CASCADE NETWORKS

Cascade design refers to the procedure in which the designer chooses to factor a complex circuit requirement of order $n \geq 2$ that is difficult or impossible to realize in its given form into a number of simpler specifications that result in more practical circuits and are more readily implemented. These are then connected in a chain, that is, in a *cascade circuit,* to realize the specified requirements. Typically, one factors a required high-order transfer function

$$H(s) = \frac{V_{\text{out}}}{V_{\text{in}}} = \frac{N(s)}{D(s)} = \frac{b_{2m}s^{2m} + b_{2m-1}s^{2m-1} + \cdots + b_1 s + b_0}{s^{2n} + a_{2n-1}s^{2n-1} + \cdots + a_1 s + a_0} \tag{1}$$

with $n \geq m$ and $n > 1$, into a number of lower-order functions $T_j(s)$. $H(s)$ in Eq. (1) is a ratio of two polynomials $N(s)$ and $D(s)$ of degrees $2m$ and $2n$, respectively. We selected, without loss of generality, the coefficient $a_{2n} = 1$, because numerator and denominator can always be divided by a_{2n}. The realization then connects the functions $T_j(s)$ such that

$$H(s) = T_1(s)T_2(s)\cdots T_j(s)\cdots T_n(s) = \prod_{j=1}^{n} T_j(s) \tag{2}$$

that is, the total transfer function is obtained from the product of lower-order functions. Examples of cascade synthesis extend from the early days of electronics when *RLC* circuits were isolated by vacuum-tube amplifiers to the present-day discrete or fully integrated realizations of active filters (*RC* circuits augmented or isolated by operational amplifiers) and switched-capacitor circuits where resistors in active filters are replaced by periodically switched capacitors. Cascade design is used widely not only for the implementation of magnitude and phase responses but also for group-delay equalizers. The goal is to realize $H(s)$ via simpler circuits in an efficient way with low sensitivities to component tolerances. The sensitivity issue will be addressed below.

In Eq. (1), we have labeled the degrees of the numerator and denominator polynomials $N(s)$ and $D(s)$ as $2m$ and $2n$, respectively, to emphasize the fact that we assume the degrees to be even. Both $N(s)$ and $D(s)$ can therefore be factored into the product of second-order pole-zero pairs, as expressed in the following form:

$$H(s) = \frac{N(s)}{D(s)} = \frac{\prod_{i=1}^{m}(\beta_{2i}s^2 + \beta_{1i}s + \beta_{0i})}{\prod_{j=1}^{n}(s^2 + \alpha_{1j}s + a_{0j})} = \prod_{j=1}^{n}\frac{\beta_{2j}s^2 + \beta_{1j}s + \beta_{0j}}{s^2 + \alpha_{1j}s + \alpha_{0j}} = \prod_{j=1}^{n} T_j(s) \tag{3}$$

The notation assumes that both N and D are of degree $2n$; if $m < n$, the numerator will contain $2(n - m)$ factors of unity. If the degree of $H(s)$ is odd, the function can always be factored into the product of even terms as shown in Eq. (3) and a first-order factor. First-order sections can easily be realized by a passive *RC* network and can be appended to the higher-order circuit as an additional term.

Example 1. We now illustrate the decomposition of a higher-order transfer function into a product of lower-order transfer functions. The transfer function of a sixth-order filter, with a normalized frequency parameter, is

$$H(s) = \frac{s^5 + 2.5s^4 + 0.5625s}{s^6 + 0.390s^5 + 3.067s^4 + 0.785s^3 + 3.056s^2 + 0.387s + 0.989}$$

To realize this filter as a cascade circuit, the numerator and denominator are factored by a suitable root-finding algorithm as follows:

$$\begin{aligned} H(s) &= \frac{s(s^2 + 0.25)(s^2 + 2.25)}{(s^2 + 0.09s + 0.83)(s^2 + 0.2s + 1.01)(s^2 + 0.1s + 1.18)} \\ &= T_1(s)T_2(s)T_3(s) \\ &= \frac{s}{s^2 + 0.09s + 0.83}\,\frac{s^2 + 0.25}{s^2 + 0.2s + 1.01}\,\frac{s^2 + 2.25}{s^2 + 0.1s + 1.18} \end{aligned} \tag{4}$$

In this case, $n = 3$ and the numerator is odd, that is, $N(s)$ is factored into a product of two second-order factors and a first-order factor, and the function is presented as the product of three terms according to Eq. (2).

Example 2. This example illustrates how cascading of lower-order sections enables the realization of an amplifier with a large inverting gain, $K = -1{,}000{,}000$. The bandwidth must be at least 5 kHz and the smallest resistor used should be 1 kΩ to minimize currents and loading effects.

If one attempts to realize the amplifier in the usual way with a 741-type operational amplifier as shown in Fig. 1(a), a 1000 MΩ = 1 GΩ = 10^9 Ω resistor is required. This resistor is too large to be realized; it is essentially an open circuit and leaves the operational amplifier operating in an open loop. The bandwidth would be less than 1 Hz. Because of such processing or technology constraints, it is convenient, even necessary, to partition the prescribed gain into several factors, such as $K = K_1K_2K_3 = (-100) \times (-100) \times (-100)$, and connect the resulting circuits in cascade [Fig. 1(b)]. In this configuration, the second amplifier picks up the output of the first one with gain -100 and multiplies it by a second factor -100, and so on, to realize $K = -1{,}000{,}000$ as required. Of course, with this large amplification of the factor 10^6, the designer must pay careful attention to avoid signal-level distortion. This issue will not be addressed here. At the expense of additional circuitry, cascade design enabled us to realize the specifications with the required gain, bandwidth (it is larger than 8 kHz), and practical component values. Without the cascade method, the specifications placed such demands on the components to render the circuit unrealizable.

