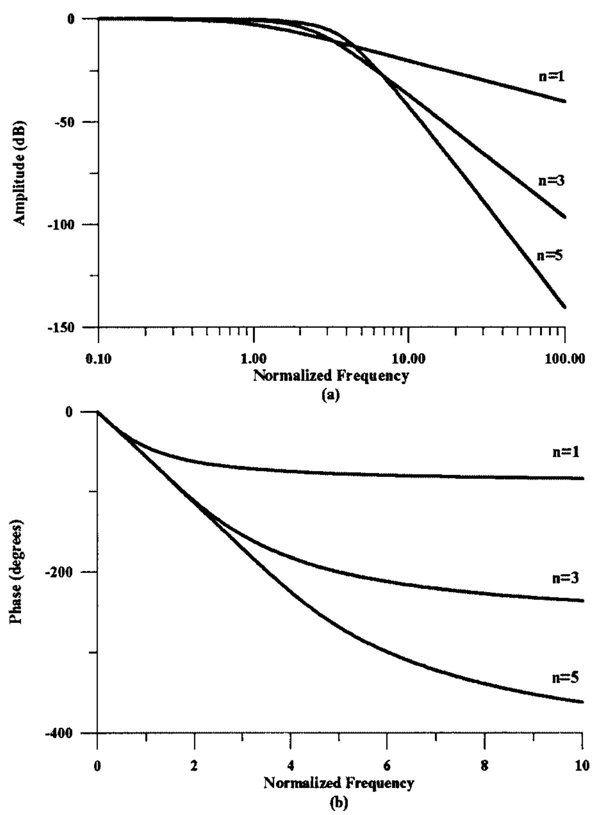

FIGURE 2.12
(a) Magnitude and (b) phase response of the first three Bessel filters of odd orders.

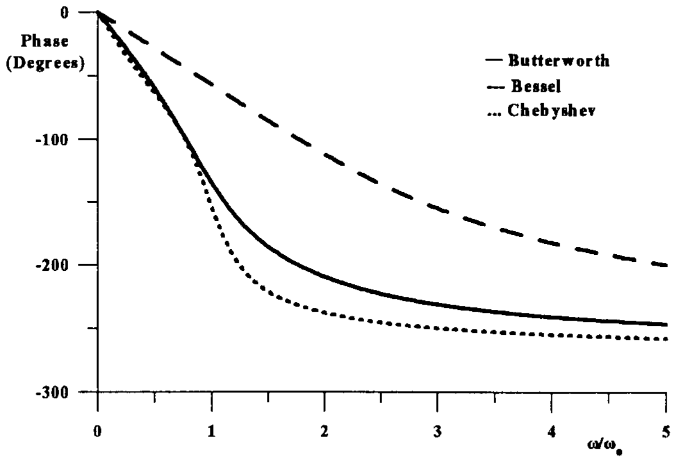

FIGURE 2.13
Phase response of Bessel, Butterworth, and Chebyshev (1 dB ripple) third-order filters.

view, the Bessel filter is at a disadvantage, but its phase response, as far as linearity is concerned, is by far superior—particularly when compared to the Chebyshev phase response.

As a consequence of their superior phase response, the time response of the Bessel-Thomson filters also displays superior performance concerning fidelity to the input waveforms over the other lowpass filters. In other words, they transmit, for example, square pulses with lower distortion than the other filters. This can be easily seen in Fig. 2.14, where the step response of the three filters considered above is shown. Clearly, the rise time, the settling time, and the overshoot are lower for the Bessel filter than for the other two.

It can also be seen that the Bessel response rises to 50 percent of its final value at $t_d = 1s$, which is equal to the unit (1s) delay it approximates. This justifies the characterization of t_d as the delay time.

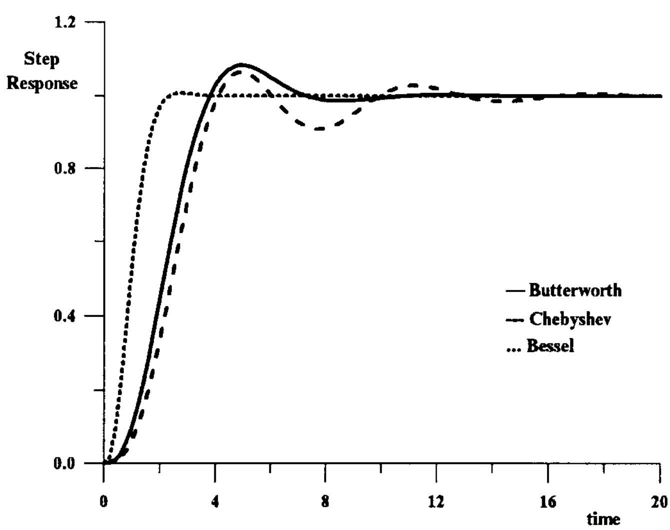

FIGURE 2.14
Step response of the Bessel, Butterworth, and Chebyshev (1 dB ripple) third-order filters.

This observation is nearly true for the corresponding response of the Bessel filters of any order n. However, as n increases, the rise time becomes shorter, and thus the step response comes nearer to the undistorted step. This is so because, as n increases, more frequencies of the infinite spectrum of the input step fall within the bandwidth of nearly linear phase, and thus the approximation comes closer to the input.

The reduced rise time with increasing n can also be proved, if the magnitude response of the Bessel filters is considered. Clearly, from Eq. (2.46) and the plots in Fig. 2.12(a), the 3 dB bandwidth increases with increasing n. Then, from Eq. (1.33), it follows that the rise time should decrease with increasing n.

