Behzad Razavi

The Biquadratic Filter

The biquadratic filter, also known as the "biquad," dates back to the 1960s [1]–[3] but still serves as an essential building block in analog filter design. In this article, we study this circuit's properties and design issues.

The General Biquad

The biquad is a second-order filter whose transfer function is given, in the general case, by

$$H(s) = \frac{a_1 s^2 + b_1 s + c_1}{a_2 s^2 + b_2 s + c_2}. \quad (1)$$

Here, the numerator coefficients can be chosen to yield a low-pass, band-pass, or high-pass response. For example, $a_1 = b_1 = 0$ leads to a low-pass filter (LPF), the focus of our study here. To realize higher-order filters, biquad sections can be cascaded.

The Need for Complex Poles

We typically begin the design of filters by deciding on the order and shape of their frequency response. For example, Wi-Fi receivers commonly employ a fifth-order LPF to suppress unwanted channels. However, for a given order, the roll-off from the passband to the stopband can be made sharper if some peaking or ripple is tolerable.

It is in this spirit that we turn to transfer functions having *complex* poles. We explain the thought process behind this point by means of an example. Consider a low-pass biquad characterized by the following two (equivalent) transfer functions:

$$H(s) = \frac{\omega_n^2}{s^2 + \frac{\omega_n}{Q} s + \omega_n^2} \quad (2)$$

$$= \frac{\omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2}. \quad (3)$$

Here, ω_n denotes the natural frequency, Q the quality factor (also called the pole Q), and ζ the damping factor. The first form is common in filter design and the second in control theory (e.g., in phase-locked loops). Noting that $Q = 1/(2\zeta)$, we will use the two forms interchangeably.

We intuitively observe that, if $Q \to \infty$, the two poles approach $\pm j\omega_n$ and the system becomes unstable. Thus, the value of Q determines how much the poles depart from the real axis and how much peaking $|H(s = j\omega)|$ has. The two poles can be expressed as

$$\omega_{p1,2} = -\frac{\omega_n}{2Q} \pm j\omega_n \sqrt{1 - \frac{1}{4Q^2}}, \quad (4)$$

taking on a complex value if $Q > 1/2$. But complex poles do not necessarily imply peaking. Writing

$$|H(j\omega)|^2 = \frac{\omega_n^4}{(\omega_n^2 - \omega^2)^2 + \left(\frac{\omega_n \omega}{Q}\right)^2}, \quad (5)$$

we find that the denominator falls to a minimum at $\omega_\alpha = \omega_n \sqrt{1 - 1/(2Q^2)}$ if $Q > \sqrt{2}/2$. In such a case, $|H(j\omega)|$ exhibits a peak equal to $Q/\sqrt{1 - 1/(4Q^2)}$ (Figure 1). It is helpful to remember that for $Q = 1$, the peaking is about 1.15 dB, and it occurs at $\omega_a = 0.71\omega_n$.

With these preliminary developments, we can now evaluate the stopband attenuation of $|H|$ at a given frequency for the case of real or complex poles. As an example, we seek the rejection at a frequency
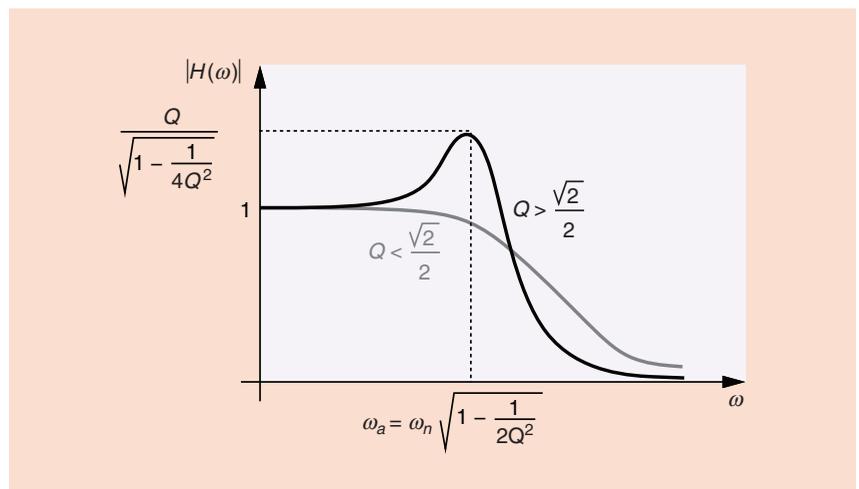
FIGURE 1: Peaking in biquad frequency response.

equal to twice the –3-dB bandwidth of the filter. This choice represents a scenario of interest in radio-frequency (RF) receivers where the adja-cent channel must be suppressed. The –3-dB bandwidth is given by

$$\omega_{-3dB}^2 = \frac{\omega_n^2}{2}\left[2 - \frac{1}{Q^2} + \sqrt{\left(2 - \frac{1}{Q^2}\right)^2 + 4}\right].$$

(6)

If $Q = 1/2$, the two poles are real and equal (as in an open-loop cascade of two first-order RC sections), $\omega_{-3dB} = 0.64\omega_n$, and the attenuation provided by $|H|$ in (5) at $2\omega_{-3dB}$ is 0.379. On the other hand, if $Q = 1$, $\omega_{-3dB} = 1.27\omega_n$ and the attenuation at $2\omega_{-3dB}$ reaches 0.166. That is, by allowing 1.15 dB of peaking at the edge of the passband, we improve the rejection at $2\omega_{-3dB}$ by $20\log(0.379/0.166) \approx 7.2\,dB$.