In a similar manner, high-order active filters described by a transfer function, $H(s)$, are most frequently designed as a cascade of low-order circuits. The method results in filters that can be adjusted (or *tuned*) easily and have low sensitivity to component tolerances Ref. 1. In addition, the design method is completely general, in that transfer func-

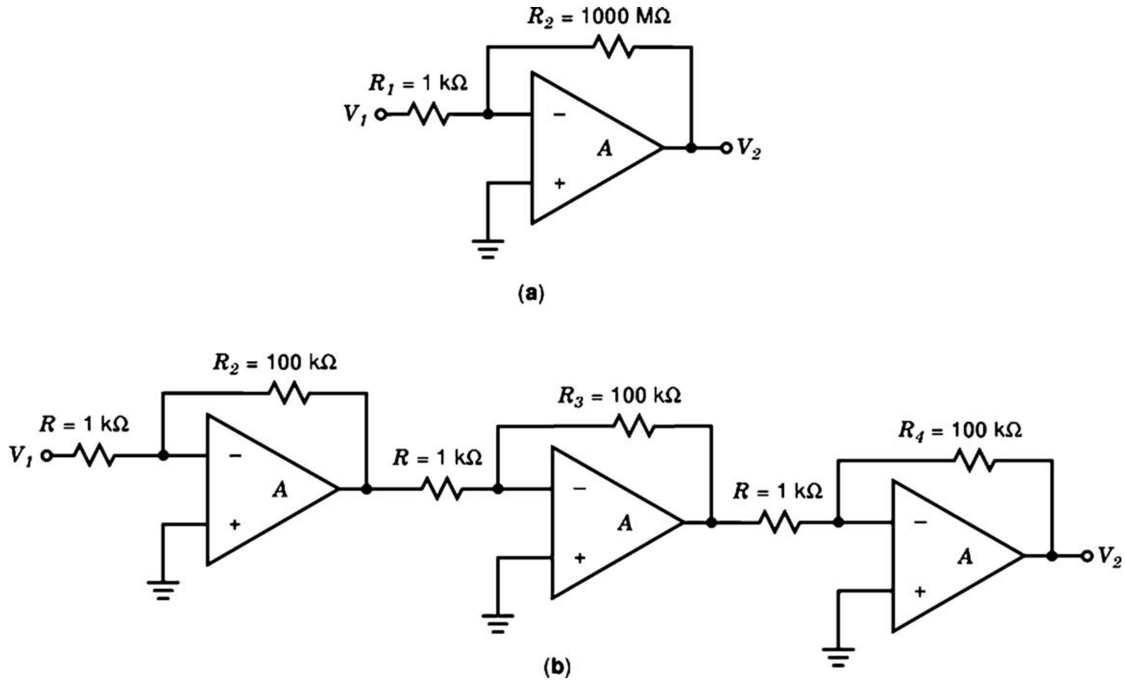

Figure 1. (a) Proposed operational amplifier circuit to realize $K = -R_2/R_1 = -1{,}000{,}000$. The circuit is *not* practical because of the large resistor R_2 and because the minimal feedback limits the bandwidth to very low values; (b) practical realization as a cascade of three amplifiers with gain $K_\mathrm{i} = -100$.

tions of arbitrary form can be realized. Naturally, to keep the circuits stable, that is, to prevent them from oscillating, the poles, the roots of $D(s)$, are restricted to the left half of the complex s plane ($\alpha_{0j} > 0$, $\alpha_{1j} > 0$). The transmission zeros, the roots of $N(s)$, however, can be anywhere in the s plane; their location depends only on the desired response. An additional advantage is that in cascade design the designer can focus on the low-order sections $T_\mathrm{j}(s)$ that are generally simpler to implement than the high-order function $H(s)$. The low-order sections are then simply connected in a chain, or *cascaded,* as shown in Fig. 2. If $T_\mathrm{j}(s)$ is defined as a voltage transfer function,

$$T_j(s) = \frac{V_{\mathrm{o}\,j}}{V_{\mathrm{i}\,j}} \tag{5}$$

the total transfer behavior realized by that circuit is derived to be

$$\begin{aligned} H(s) &= \frac{V_\mathrm{out}}{V_\mathrm{in}} = \frac{V_\mathrm{o1}}{V_\mathrm{in}}\frac{V_\mathrm{o2}}{V_\mathrm{o1}}\frac{V_\mathrm{o3}}{V_\mathrm{o2}}\cdots\frac{V_{\mathrm{o}(n-1)}}{V_{\mathrm{o}(n-2)}}\frac{V_\mathrm{out}}{V_{\mathrm{o}(n-1)}} \\ &= T_1(s)T_2(s)\ldots T_{n-1}(s)T_n(s) \end{aligned} \tag{6}$$

as required by Eq. (2). Equation (7) holds because the connection guarantees that the input voltage of section j is equal to the output voltage of section $j-1$, $V_{\mathrm{i}j} = V_{\mathrm{o}(j-1)}$.