The bandwidth ω_o of the maximally flat delay is defined [4] as the reciprocal of the delay at $\omega = 0$, i.e.,

$$\omega_o = \frac{1}{T} \tag{2.47}$$

Since the normalized delay is 1, ω_o will also be 1, i.e.

$$\omega_o = 1 \tag{2.48}$$

Clearly, the meaning of Eq. (2.47) is that the product bandwidth times delay is constant. That is, large bandwidth corresponds to short delay and vice versa.

Letting $s = ju$, where

$$u = \omega T = \frac{\omega}{\omega_o}$$

we can create two tables, the first giving the u values (for each n) for certain deviations of the time delay from its ideal value (i.e., its value at $\omega = 0$), and the second giving the values of u (for each n) in which the attenuation is certain decibels below its value at $\omega = 0$. From these tables, the designer can select the value of n and consequently determine the corresponding Bessel-Thomson delay function that suits best the specifications (see Reference 3).

2.5.2 Other Delay Functions

Another class of functions approximating in fact the phase of e^{-sT} according to the maximally flat criterion at $\omega = 0$ are the allpass Padè approximations [6]. All of the zeros of these functions lie on the RHP located symmetrically to the poles with respect to the $j\omega$-axis. The magnitude response of these is unity for all ω and their useful bandwidth is twice that of the Bessel-Thomson delays of corresponding orders. However, in spite of these useful characteristics, their step response displays a very narrow precursor of height about equal to their final value, a highly undesirable characteristic (see Fig. 2.15). To avoid the appearance of this precursor in the step response one may use lowpass Padè delay approximations [7] or other more useful delay functions [8, 9, 10]. Other delay functions which approximate the group delay according to the Chebyshev criterion have also been proposed [11]. These display improved characteristics over the Padè delay functions of corresponding orders.

In general, the function that will be selected for delaying a signal will depend greatly on the type of signal. Thus, if the signal is in the form of a step, a lowpass delay is more suitable than an allpass. In the case of a signal with a certain bandwidth though, an allpass function with linear phase may satisfy the specifications more effectively. In practice, on

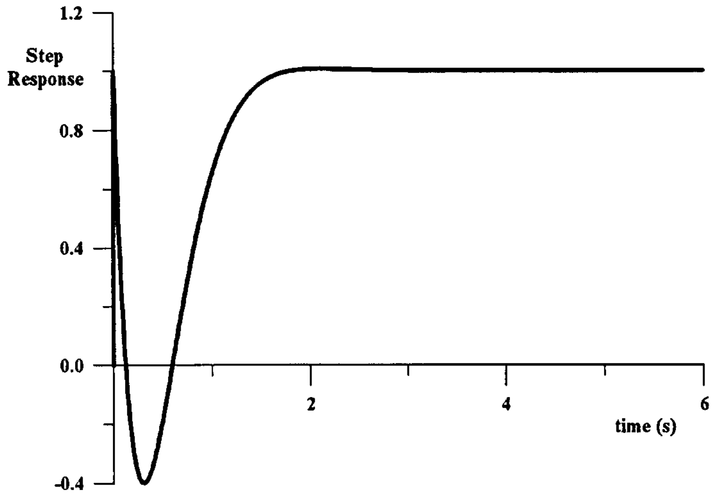

FIGURE 2.15
Step response of the second-order allpass Padè delay function.

many occasions, the combination of a lowpass filter with sharp cutoff and an allpass function filter connected in cascade results in the desirable solution, as explained next.

2.5.3 Delay Equalization

As it was mentioned in the previous section, the step response of the Bessel-Thomson filters, due to their linear phase response, makes them more suitable in pulse transmission than the corresponding Butterworth, Chebyshev or Cauer filters. However, from the selectivity point of view their performance is very poor compared to the other filters.

To achieve both phase linearity and good selectivity in the amplitude response, a practical solution is to use suitable allpass functions of second-order in order to modify the phase of the filter that has the desirable magnitude response. Their use will not affect the filter magnitude response, since they are allpass. Such functions will be of the form

$$F(s) = \frac{s^2 - \beta s + \gamma}{s^2 + \beta s + \gamma} \tag{2.49}$$

They will be selected by means of a computer optimization program, which will determine the most suitable coefficient values β and γ in each case.

This procedure, called phase equalization, proves to be very useful in problems where the required filter should possess high selectivity and at the same time linear phase response, i.e., constant group delay.

2.6 Frequency Transformations

The filter functions that were reviewed in the previous sections refer to lowpass filters. They are given in normalized form, i.e., their passband width is unity. In all these tables, the normalized frequency s_n is implied for sinusoidal excitation $s_n = j\Omega$ with

$$\Omega = \frac{\omega}{\omega_c}$$

where ω is the real frequency variable, and ω_c is the actual cutoff frequency of the desired lowpass filter. Following this convention, the normalized cutoff frequency Ω_c of all filters is

$$\Omega_c = 1$$

We will now show by means of suitable transformations how we can obtain denormalized lowpass, highpass, bandpass, bandstop filters, and delays using data obtained from the tabulated normalized lowpass functions. In all cases, we refer to frequency response, when $s = j\omega$.

2.6.1 Lowpass-to-Lowpass Transformation

In the normalized lowpass function, if we substitute s/ω_c for s_n, we will obtain the denormalized lowpass function with ω_c being its cutoff frequency. For example, for the first-order Butterworth function

$$F(s_n) = \frac{1}{s_n + 1} \tag{2.50}$$

we will get

$$F(s) = \frac{1}{s/\omega_c + 1} = \frac{\omega_c}{s + \omega_c} \tag{2.51}$$

which is also lowpass with its cutoff frequency at ω_c.