Realization of Complex Poles

It can be shown that a passive network consisting of only resistors and capacitors does not provide complex poles [4]. We must therefore seek active implementations that exploit feedback to create such poles.

Let us begin with the negative-feedback system shown in Figure 2, where two lossless integrators appear in the loop. Note that Y is negated as it enters the input summer. This negation can be removed if one integrator inverts and the other does not. We have for this system

$$\frac{Y}{X}(s) = \frac{k_1 k_2}{s^2 + k_1 k_2},$$

(7)

obtaining imaginary poles if $k_1 k_2 > 0$. To stabilize the system, we must add a term proportional to s in the denominator. This can be accomplished by a number of techniques, for example, 1) we can add a zero to the open-loop transfer function, as practiced in type II phase-locked loops, or 2) we can make one of the integrators lossy, e.g., we can change k_1/s to $k_1/(s+a)$. The latter is realized if a fraction of the integrator's output is returned to its input without phase shift. Illustrated in Figure 3(a), such an arrangement yields

$$\frac{B}{A}(s) = \frac{k_1}{s + \alpha k_1}.$$

(8)

The circuit implementation is straightforward [Figure 3(b)] and gives

$$\frac{V_{out}}{V_{in}}(s) = \frac{-R_2}{R_1(R_2 C_1 s + 1)}.$$

(9)

We can now incorporate the lossy integrator of Figure 3(a) in the architecture of Figure 2 [Figure 4(a)]. In
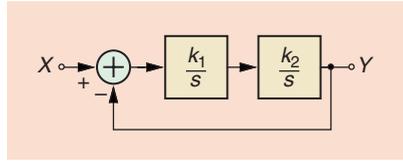
FIGURE 2: A filter using two lossless integrators in a loop.
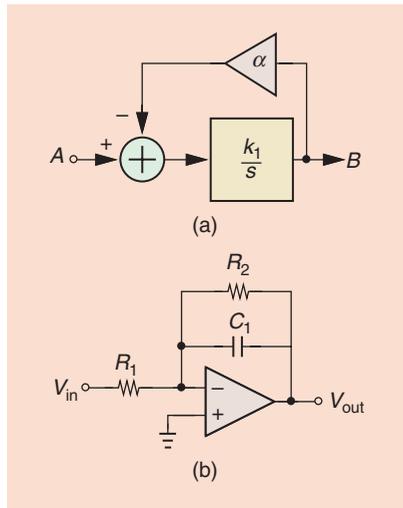
FIGURE 3: (a) A lossy integrator and (b) its circuit implementation.
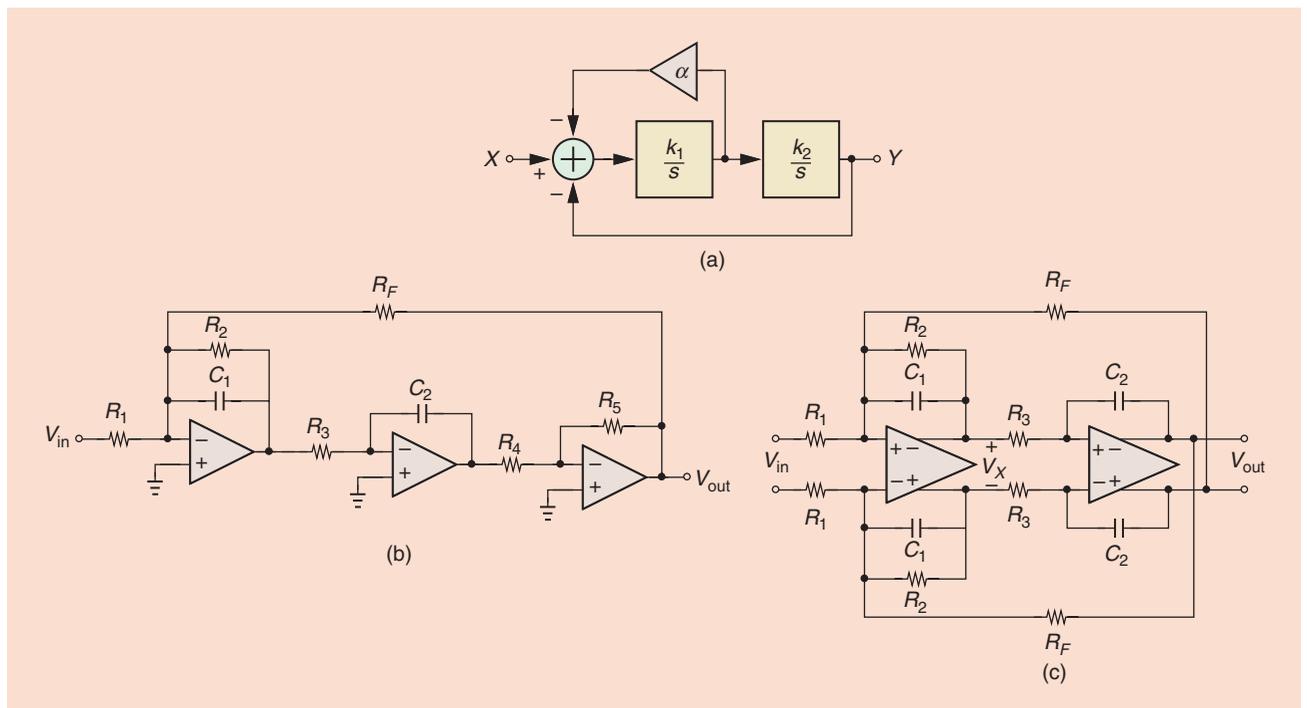
FIGURE 4: (a) The use of a lossy integrator in a biquad loop, (b) the Tow–Thomas biquad, and (c) its differential version.