SENSITIVITY

The *sensitivity* of a cascade realization of a transfer function $H(s, x)$ to a component x is calculated via derivatives,

$$\frac{\partial H(s,x)}{\partial x} = T_1T_2\cdots T_{j-1}T_{j+1}\cdots T_n\frac{\partial T_j(s,x)}{\partial x} \tag{7}$$

where we assumed that the component x is located in section j, and partial derivatives are used because H is a function of s and of all circuit components. Customarily, both the transfer function deviation ∂H and the component deviation ∂x are normalized to H and x, respectively, so that we obtain from Eq. (8) with Eq. (2)

$$\frac{x}{H}\frac{\partial H(s,x)}{\partial x} = \frac{x}{T_j}\frac{\partial T_j(s,x)}{\partial x} \tag{8}$$

These quantities are the classical sensitivities, defined as

$$S_x^{F(x)} = \frac{x}{F}\frac{\partial F(x)}{\partial x} = \frac{\partial F/F}{\partial x/x} = \frac{\partial \ln F}{\partial \ln x} \tag{9}$$

Equation (6) can be rewritten in the form

$$\frac{\partial F}{F} = S_x^{F(x)}\left(\frac{\partial x}{x}\right) \tag{10}$$

which means that the percentage error in a function $F(x)$ caused by a component with tolerance ∂x is computed by multiplying the percentage error of the component by the sensitivity. From Eq. (9) we have the important result

$$S_x^{H(x)} = S_x^{T_j(x)} \tag{11}$$

which says that in a cascade circuit the sensitivity of the total transfer function H to a component is equal to the sensitivity of the section T_j that contains the component. In contrast to noncascaded implementations, in which the sensitivity to a component may depend on *all* poles and zeros and may become very large Ref. 1, here it depends only on the pole and zero in section j and is not affected by

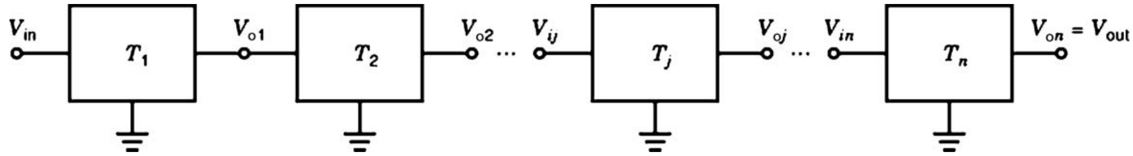

Figure 2. Cascade realization of a $2n$th-order transfer function. The realized transfer function is the product of the individual function realized by the blocks T_1.

any of the other sections. Equation (12) verifies formally the intuitive expectation that the cascade circuit with good sensitivity behavior should be built from low-sensitivity sections.

CASCADING CONDITION

A general condition that must be satisfied for Eq. (7) to hold in the simple form shown is that the sections do not load each other and do not otherwise interact. This condition requires that the output impedance Z_{out} of any section ideally is zero and its input impedance Z_{in} infinite. Figure 3 illustrates this requirement for two sections T_1 and T_2, defined as

$$T_1(s) = \frac{V_{n1}}{V_1} \quad \text{and} \quad T_2(s) = \frac{V_{n2}}{V_{i2}}$$

V_{n1} and V_{n2} are the voltages at the respective internal nodes n_1 and n_2 in Fig. 3. The complete sections, realistically modeled to include a finite input impedance $Z_i(s)$ and output impedance $Z_o(s)$, are contained in the dashed boxes. Assuming that V_1 is an ideal voltage source so that Z_{i1} has no effect and that there is an open circuit at V_{o2} so that no current flows through Z_{o2}, analysis of the configuration in Fig. 3, starting from the output, yields

$$\begin{aligned} V_{o2} &= T_2(s)V_{i2} = T_2(s)\,\frac{Z_{i2}(s)}{Z_{i2}(s)+Z_{o1}(s)}\,V_{n1} \\ &= T_2(s)\,\frac{Z_{i2}(s)}{Z_{i2}(s)+Z_{o1}(s)}\,T_1(s)V_1 \end{aligned} \tag{12}$$

Thus, the transfer function of the two-section cascade is

$$H(s) = \frac{V_{o2}}{V_1} = T_1(s)T_2(s)\frac{Z_{i2}(s)}{Z_{i2}(s)+Z_{o1}(s)} \tag{13}$$

rather than the simple product T_1T_2 intended. The problem is that the current through Z_{o1} is not zero, that is, section 1 is loaded by section 2. Equation (15) indicates that the desired result is achieved provided that

$$|Z_{i2}(\,j\omega)| \gg |Z_{o1}(\,j\omega)| \tag{14}$$

Ideally, $Z_{i2} = \infty$ and $Z_{o1} = 0$. Condition (14) is satisfied in the circuit in Fig. 1(b) because the operational amplifier circuits with feedback and a gain of 100 have very small output resistance ($R_{\text{out}} < 50\ \Omega$) compared to the input resistance of the following stage ($R_{\text{in}} = 1\ \text{k}\Omega$). The condition to permit cascading is, therefore, that the input impedance of the loading section must be much larger than the output impedance of the driving section so that the voltage divider factor that multiplies the ideal product T_1T_2 in Eq. (15) is as close to unity as possible. The voltage divider ratio is written on purpose in terms of impedances as functions of s or $j\omega$ to emphasize that, in practice, the designer needs to contend with more than a resistor ratio that would multiply T_1T_2 simply by a frequency-independent constant.

To illustrate this point consider the example of a second-order low-pass filter being built by cascading two identical first-order sections as in Fig. 4. The sections have the transfer function,

$$T_1(s) = T_2(s) = \frac{1}{sCR+1} \tag{15}$$

The intent is to multiply the transfer functions T_1 and T_2 so that

$$\begin{aligned} H(s) &= \frac{V_{o2}}{V_1} = T_1(s)T_2(s) = \frac{1}{sCR+1}\,\frac{1}{sCR+1} \\ &= \frac{1}{(sCR)^2+2sCR+1} \end{aligned} \tag{16}$$

However, because of the finite impedances identified in the figure,

$$Z_{i2} = R + \frac{1}{sC} \quad \text{and} \quad Z_{o1} = \frac{1}{sC+1/R}$$

by Eq. (15) the proposed circuit in Fig. 4 realizes the transfer function

$$\begin{aligned} H(s) &= \frac{V_{o2}}{V_1} = \left(\frac{1}{sCR+1}\right)^2 \frac{R+\dfrac{1}{sC}}{\left(R+\dfrac{1}{sC}\right)+\left(\dfrac{R}{sCR+1}\right)} \\ &= \frac{1}{(sCR+1)^2}\,\frac{(sCR+1)^2}{(sCR)^2+3sCR+1} = \frac{1}{(sCR)^2+3sCR+2} \end{aligned} \tag{17}$$

The desired performance is completely altered as can be confirmed by direct analysis. It is seen that the finite frequency-dependent input and output impedances substantially change the desired transfer function. If cascading must be used but Eq. (16) is not satisfied, buffering using a voltage follower can be employed as in Fig. 5 so that no current is drawn from the first section and the second section is driven by a nearly ideal voltage source.