Clearly, with this transformation, the normalized frequency band $0 \le |\Omega| \le 1$ is transformed to the denormalized frequency band

$$0 \le |\omega| \le \omega_c$$

as shown in Fig. 2.16.

As can be seen from Eqs. (2.50) and (2.51) the shape of the frequency response does not change with this transformation. Following this, it is obvious that transmission zeros at Ω_1 will appear at the frequency $\omega_1 = \Omega_1\omega_c$ in the denormalized magnitude response. This is shown clearly in Fig. 2.17.

Finally, it should be mentioned that the gain of the normalized filter functions is assumed to be normalized, i.e., its maximum value is equal to unity. It is usual that in filter design we are not so interested in the actual value of the magnitude in the corresponding response, but in its relative value (or relative attenuation), which determines the filter selectivity.

FIGURE 2.16
Lowpass-to-lowpass transformation.

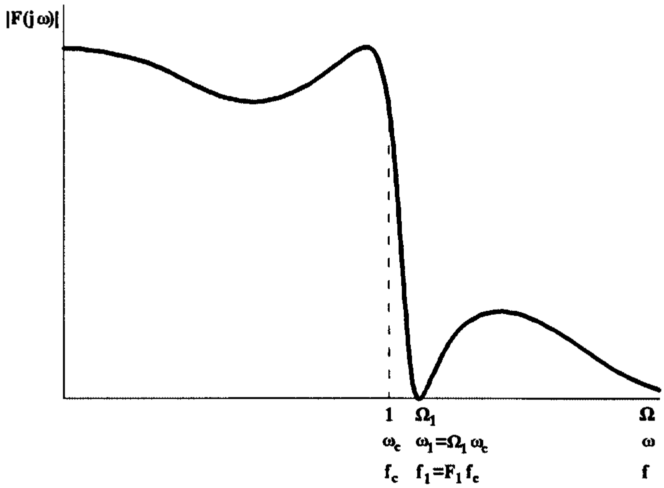

FIGURE 2.17
Lowpass-to-lowpass frequency transformation.

2.6.2 Lowpass-to-Highpass Transformation

Applying the transformation

$$s_n \rightarrow \frac{1}{s}$$

a lowpass function is transformed to a highpass.

The frequencies 0 and ∞ of the lowpass function are transformed to ∞ and 0, respectively, while the cutoff frequency of the lowpass, which is 1 in the normalized function, is transformed to itself in the new function. Thus, the passband of the lowpass $0 \le |\Omega| \le 1$ is transformed to the passband $0 \le |\Omega| \le \infty$ of the highpass function as shown in Fig. 2.18(a).

Following the same argument as in the case of the lowpass-to-lowpass transformation, if we substitute ω_c/s for s_n in Eq. (2.50) we will get

$$F(s) = \frac{s}{s+\omega_c} \tag{2.52}$$

which is a highpass function with ω_c its cutoff frequency. The mapping of the lowpass passband $0 \le |\Omega| \le 1$ to the highpass passband $\omega_c \le |\omega| \le \infty$ is shown pictorially in Fig. 2.18(b).

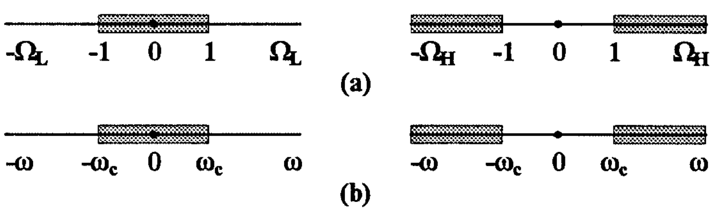

FIGURE 2.18
Lowpass-to-highpass transformation.

2.6.3 Lowpass-to-Bandpass Transformation

Applying the transformation

$$s_n \to s + \frac{1}{s} = \frac{s^2 + 1}{s}$$

a lowpass function is transformed to a bandpass function with a passband width equal to that of the lowpass, i.e., equal to 1. There are two bandpass cutoff frequencies, Ω_{c1} and Ω_{c2}, such that

$$\Omega_{c1}\Omega_{c2} = 1$$

and

$$\Omega_o^2 = \Omega_{c1}\Omega_{c2} = 1 \qquad (2.53)$$

Ω_o is called the normalized center frequency of the bandpass function. The passband mapping is shown in Fig. 2.19(a).

The bandwidth of the passband is

$$\Omega_{c2} - \Omega_{c1} = 1 \qquad (2.54)$$

since this has to be equal to that of the lowpass. Solving Eqs. (2.53) and (2.54) for Ω_{c1} and Ω_{c2} gives the following:

$$|\Omega_{c1}| = -\frac{1}{2} + \frac{\sqrt{5}}{2}$$

$$|\Omega_{c2}| = \frac{1}{2} + \frac{\sqrt{5}}{2} \qquad (2.55)$$

Using this transformation in the example we considered before, the lowpass Butterworth function [Eq. (2.50)] will be transformed to the function

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

which is bandpass, since it becomes zero at $\Omega = 0$ and ∞.

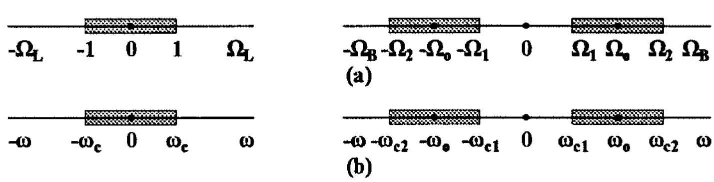

FIGURE 2.19
Lowpass-to-bandpass transformation.