To repeat the important condition that must be satisfied if two circuits are to be connected in cascade: the trailing section must not load the leading section. The circuits to be cascaded can be of low order, high order, active, passive, or any desired combination; the loading restriction does not change. For designing high-order active filters, the cascade connection normally consists of first- and second-order active building blocks because they can be cascaded directly with no need of buffering: since active circuits usually have an operational amplifier output as their output terminal, unbuffered cascading is possible because the operational amplifier output resistance in a feedback network is very small (see Example 2). On the other hand, as in the example of Fig. 4, it can be expected intuitively that passive circuits can generally not be connected in cascade without

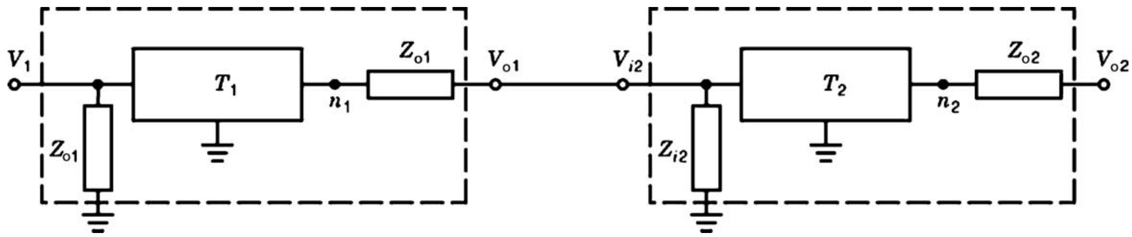

Figure 3. Two-section cascade with finite input and output impedances. The second section *loads* the first one unless its input impedance, Z_{i2}, is much larger than the output impedance, Z_{o1}, of the first section.

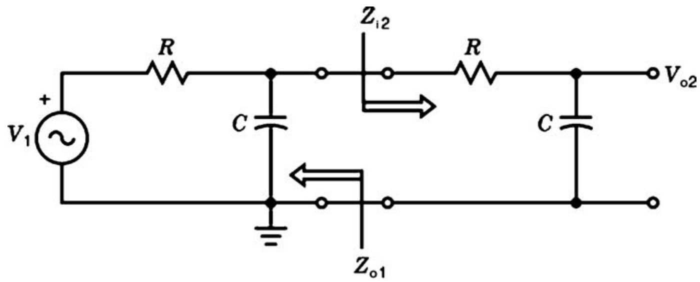

Figure 4. Cascade connection of two first-order passive low-pass sections with finite input and output impedances. The two modules interact so that the realized transfer function is *not* equal to the product of the two first-order functions. Isolating the sections is required as shown in Fig. 5.

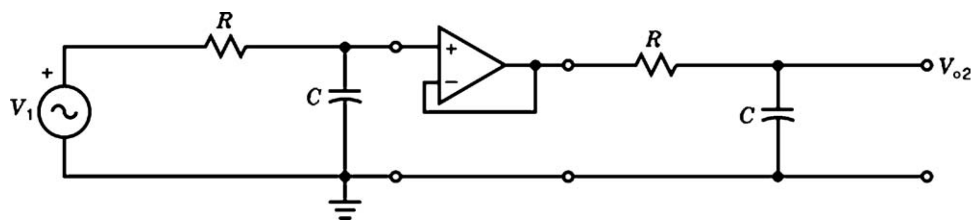

Figure 5. Two low-pass modules isolated by an operational amplifier. The unity-gain voltage follower isolates the sections' performance so that the total transfer function is the product of the two first-order modules.

buffering because the condition in Eq. (16) will rarely be satisfied.

THE DESIGN APPROACH AND TRADE-OFFS

The problem to be addressed is how the transfer function of Eq. (1) can be realized in an efficient way with simple low-order circuits and low sensitivities to component tolerances. As was mentioned, the sensitivity behavior of high-order filter realizations shows Ref. 1 that, in general, it is not advisable to realize the transfer function $H(s)$ of Eq. (1) in the so-called *direct form,* using only one or maybe two operational amplifiers embedded in a high-order passive RC network. Although it is possible in principle to realize Eq. (1) in direct form, the resulting circuits are normally so sensitive to component tolerances that reliable performance cannot be expected in practice. A further disadvantage of the direct synthesis method is the use of a very large number of passive components to realize a function of given order.

In the cascade approach, the high-order function $H(s)$ is factored into functions of second order as indicated in Eq. (3). The resulting biquadratic functions are realized by the methods discussed elsewhere in this encyclopedia and connected in cascade such that their product implements the prescribed function $H(s)$. The cascade method is used widely in industry because it is applicable generally, is well understood, very easy to implement, and efficient in its use of active devices (as few as one operational amplifier per pole pair). It uses a modular approach and results in filters that for the most part show satisfactory performance in practice. One of the main advantages of cascade filters is that they are very easy to tune because each biquadratic section, referred to as *biquad* in the literature, is responsible for the realization of only one pole pair and zero pair: the realizations of the individual critical frequencies of the filter are *decoupled* from each other. The disadvantage of this decoupling is that for filters of high order, say larger than order eight ($n > 4$), with stringent requirements and tight tolerances, cascade designs are often found to be still too sensitive to component variations in the passband. In these cases, ladder simulations may lead to more reliable circuits Ref. 1.

As shown in Eq. (3), the high-order transfer function $H(s)$ is factored into a product of second-order blocks,

$$T_j(s) = k_j \frac{\beta_{2j}s^2 + \beta_{1j}s + \beta_{0j}}{s^2 + s\omega_{0j}/Q_j + \omega_{0j}^2} = k_j t_j(s) \quad (18)$$

where the denominator is expressed in terms of the usual filter parameters, the quality factor Q and the pole frequency ω_0. We have also introduced a suitably defined *gain constant,* k_j, for example such that the leading coefficient in the numerator of the *gain-scaled* transfer function $t_j(s)$ is unity or such that $|t_j(j\omega_{0j})| = 1$. The sections $T_j(s)$ are quite arbitrary; for example, they can realize a low-pass function ($\beta_{2j} = \beta_{1j} = 0$), a bandpass function ($\beta_{2j} = \beta_{0j} = 0$), a high-pass function ($\beta_{1j} = \beta_{0j} = 0$), or a pair of finite transmission zeros on the $j\omega$ axis ($\beta_{1j} = 0$), as the process does not change. Because second-order filter sections can be built to realize arbitrary functions of the form of Eq. (21), cascade design is very general in permitting the realization of any type of stable transmission requirement. If we then may assume that the output impedances of the biquadratic functions are sufficiently small (compared to the input impedances), all second-order blocks can be connected in cascade, Fig. 2, without causing mutual interactions due to loading, and the product of the biquadratic functions is realized as required by Eq. (7).

Although this simple process leads in a straightforward way to a possible cascade design, it leaves several questions unanswered:

1. Which zero should be assigned to which pole in Eq. (3) when the biquadratic functions $T_j(s)$ are formed? Since we have n pole pairs and n zero pairs (counting zeros at 0 and at ∞) we can select from n factorial, $n! = 1 \times 2 \times 3 \times \cdots \times n$, possible pole–zero pairings.
2. In which order should the biquadratic sections in Eq. (7) be cascaded? Does the cascading sequence make a difference? For n biquadratic sections, we have $n!$ possible sequences.
3. How should the gain constants k_j in Eq. (21) be chosen to determine the signal level for each biquad? In other words, what is the optimum gain distribution?

Because the total transfer function is the product of the biquadratic sections, the selections in steps 1–3 are quite arbitrary as far as $H(s)$ is concerned. However, they do determine significantly the dynamic range, that is, the distance between the maximum possible undistorted signal and the noise floor: the maximum and minimum signal levels throughout the cascade filter can be shown to depend on the choices in steps 1–3. Although the sensitivities to component tolerances are functions of pole–zero pairing, the effect usually is not very strong. For a detailed treatment see Refs. 2 and 3. Also, the selection of pole–zero pairing for best sensitivity can be shown to conflict often with the choice necessary for best dynamic range. Since the implementation of the blocks $T_j(s)$ normally makes use of active devices, such as operational amplifiers, depending on the operating frequency, the maximum undistorted signal voltage that a filter can process is limited either by the power supply or by the slew rate of the operational amplifiers.

The optimal cascading routine to be discussed (pole–zero pairing, section ordering) and gain assignment, is entirely general and does not depend on the transfer function or its type. It is independent of the actual implementation of the second-order building blocks. The designer may choose any convenient technology and the circuit architecture that seems preferable from the point of view of sensitivity, numbers and kinds of elements, values and element value spreads, power consumption, or other practical considerations.

DYNAMIC RANGE

Since it is the dominant effect of pole–zero pairing, section ordering, and gain assignment, we will be concerned here only with dynamic range issues. To help tackle the problem, let us label the maximum signal level that can be handled with no distortion as $V_{\mathrm{o,max}}$. We assume that it is measured at the *output* of the biquadratic sections. This assumption will always be correct in single-amplifier biquadratic sections for which section output and operational amplifier output are the same. In multiamplifier biquadratic sections, each operational amplifier output must be evaluated and the maximum operational amplifier output voltage in the biquadratic section must be determined. To avoid overdriving any operational amplifier sooner than any other one inside a biquadratic section, it is intuitively reasonable that any available design freedom in the biquadratic sections should be chosen such that all operational amplifiers "see" the same signal level. If this is not possible, that is, if the undistorted output voltage V_{oi} of section i is only a fraction $1/q_i$ of the maximum internal voltage, with $q_i > 1$, V_{oi} in the following equations must be replaced by $q_i V_{oi}$. For simplicity, we shall assume in our discussion that the sections can be designed such that $q_i = 1$.

We must ensure then that the signal level at any section output, $|V_{oi}(j\omega)|$, satisfies

$$\max|V_{oi}(j\omega)| < V_{\mathrm{o,max}}, \quad 0 \le \omega \le \infty, \quad i = 1, \ldots n \qquad (19)$$

Note that this condition must indeed be satisfied for *all* frequencies and not only in the passband because large signals even outside the passband must not be allowed to overload and saturate the operational amplifiers: when operational amplifiers are overdriven, their operation becomes *nonlinear*. The circuit, however, may still act as a filter and remove the higher harmonics that are generated by the nonlinear operational amplifier operation. The problem that arises when saturating the operational amplifiers is, therefore, not so much harmonic distortion of the signal but changed operating points, intermodulation distortion, and deviations of the magnitude response Ref. 1.