Denormalization of the bandpass function to a center frequency ω_o is obtained by substituting ω/ω_o for Ω in Eqs. (2.53) to (2.55). The bandwidth of the denormalized function will then be

$$B = \omega_{c2} - \omega_{c1} = \omega_o$$

If the bandwidth should be other than ω_o, the transformation has to be modified. Thus, we may obtain the denormalized bandpass function straight from the normalized lowpass function by applying to the latter the following transformation:

$$s_n \rightarrow \frac{\omega_o}{B}\left(\frac{s}{\omega_o} + \frac{\omega_o}{s}\right)$$

Corresponding mapping of the passbands and stopbands are shown in Fig. 2.19(b). In this case,

$$\omega_o^2 = \omega_{c1}\omega_{c2}$$

$$B = \omega_{c2} - \omega_{c1} \tag{2.56}$$

2.6.4 Lowpass-to-Bandstop Transformation

Applying the transformation

$$s_n \rightarrow \frac{1}{s + \frac{1}{s}} = \frac{s}{s^2 + 1}$$

a normalized lowpass function is transformed to a bandstop (bandreject) function, the stopband of which is between the normalized frequencies Ω_{c2} and Ω_{c1} as shown in Fig. 2.20(a).

To obtain the required denormalized bandstop function from the normalized lowpass function, the suitable transformation is

$$s_n \rightarrow \frac{B}{\omega_o\left(\frac{s}{\omega_o} + \frac{\omega_o}{s}\right)}$$

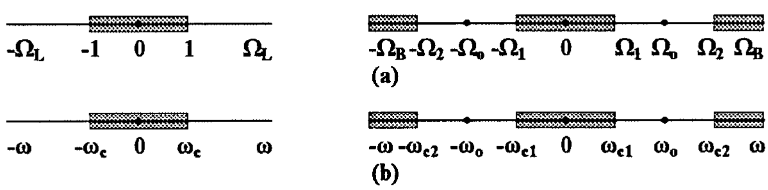

FIGURE 2.20
Lowpass-to-bandstop transformation.

where ω_o is the centre frequency in the stopband with

$$\omega_o^2 = \omega_{c1}\omega_{c2}$$

and B is the bandwidth of the stopband

$$B = \omega_{c2} - \omega_{c1}$$

This is shown in Fig. 2.20(b).

In the case of the example considered above, we will obtain the normalized bandstop function

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

This is zero at $\Omega = 1$ and 1 at $\Omega = 0$ and ∞, which shows the bandstop behavior of the function.

2.6.5 Delay Denormalization

As we have seen in Section 2.5, the functions that approximate the ideal delay are also given in tables or in the form of recursion formulas. The denormalized function is obtained from the normalized one by substituting $s\tau$ for s_n, where τ is the required delay in seconds.

2.7 Design Tables for Passive LC Ladder Filters

All filter functions introduced in this chapter can be realized by passive networks. There is an abundance of books and papers in the literature describing how this can be done. However, in this book, we are interested only in the realization using doubly terminated LC ladders, since their simulation by active RC networks results in low sensitivity filters (see Chapter 6).

The general structure of such a network is shown in Fig. 2.21. In the case of lowpass polyonimic filters (all zeros at infinity), all Z_is are inductors, and all Y_is are capacitors. In the case of lowpass functions with finite transmission zeros (e.g., Cauer filters), the Z_is will be parallel tuned LC subcircuits, or the Y_is will be series tuned LC combinations.

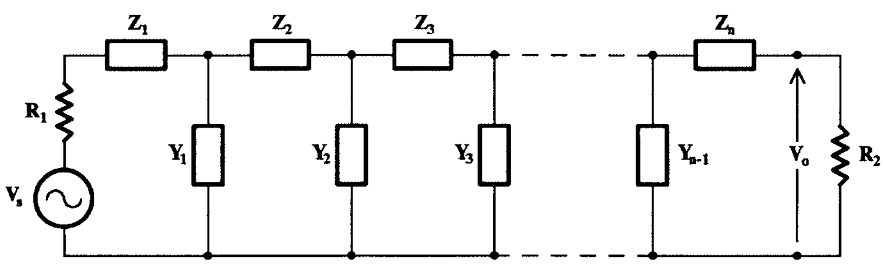

FIGURE 2.21
The general form of a doubly terminated ladder.

Design tables for realizing all the lowpass functions that we have introduced above for various values of A_{max} (or ε), A_{min}, and orders, as well as for various values of the ratio R_2/R_1 are given in many books [4, 12, 13, 14]. Although, from the sensitivity point of view, as we shall see later, the equally terminated ladder ($R_1 = R_2$) is most desirable, in some cases (e.g., even-order Chebyshev and Cauer filters), this is not exactly possible.

These tables are used by the passive filter designer and, of course, by the active RC filter designer who will choose to design a filter by simulating the passive ladder by active elements. Therefore, once the designer has chosen the normalized lowpass filter function from the specifications of his problem, he may use these tables to obtain the passive circuit, which realizes this function. Then, he can obtain the required denormalized filter by applying suitable transformations to the element values as explained below.

As we shall see in later Chapters (e.g., Chapter 6), one powerful method for designing RC active filters is to simulate passive LC ladder filters, either topologically or functionally, using RC active circuits. Thus, the tables of LC passive filters greatly simplify this design and are used accordingly. It is for this very reason that we have introduced these tables for the design of LC ladder filters here.

However, for various reasons, there are no corresponding tables available for the design of RC active filters in general use. For example, an RC active cannot be transformed to a corresponding bandpass or bandstop circuit by the transformation of its elements as in the case of the passive LC circuits, while available tables [15] do not cover all useful active RC circuits and Cauer filter design.