The lower limit of the useful signal range is set by the noise floor. If in the passband of a cascade filter the signal at an internal stage becomes very small, it must be amplified again to the prescribed output level. From any point in the cascade of filter stages, say at the output of stage i, signal *and* noise are amplified by the same amount, namely,

$$H_i^+(s) = \prod_{k=i+1}^{n} T_k(s) \qquad (20)$$

Consequently, the signal-to-noise ratio will suffer if in the cascade filter the signal suffers in-band attenuation, that is, if it is permitted to become very small. The function $H^+{}_i(s)$, defined in Eq. (23), is referred to as the noise gain from the output of section i to the filter output. Thus, the second condition to be satisfied by the output voltage of any biquadratic section is

$$\min|V_{oi}(j\omega)| \to \max \quad \text{for } \omega_L \le \omega \le \omega_U, \quad i = 1, \ldots n \quad (21)$$

ω_L and ω_U are the lower and upper, respectively, corners of the passband. In this case we are, of course, only concerned with signal frequencies in the passband, because in the stopband the signal-to-noise ratio is of no interest. Note, however, that for a white noise input the output noise spectrum of a filter section has the same shape as the square of the transfer function magnitude, which means that the highest noise response occurs at the pole frequencies with the highest Q values. Since these are mostly found just beyond the specified corners of the passband they would not be included in the measurement defined in Eq. (24). Therefore, to avoid decreased dynamic range caused by possibly large noise peaks at the passband corners, it is advisable to extend the frequency range beyond the specified passband corners, ω_L and ω_U, into the transition band to cover the pole frequencies with the highest Q values.

The steps of pole–zero pairing, section ordering, and gain assignment will now be chosen such that the conditions in Eqs. (22) and (24) are satisfied. It must be emphasized that these steps do *not* just amount to minor adjustments when designing a cascade filter, but that the cascade circuit will likely not perform to specifications in practice unless these steps are taken at the design stage.

Pole-Zero Pairing

According to Eqs. (22) and (24), the pole–zero pairing should be chosen such that, in a given section, $M_i = \max|V_{oi}(j\omega)|$ is minimized at all frequencies, and $m_i = \min|V_{oi}(j\omega)|$ is maximized in the passband. In other words, using Eq. (21), $|t_i(j\omega)|$ should be as flat as possible in the frequency range of interest. Notice that M_i may lie outside and m_i at the edge of the passband, and that the actual minimum of the magnitude $|t_i(j\omega)|$ lies in the stopband and is of no concern. As the values of M_i and m_i change when the relative positions of the poles and the zeros of $t_i(s)$ are altered, the pole–zero assignment must be chosen such that the ratio M_i/m_i is as close to unity as possible, which means that for each biquadratic function the "measure of flatness"

$$d_i = \frac{M_i}{m_i} - 1, \qquad i = 1, 2, \ldots, n \tag{22}$$

should be minimized. The optimal pole–zero assignment for the total $2n$th-order cascade filter is then the one that minimizes the maximum value of d_i:

$$d_{max} = \max\{d_i\} \to \min, \qquad i = 1, 2, \ldots, n \tag{23}$$

Algorithms that accomplish this task are available in the literature Refs. 4 to 7. Even in fairly simple low-order cases the problem of pole–zero assignment can be quite computation intensive; it requires substantial software and computer resources.

If the appropriate computing facilities are not available, a simple solution that provides good suboptimal results is simply to assign each zero or zero-pair to the closest pole Refs. 4–8. On occasion, depending on system requirements, we may also preassign some pole–zero pair(s) and leave them out of the remaining pairing process. For instance, if the numerator contains a term $s2$, we may prefer to factor it into $s \times s$ instead of $s^2 \times 1$, that is, we may prefer to realize two second-order bandpass sections instead of a high-pass and a low-pass section.

Example 3. Determine the optimal pole-zero pairing for the transfer function of Eq. (5) of Example 1. The transfer function was

$$H(s) = \frac{s}{s^2 + 0.09s + 0.83}\,\frac{s^2 + 0.25}{s^2 + 0.2s + 1.01}\,\frac{s^2 + 2.25}{s^2 + 0.1s + 1.18}$$

The zeros are located at $z_1 = 0$ and $z_2 = \infty, z_{3,4} = \pm j0.5$, and at $z_{5,6} = \pm j1.5$, and the poles are at $p_{1,2} = -0.045 \pm j0.9099, p_{3,4} = -0.1 \pm j1.0$, and $p_{5,6} = -0.05 \pm j1.085$. According to the approximate assignment rule just stated, we should pair $(z_{1,2}, p_{3,4})$, $(z_{3,4}, p_{1,2})$, and $(z_{5,6}, p_{5,6})$. This choice is indicated in Eq. (28):

$$\begin{aligned} H(s) &= T_1(s)T_2(s)T_3(s) \\ &= \frac{s^2 + 0.25}{s^2 + 0.09s + 0.83}\,\frac{s}{s^2 + 0.2s + 1.01}\,\frac{s^2 + 2.25}{s^2 + 0.1s + 1.18} \end{aligned} \tag{24}$$

Section Ordering

After the pole-zero assignment has been solved, the optimal ordering sequence must be determined out of the $n!$ possibilities in which the biquadratic sections can be connected to form the cascade network. For example, for the sixth-order network with three sections in Example 3, there exist six possible ways to cascade the biquadratic functions:

$$T_1T_2T_3 \quad T_1T_3T_2 \quad T_2T_1T_3 \quad T_2T_3T_1 \quad T_3T_1T_2 \quad T_3T_2T_1$$

The best sequence is the one that finds the ordering that maximizes the dynamic range. The procedure is completely analogous to the earlier discussion where pole–zero pairs were chosen to keep the transfer functions of the individual sections as flat as possible. Now the cascade connection is designed such that the transfer functions

$$H_i(s) = \prod_{k=1}^{i} T_k(s), \qquad i = 1, \ldots, n \tag{25}$$

from filter input to the output of the ith intermediate biquadratic section are as flat as possible. H_n is, of course, equal to the total transfer function H. This will help ensure that the maximum signal voltages do not overdrive the operational amplifiers and that, over the passband, the smallest signal stays well above the noise floor. Consequently, the relationships in Eqs. (22) and (24) must be satisfied,

$$\min|V_{oi}(j\omega)| < V_{o,\max} \quad 0 \le \omega \le \infty \tag{26}$$

$$\min|V_{oi}(j\omega)| \to \max \quad \text{for} \quad \omega_L \le \omega \le \omega_U \tag{27}$$

where $V_{oi}(s)$ is now the output voltage of the cascade of the first i sections when driven by an input signal $V_{in}(s)$. With $H_i(s)$ given in Eq. (30), we define the two measures

$$\begin{aligned} M_i &= \frac{\max|V_{oi}(j\omega)|}{|V_{in}(j\omega)|} = \max\left|\frac{V_{oi}(j\omega)}{V_{in}(j\omega)}\right| \\ &= \max|H_i(j\omega)| \quad \text{for} \quad 0 \le \omega \le \infty \end{aligned} \tag{28}$$

and

$$\begin{aligned} m_i &= \frac{\min|V_{oi}(j\omega)|}{|V_{in}(j\omega)|} = \min\left|\frac{V_{oi}(j\omega)}{V_{in}(j\omega)}\right| \\ &= \min|H_i(j\omega)| \quad \text{for} \quad \omega_L \le \omega \le \omega_U \end{aligned} \tag{29}$$

and require again that the flatness criterion of Eq. (25) be minimized, now, however, by choice of the cascading sequence,

$$d_i = \frac{M_i}{m_i} - 1 \rightarrow \min \tag{30}$$

The optimal sequence is the one that minimizes the maximum number d_i as prescribed in Eq. (26). Note that we do not have to consider d_n because, with all sections connected in the cascade filter, d_n is nothing but a measure of the prescribed passband variations (the *ripple*). With the problem identified, the optimum cascading sequence can be found in principle by calculating d_i for all $n!$ sequences and selecting the one that satisfies Eq. (35). As in the pole–zero assignment problem, a brute-force optimization approach involves a considerable amount of computation, and more efficient methods have been developed that use linear programming techniques, such as the "branch and bound" method of Refs. 5–7, or "back track programming" Ref. 9. The necessary computer algorithms are described in the literature.

If the required software routines are not available, the designer must pick a cascading sequence that is based on experience or intuition. A selection that is often very close to the optimum is the one that chooses the section sequence in the order of increasing values of Q_i, that is,

$$Q_1 < Q_2 < \cdots < Q_n \tag{31}$$

so that the section with the flattest transfer function magnitude (the lowest Q) comes first, the next flattest one second, and so on. The possible choices are frequently further limited by other considerations. For example, it is often desirable to have as the first section in the cascade a low-pass or a bandpass section so that high-frequency signal components are kept from the amplifiers in the filter in order to minimize slew-rate problems. Similarly, the designer may wish to employ a high-pass or a bandpass section as the last section in order to eliminate low-frequency noise, dc offset, or power-supply ripple from the filter output. In such situations, the optimum sequencing is performed only on the remaining sections.

The following example illustrates some of the steps discussed.

Example 4. Continue Example 3 to find the optimal cascading sequence for the three second-order sections. Since the coefficient of s in the denominators of the second-order sections of Eq. (28) equals ω_{0i}/Q_i, and the constant coefficient equals $\omega^2{}_{0i}$, we have

$$Q_1 = \frac{\omega_{01}}{0.09} = \frac{\sqrt{0.83}}{0.09} \approx 10.12, \qquad Q_2 = \frac{\sqrt{1.01}}{0.2} \approx 5.03,$$
$$Q_3 = \frac{\sqrt{1.18}}{0.1} \approx 10.86$$

Using Eq. (31) and using the section numbering in Eq. (28), the optimal ordering is, therefore, $T_2T_1T_3$. If instead the design were to emphasize the elimination of high-frequency signals from the filter and low-frequency noise from the output, the ordering $T_2T_3T_1$,

$$\begin{aligned} H(s) &= T_2(s)T_3(s)T_1(s) \\ &= \frac{s}{s^2+0.2s+1.01}\,\frac{s^2+2.25}{s^2+0.1s+1.18}\,\frac{s^2+0.25}{s^2+0.09s+0.83} \end{aligned} \tag{32}$$

would be preferred because the bandpass section, T_2, has the best high-frequency attenuation, and section T_1 provides reasonable attenuation at low frequencies (0.25/0.83 = 0.3 ≈ −10.4 dB), whereas section T_3 has a high-frequency gain of 1 and amplifies low-frequency noise by more than 2.25/1.18 = 1.91 ≈ 5.6 dB. This suboptimal ordering gives almost identical results to the optimal one, $T_2T_1T_3$, because $Q_1 \approx Q_3$.

Gain Assignment

The last step in the realization of a cascade filter is the assignment of the gain constants. Generally, the selection is again based on dynamic range concerns with the goal of keeping the signals below amplifier saturation limits and above the system noise floor. To get a handle on the process, we note that the circuit is linear and all voltages rise in proportion to V_{in}. It is clear then that the maximum undistorted input signal can be processed if we choose the gain constants such that all internal output voltages V_{oi}, $i = 1, \ldots, n-1$, are equal in magnitude to the presumably prescribed magnitude of the output voltage, V_{on}:

$$\begin{aligned} \max|V_{oi}(j\omega)| &= \max|V_{on}(j\omega)| = \max|V_{out}(j\omega)|, \\ i &= 1 = 1, \ldots n-1 \end{aligned}$$

Assuming as before that the output voltage of the biquadratic sections reaches the critical magnitude, this choice ensures that for a given signal level none of the operational amplifiers in the blocks of Fig. 2 is overdriven sooner than any other one. Note, however, the earlier comments about precautions necessary in multiamplifier biquads.