One final point here: the lowpass filter function realized by the circuit in Fig. 2.21 as the voltage ratio V_o/V_s is in fact realized within a constant multiplier that is lower than unity. The reason for this is that at dc ($\omega = 0$), the transfer voltage ratio V_o/V_s reduces to

$$\frac{V_o}{V_s} = \frac{R_2}{R_1 + R_2}$$

which is always less than 1 except for $R_1 = 0$, when it is equal to 1.

2.7.1 Transformation of Elements

In the case of passive filters, as stated above, one can obtain the denormalized highpass, bandpass, or bandstop filters by applying the previously introduced frequency transformations to the impedances of the elements of the normalized lowpass filter.

This approach is not applicable in the case of active RC filters except for the case of obtaining a highpass from the normalized lowpass filter. We examine the element transformation in more detail below:

2.7.1.1 LC Filters

TABLE 2.4

Lowpass-to-lowpass	Transformation $s_n \to \frac{s}{\omega_c}$ with ω_c the cutoff frequency	
Element	Impedance	New element value
L_n	$s_n L_n = \frac{s}{\omega_c} L_n$	$L = L_n/\omega_c$
C_n	$\frac{1}{s_n C_n} = \frac{\omega_c}{s C_n}$	$C = C_n/\omega_c$

Lowpass-to-highpass	Transformation $s_n \to \frac{\omega_c}{s}$ with ω_c the cutoff frequency	
Element	Impedance	New element value
L_n	$s_n L_n = \frac{\omega_c L_n}{s}$	$C = \frac{1}{\omega_c L_n}$
C_n	$\frac{1}{s_n C_n} = \frac{s}{\omega_c C_n}$	$L = \frac{1}{\omega_c C_n}$
Lowpass-to-bandpass	Transformation $s_n \to \frac{\omega_o}{B}\left(\frac{s}{\omega_o} + \frac{\omega_o}{s}\right)$	
Element	Impedance	New element values
L_n	$s_n L_n = \frac{\omega_o}{B}\left(\frac{s}{\omega_o} + \frac{\omega_o}{s}\right) L_n$	L and C in series $L = \frac{L_n}{B}$ $C = \frac{B}{\omega_o^2 L_n}$
C_n	$\frac{1}{s_n C_n} = \frac{1}{C_n \frac{\omega_o}{B}\left(\frac{s}{\omega_o} + \frac{\omega_o}{s}\right)}$	L and C in parallel $L = \frac{B}{C_n \omega_o^2}$ $C = \frac{C_n}{B}$
Lowpass-to-bandstop	Transformation $s_n \to \frac{1}{\frac{\omega_o}{B}\left(\frac{s}{\omega_o} + \frac{\omega_o}{s}\right)}$	
Element	Impedance	New element values
L_n	$s_n L_n = \frac{L_n}{\frac{\omega_o}{B}\left(\frac{s}{\omega_o} + \frac{\omega_o}{s}\right)}$	L and C in parallel $L = \frac{B L_n}{\omega_o^2}$ $C = \frac{1}{B L_n}$
C_n	$\frac{1}{s_n C_n} = \frac{\frac{\omega_o}{B}\left[\frac{s}{\omega_o} + \frac{\omega_o}{s}\right]}{C_n}$	L and C in series $L = \frac{1}{B C_n}$ $C = \frac{B C_n}{\omega_o^2}$

All these element transformations are summarized in Table 2.5.

TABLE 2.5

Element Transformations

Elements of lowpass filter	Corresponding elements of the denormalized Highpass	Bandpass	Bandstop
L_n	$C = 1/(L_n \omega_c)$	$L = L_n/B$ $C = B/(\omega_o^2 L_n)$	$L = L_n B/\omega_o^2$ $C = 1/(B L_n)$
C_n	$L = 1/(C_n \omega_c)$	$L = B/(\omega_o^2 C_n)$ $C = C_n/B$	$L = 1/(B C_n)$ $C = C_n B/\omega_o^2$

As an example, consider the design of a bandpass filter having center frequency at 1 krad/s and 100 rad/s bandwidth. Assume that, from additional specifications, a sixth-order Butterworth bandpass filter has been found suitable.

Clearly, the sixth-order bandpass will be obtained from the third-order lowpass Butterworth filter by applying the lowpass-to-bandpass transformation

$$s_n \rightarrow \frac{\omega_o}{B}\left(\frac{s}{\omega_o} + \frac{\omega_o}{s}\right)$$

where $\omega_o = 1$ krad/s and B = 100 rad/s.

From the tabulated Butterworth filters, Table A.1, we find

$$F(s_n) = \frac{1}{s_n^3 + 2s_n^2 + 2s_n + 1}$$

At this point, we may proceed in one of the following two alternative ways:

1. We may apply the lowpass-to-bandpass transformation to obtain either the normalized or the denormalized bandpass function and proceed to realize it, i.e., to determine the circuit.
2. Alternatively, we may realize the lowpass filter and apply the lowpass-to-denormalized bandpass transformation to the elements of the lowpass making use of the Table 2.5.

The second method has the advantage that the realization of the lowpass filter reduces to choosing the circuit from the available design tables. In cases, when there are no available design tables, the advantage still remains, since the realization of the lowpass filter is simpler than that of the bandpass, because the order of the latter is double that of the lowpass. On the other hand, the second method is not applicable to the synthesis of active RC networks. So, the choice of the most suitable method will depend on the type of the circuit (passive LC or active RC) we choose to design.

Suppose we choose to realize $F(s_n)$ by a passive LC equally terminated ladder. Using the corresponding table, given for example in Reference 12, we find that the suitable circuit is that appearing in Fig. 2.22.

Using Table 2.5, we can easily obtain the element denormalization for $\omega_o = 1$ krad/s and $B = 100$ rad/s. The denormalized w.r.t. frequency bandpass circuit is as shown in Fig. 2.23.

This circuit is still normalized w.r.t. impedance level, since all component values are referred to terminating resistances of 1 Ω. If we want to raise the impedance level to a practical value, e.g., 600 Ω, we should multiply the impedance of each component by 600, when we obtain the component values in parentheses in Fig. 2.23. We treat impedance denormalization in Section 2.7.

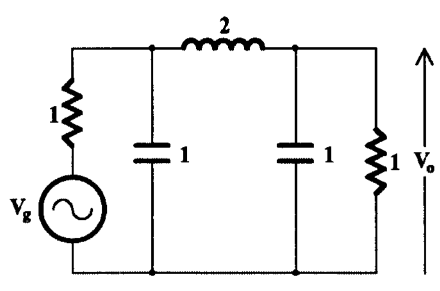

FIGURE 2.22
Butterworth normalized lowpass filter of third-order.

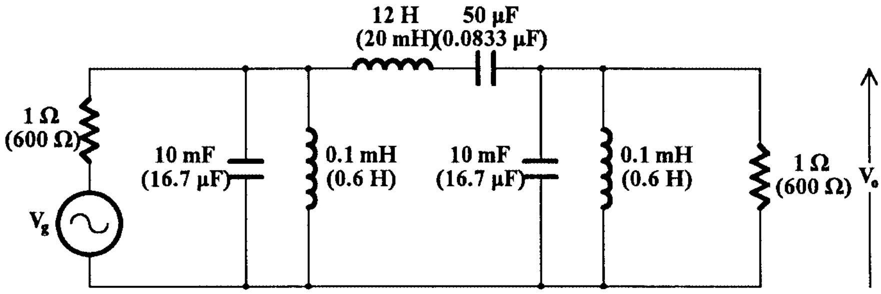

FIGURE 2.23
Frequency denormalized bandpass filter.

2.7.1.2 Active RC Filters

The procedure described above can be applied to obtain the denormalized lowpass or highpass RC active filter once the normalized lowpass function has been selected.

It is not possible to obtain the normalized or denormalized bandpass and bandstop circuits straight from the normalized lowpass, because there is no suitable transformation for this purpose. Clearly with no inductances in these circuits, it is impossible to apply the element transformations for bandpass and bandstop of Table 2.5. In this case, first the normalized bandpass or bandstop function is obtained using the corresponding frequency transformation. Next, the normalized filter is synthesized by a suitable method, as we shall see in later chapters, and then the denormalized bandpass or bandstop filter is obtained by properly scaling the filter time constants.

It has to be emphasized here that frequency transformation must be applied only to time constants, i.e., either to the capacitances or to those resistances that determine the time constants and not to those that determine the gain of the active element.

The following two examples will clarify this, while for a more formal proof the interested reader should refer to References 6 and 16.

Consider first the simple RC circuit in Fig. 2.24(a). The transfer voltage ratio V_2/V_1 is the following:

$$F(s) = \frac{V_2}{V_1} = \frac{1}{RCs+1} = \frac{1}{RC}\cdot\frac{1}{s+1/RC} \tag{2.57}$$

If we interchange the position of the elements without changing its topology, as shown in Fig. 2.24(b), the new transfer voltage ratio will be

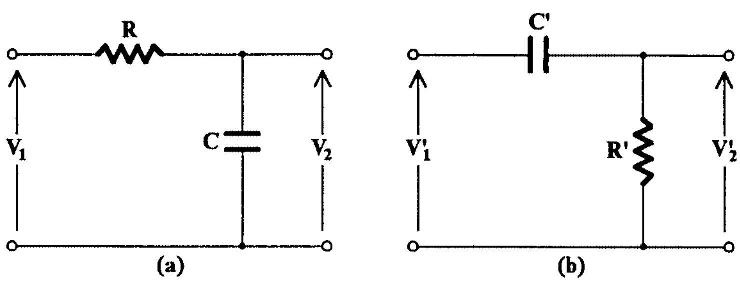

FIGURE 2.24
Simple (a) lowpass and (b) highpass RC circuits.

$$F'(s) = \frac{V'_2}{V'_1} = \frac{sC'R'}{sC'R' + 1} = \frac{s}{s + 1/C'R'} \tag{2.58}$$

It can be seen that, with this interchange of elements, the lowpass circuit in Fig. 2.24(a) has been transformed to the highpass of Fig. 2.24(b). The two filters will have the same cutoff frequency if

$$RC = C'R' \tag{2.59}$$

As a second example, consider the two RC active circuits in Fig. 2.25(a) and (b), using an operational amplifier as the active element (see Chapter 3). Circuit (b) is obtained from circuit (a) by changing all resistors to capacitors and vice versa.

The transfer function V_2/V_1 of circuit (a) is

$$F(s) = \frac{V_2}{V_1} = -\frac{1}{CR_1}\frac{1}{s + \frac{1}{CR_2}} = -\frac{R_2/R_1}{sCR_2 + 1} \tag{2.60}$$

and that of circuit (b) is

$$F'(s) = \frac{V'_2}{V'_1} = -\frac{C_1}{C_2}\frac{s}{s + \frac{1}{C_2R}} = -\frac{sC_1R}{sC_2R + 1} \tag{2.61}$$

Clearly, again the lowpass circuit has been transformed to a highpass simply by changing resistors to capacitors and vice versa in the lowpass circuit. For the two circuits to have the same cutoff frequency, the following relationship should hold:

$$CR_2 = C_2R \tag{2.62}$$

In both these examples, if the required denormalized cutoff frequency is ω_c, the time constants in Eqs. (2.59) and (2.62) have to be divided by ω_c. This means that either the capacitance or the resistance which determine the time constant should be divided by ω_c and not both. Compare this case with LC filters where both L and C are divided by ω_c.