For the analysis it is convenient to use the notation of Eqs. (2), (21), and (30), that is,

$$H(s) = \prod_{j=1}^{n} T_j(s) = \prod_{j=1}^{n} k_j \prod_{j=1}^{n} t_j(s) \tag{34}$$

and, for the intermediate transfer functions,

$$H_i(s) = \prod_{j=1}^{i} T_j(s) = \prod_{j=1}^{i} k_j \prod_{j=1}^{i} t_j(s) \tag{35}$$

Furthermore, we introduce the constant

$$K = \prod_{j=1}^{n} k_j \tag{36}$$

such that

$$\max|H(j\omega)| = K \max \prod_{j=1}^{n} |t_j(j\omega)| = KM_n \tag{37}$$

is the prescribed gain. Similar to the definition of M_n, let us denote the maxima of the intermediate $n-1$ gain-scaled transfer functions by M_i, that is,

$$\max\left|\prod_{j=1}^{i} t_j(j\omega)\right| = M_i, \qquad i = 1,\ldots,n-1 \tag{38}$$

To make $\max|V_{o1}(j\omega)| = \max|V_{out}(j\omega)|$, we then obtain the equation $k_1M_1 = KM_n$, that is,

$$k_1 = K\frac{M_n}{M_1} \tag{39}$$

Similarly, $k_1k_2M_2 = KM_n$, that is, with Eq. (45),

$$k_2 = \frac{M_1}{M_2} \tag{40}$$

results in $\max|V_{o2}(j\omega)| = \max|V_{out}(j\omega)|$. Proceeding in the same manner yields

$$k_j = \frac{M_{j-1}}{M_j}, \qquad j = 2,\ldots,n \tag{41}$$

Choosing the gain constants as in Eqs. 45, 46, 47 guarantees that all operational amplifiers "see" the same maximum voltage to ensure that the largest possible signal can be processed without distortion. Note that changing the total gain of the n-section cascade filter affects only K, that is, k_1. The voltages in all other sections $T_i(s)$, $i = 2, \ldots, n$, increase or decrease proportionally, but their *relative* magnitudes stay the same, as determined by k_i in Eq. (47).

Example 5. Continue Example 4 to find the optimal gain constants for the three second-order sections so that their maximum output levels are equalized. The specified mid-band filter gain is 10 dB. Use the section ordering in Eq. (38).

From Eq. (43), we find

$$KM_3 = 3.16 \tag{42}$$

corresponding to the prescribed gain of 10 dB. The maxima M_i can be computed as

$$M_1 = \frac{1}{0.2} = 5, \qquad M_2 \approx 40, \qquad M_3 \approx 116$$

by evaluating the functions in Eqs. (43) and (44). With M_i known, and using Eq. (48), Eqs. (45) and (47) give

$$k_1 = \frac{KM_3}{M_1} = \frac{3.16}{5} = 0.63, \qquad k_2 = \frac{M_1}{M_2} \approx \frac{5}{40} = 0.125,$$
$$k_3 = \frac{M_2}{M_3} \approx \frac{40}{116} = 0.345$$

that is

$$H(s) = \frac{0.63s}{s^2+0.2s+1.01}\,\frac{0.125(s^2+2.25)}{s^2+0.1s+1.18}\,\frac{0.345(s^2+0.25)}{s^2+0.09s+0.83}$$

These values result in all section outputs being equal to $KM_3 = 3.16$ times the input voltage level for a uniform gain of 10 dB. If the designer were to find out later that system performance would improve for a different filter gain, say, 29 dB rather than 10 dB, it is necessary only to alter the first section in the cascade from $k_1 = 3.16$ to

$$k_1 = KM_3/M_1 = 28.2/5 = 5.64$$

to achieve the new circuit for which dynamic range is still optimized.

To demonstrate that gain equalization is very important in cascade realizations, consider the case where equalization is not performed. Had the designer chosen all $k_i = 1$ as in Eq. (38), the output levels would have been 13.9 dB at V_{o1}, 31.9 dB at V_{o2}, and 41.3 dB at V_{o3}. Quite apart from the specified gain of 10 dB not being realized, the difference of a factor of 41.3 dB − 13.9 dB = 27.4 dB = 23.4 in operational amplifier output voltages would likely result in gross distortions unless the input voltage is kept very small.

BIBLIOGRAPHY

1. R. Schaumann M. S. Ghausi K. R. Laker *Design of Analog Filters: Passive, Active RC and Switched Capacitor*, Englewood Cliffs, NJ: Prentice Hall, 1990.
2. R. Schaumann and M. A. Van Valkenburg, *Design of Active Filters*, New York, NY, Oxford University Press, 2001.
3. M. S. Ghausi K. R. Laker *Modern Filter Design*, Englewood Cliffs, NJ: Prentice Hall, 1981.
4. G. S. Moschytz *Linear Integrated Networks—Design*, New York: van Nostrand Reinhold, 1975.
5. S. Halfin An optimization method for cascaded filters, *Bell Syst. Tech. J.*, **49**: 185–190, 1970.
6. S. Halfin Simultaneous determination of ordering and amplifications of cascaded subsystems, *J. Optimiz. Theory Appl.*, **6**: 356–360, 1970.
7. E. Lüder Optimization of the dynamic range and the noise distance of RC-active filters by dynamic programming, *Int. J. Circuit Theory Appl.*, **3**: 365–170, 1975.
8. A. S. Sedra P. O. Brackett *Filter Theory and Design: Active and Passive*, Portland, OR: Matrix, 1978.
9. W. M. Snelgrove A. S. Sedra Optimization of dynamic range in cascade active filters. In Proc. IEEE Int. Symp. Circuit Syst., 1978, pp. 151–155.

ROLF SCHAUMANN
Portland State University,
Portland, OR