The above is part of the so called RC:CR transformation, by means of which a lowpass RC circuit, passive or active, is transformed to the corresponding highpass, under the con-

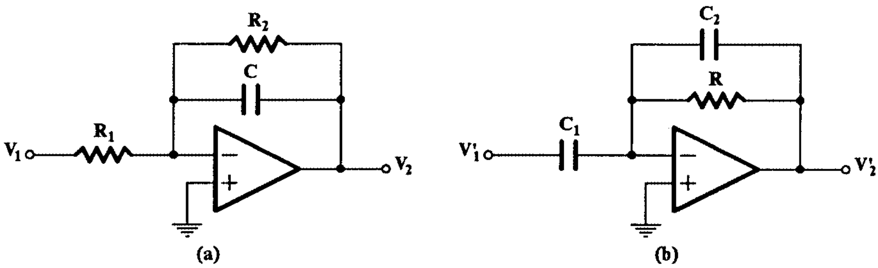

FIGURE 2.25
(a) Lowpass and (b) highpass RC active filters.

dition that the pertinent transfer function is dimensionless, i.e., ratio of voltages or currents. According to this transformation the element R_i is transformed to the element C_i with nominal value $1/R_i$ and vice versa. Resistances, which determine the voltage gain, or the current gain of the active element, do not change, which is not true if the active element is a voltage controlled current source (VCCS) or a current controlled voltage source (CCVS).

As a further example, consider the RC active filter in Fig. 2.26(a), which realizes the second-order Butterworth lowpass function.

$$F(s) = \frac{1}{s^2 + \sqrt{2}s + 1} \tag{2.63}$$

Applying the RC:CR transformation, this circuit is changed to that of Fig. 2.26(b), which is highpass with cutoff frequency $\Omega_c = 1$ equal to the cutoff frequency of the corresponding lowpass in Fig. 2.26(a). If we divide both resistances or both capacitances by ω_c (rad/s), their cutoff frequency becomes ω_c.

2.8 Impedance Scaling

If the desired cutoff frequency of the circuits in Fig. 2.24 is $\omega_c = 1$ rad/s, this can be achieved for

$$C = 1\text{ F} \qquad \text{R} = 1\ \Omega$$

However, the same result can be achieved if

$$C = 1\text{ mF} \qquad \text{R} = 1\ k\Omega$$

or

$$C = 1\ \mu\text{F} \qquad \text{R} = 1\ M\Omega$$

and so on.

When we select the second or third set of values of R and C instead of the first, in actual fact, we have multiplied the impedances of these components by 10^3 or 10^6, respectively,

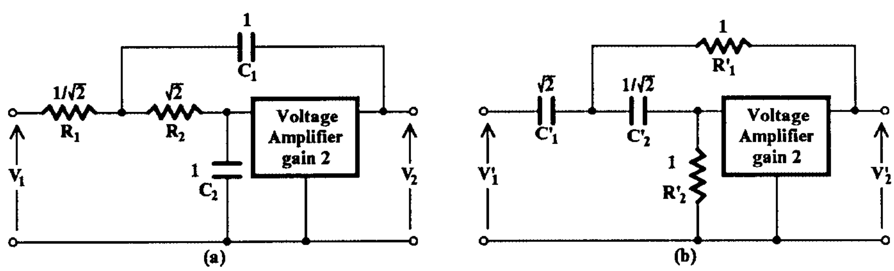

FIGURE 2.26
(a) Second-order RC active lowpass and (b) corresponding RC active highpass filter.

obtaining thus more practical values for the elements R and C. We say that we have raised the impedance level of the circuit by 10^3 and 10^6, respectively.

This denormalization of a circuit to the impedance level R_o requires multiplication by R_o of the values of all resistances, and all inductances in the circuit and division by R_o of all capacitances. In the case when the gain of the active element(s) is determined by the ratio of two resistances, which do not affect the circuit time constants, the impedance level for these resistances may be different from that for the rest of the circuit. Also, if a filter is realized as the cascade connection of low-order functions (i.e., first and second order), which are isolated from each other, again the impedance level in each section can be different from that in the other.

As an example of impedance scaling, we will calculate the values of the components in the circuit in Fig. 2.26(b) for an impedance level of 10^4 Ω and a cutoff frequency at 10^4 rad/s. In accordance with the above discussion, we multiply all resistances by 10^4 and divide all capacitances by 10^4 in order to perform impedance scaling.

Next we may choose to divide the values of the capacitors by another 10^4 or the values of the resistors by this same scaling factor in order to achieve frequency denormalization. Choosing to frequency scale the capacitances, the following set of component values is obtained:

$$R_1 = R_2 = 10\ \text{k}\Omega, \qquad C_1 = \sqrt{2} \times 10^{-8}\ \text{F}, \qquad C_2 = \frac{1}{\sqrt{2}} \times 10^{-8}\ \text{F}$$

These component values are more suitable for RC active filters (particularly if the active element is an operational amplifier) than if we had chosen to frequency scale the resistances. It should be pointed out, however, that frequency scale could have also been obtained if all resistances had been divided by 10^a and all capacitances by 10^{4-a}, since the time constants that actually matter would have been scaled by $10^a \times 10^{4-a} = 10^4$.

2.9 Predistortion

It is well understood that the components one uses in order to build up a circuit one has designed are not ideal. Thus, the equivalent of a coil is not a pure inductance, but it has some loss associated with it. This loss is modeled by a small resistance r connected in series with its inductance L. Similarly, there is some loss associated with the capacitance of a capacitor, usually negligible in today's capacitors, which is modeled by connecting a conductance g in parallel with the capacitance C.

For the sake of argument, let us suppose that the ratios r/L and g/C are both equal to d. Thus the impedance of the coil and the admittance of the capacitor will be, respectively,

$$Z_L = sL + r = L(s + d)$$

and

$$Y_C = sC + g = C(s + d)$$

The circuit transfer function, as derived by circuit analysis, is found to be a linear function of impedance or admittance ratios. This means that the poles and zeros of the transfer function will move to the left in the s-plane by the same amount, d. Thus, the frequency response of the practical circuit will differ from the expected one, i.e., it will be distorted. In some cases, this effect may not matter, but in cases of highly selective bandpass filters, the distortion in the frequency response will be serious and thus unacceptable.

To counterbalance this type of distortion in the frequency response of the practical circuit to be built out of coils, capacitors, and resistors, the designer can shift the poles and zeros of the transfer function to the right on the s-plane by the same amount d and then proceed to calculate the component values. With the application of this technique, known as predistortion, the poles and zeros of the transfer function of the practical circuit will be placed nearly at the initially wanted positions, moved there, because of the power dissipation of the practical components, coils, and capacitors. Predistortion is in effect a frequency transformation of the initial transfer function $F(s)$ to $F(p)$ where $p = s - d$. This transformation is nonreactive in that it is not applied to the reactive elements of the circuit.

In practice, the losses in coils and capacitors will not be the same. Then, the designer can add resistances in series with the coils and in parallel to the capacitors in order to obtain the same factor d in all components. It should be noted, though, that because of the introduction of dissipative components (the resistors) in the LC circuit, there will be an increase in the flat loss of the circuit. Since this loss can, in some cases, be intolerable, and the number of the additional resistors excessive, the predistortion technique may not always produce an attractive solution to the unavoidable problem of dissipation introduced by the practical reactive components.

2.10 Summary

The problem of determining a filter function satisfying the specifications of a filter has been examined in this chapter to some detail.

Filter specifications may refer mainly either to amplitude (magnitude) requirements or to phase (or delay) requirements. Determining first the order of the lowpass prototype filter function, the designers can then choose the most suitable one among the available in tables Butterworth, Chebyshev, Papoulis, or Cauer functions, if they are interested in the amplitude response. Similarly, if they are interested in the phase response, they may choose a lowpass or an allpass function among the available in tabulated form Bessel-Thomson, Padè, or Chebyshev-type delay functions.

In the case of amplitude or phase equalization, the designers will basically have to work heuristically using the computer as their main tool and a suitable optimization program.

Suitable frequency and element transformations were introduced in order to transform the lowpass prototype filter to the required denormalized lowpass, highpass, bandpass, or bandstop filter. These frequency transformations are also useful in the translation of the denormalized filter specifications to the corresponding lowpass prototype requirements.

Once the most suitable circuit has been chosen (in a way we shall see in later chapters) for the realization of the denormalized function, suitable impedance scaling should be applied to the component values in order to make the circuit more practical within its environment (signal level, source impedance, load impedance, and characteristics of the active element). Impedance scaling was also introduced, while in the final section of this chapter the concept of predistortion was introduced briefly.

Before we examine the selection of suitable circuits for the realization of filter functions we introduce in the next chapter various active elements that will be used in subsequent chapters.

References

[1] F.F. Kuo. 1966. *Network Analysis and Synthesis*, 2/e, New York: John Wiley & Sons.

[2] A. Papoulis. 1959. "On Monotonic Response Filters," *Proc. IRE*, 47, pp. 332–333.

[3] G. Daryanani. 1976. *Principles of Active Network Synthesis and Design*, New York: John Wiley & Sons.

[4] A. Zverev. 1967. *Handbook of Filter Synthesis*, New York: John Wiley & Sons.

[5] E. Kuh and D. Pederson. 1959. *Principles of Circuit Theory*, New York: McGraw-Hill.

[6] R.W. Daniels. 1974. *Approximation Methods for Electronic Filter Design*, New York: McGraw-Hill.

[7] D.F. Tuttle. 1958. *Network Synthesis*, Vol. 1, New York: John Wiley & Sons.

[8] A. Budak. 1965. "Maximally flat phase and controllable magnitude approximation," *Proc. IEEE* CT-12, p. 279.

[9] W.J. King and V.C. Rideout. 1961. "Improved transport delay circuits for analogue computer use," *Proc. 3rd Intl. Meeting of the Association pour le Calcul Analogique*, Opatija, Yugoslavia, Sept. 5.

[10] T. Deliyannis. 1970. "Six new delay functions and their realization using active RC networks," *The Radio and Electronic Engineer*, **39**(3), pp. 139–144.

[11] D. Humphreys. 1964. "Rational function approximation of polynomials to give an equiripple error," *IEEE Trans Circuit Theory* CT-11, pp. 479–486.

[12] L. Weindberg. 1962. *Network Analysis and Synthesis*, New York: McGraw-Hill.

[13] K. Skiwrzinski. 1965. *Design Theory and Data for Electrical Filters*, New York: Van Nostrand.

[14] R. Saal. 1979. *Handbook of Filter Design*, AEG-TELEFUNKEN.

[15] D.E Johnson and J.L. Hilburn. 1975. *Rapid Practical Designs of Active Filters*, New York: John Wiley & Sons.

[16] S. K. Mitra. 1969. *Analysis and Synthesis of Linear Active Networks*, New York: John Wiley & Sons.