early filter designs, off-the-shelf op-amps provided only a single-ended output, making it difficult to implement noninverting integrators. Thus, an inverting amplifier was inserted in the loop, leading to the circuit shown in Figure 4(b). Called the "Tow–Thomas biquad" after Tow [2], [3] and Thomas [5], this topology requires only two op-amps if implemented in fully-differential form [Figure 4(c)].

The biquad depicted in Figure 4(c) provides the following transfer function:

$$\frac{V_{\text{out}}}{V_{\text{in}}}(s)$$
$$= \frac{R_2 R_F / R_1}{R_2 R_3 R_F C_1 C_2 s^2 + R_3 R_F C_2 s + R_2}, \quad (10)$$

yielding

$$\omega_n = \frac{1}{\sqrt{R_3 R_F C_1 C_2}} \quad (11)$$

$$Q = R_2 \sqrt{\frac{C_1}{R_3 R_F C_2}}. \quad (12)$$

Interestingly, ω_n is independent of R_2 whereas $Q \propto R_2$, i.e., the pole Q can be adjusted without changing ω_n. Moreover, both ω_n and Q are independent of R_1, which sets the passband gain according to $V_{\text{out}}/V_{\text{in}} = R_F/R_1$. These attributes facilitate tuning of the filter (discussed below). Also, since the first integrator's output in Figure 4(c),

V_X, is equal to $-V_{\text{out}} R_3 C_2 s$, we write $V_X/V_{\text{in}} = -(V_{\text{out}}/V_{\text{in}}) R_3 C_2 s$ and conclude that the circuit can act as a bandpass filter as well.

To arrive at another biquad implementation, let us first note that the inverting amplifier in Figure 4(b) can be moved to the input [Figure 5(a)] without changing the transfer function. Next, we observe that the loss-inducing path in Figure 3(a) and realized by R_2 in Figure 3(b) need not return to the very input of the integrator; this path can even traverse additional stages placed before or after the integrator if such stages are free from phase shift [Figure 5(b)]. It is, therefore, possible to tie the left terminal of R_2 to the input of the inverting amplifier, but we must ensure negative feedback. These thoughts lead to the Kerwin–Huelsman–Newcomb (KHN) biquad depicted in Figure 5(c).

The KHN biquad's transfer function is given by

$$\frac{V_{\text{out}}}{V_{\text{in}}}(s) = \frac{c_1}{a_2 s^2 + b_2 s + c_2}, \quad (13)$$

where

$$c_1 = \frac{R_2(R_4 + R_5)}{R_4(R_1 + R_2)} \quad (14)$$

$$a_2 = R_1 R_3 C_1 C_2 \quad (15)$$

$$b_2 = \frac{R_1 R_3 C_2(R_4 + R_5)}{R_4(R_1 + R_2)} \quad (16)$$

$$c_2 = \frac{R_5}{R_4}. \quad (17)$$

It follows that

$$\omega_n = \sqrt{\frac{R_5}{R_4 R_1 R_3 C_1 C_2}} \quad (18)$$

$$Q = \sqrt{\frac{R_5 R_4}{R_1 R_3 C_1 C_2} \cdot \frac{R_1 + R_2}{R_1 R_3 C_3 (R_4 + R_5)}}. \quad (19)$$

In this case, too, ω_n is independent of R_2, but Q is not. Thus, the pole Q can be adjusted without affecting ω_n.

It is interesting to recognize that the Tow–Thomas and KHN biquads are fundamentally based on the same principle, namely, introducing loss in one integrator so as to control the Q. That is, Figure 4(a) embodies both topologies. However, the disadvantage of the KHN circuit is that it requires three op-amps even in fully differential implementations. For this reason, we focus on the Tow–Thomas biquad.

Sensitivity

The design of analog filters must deal with the sensitivity of the performance to component variations. For example, the −3-dB bandwidth, the Q, and the stopband rejection are functions of resistor and capacitor values used in the circuit. We generally ask, if one component value changes by, say, 10%, how much is the percentage change in a given filter parameter, e.g., in the bandwidth? If the bandwidth also changes by 10%, the
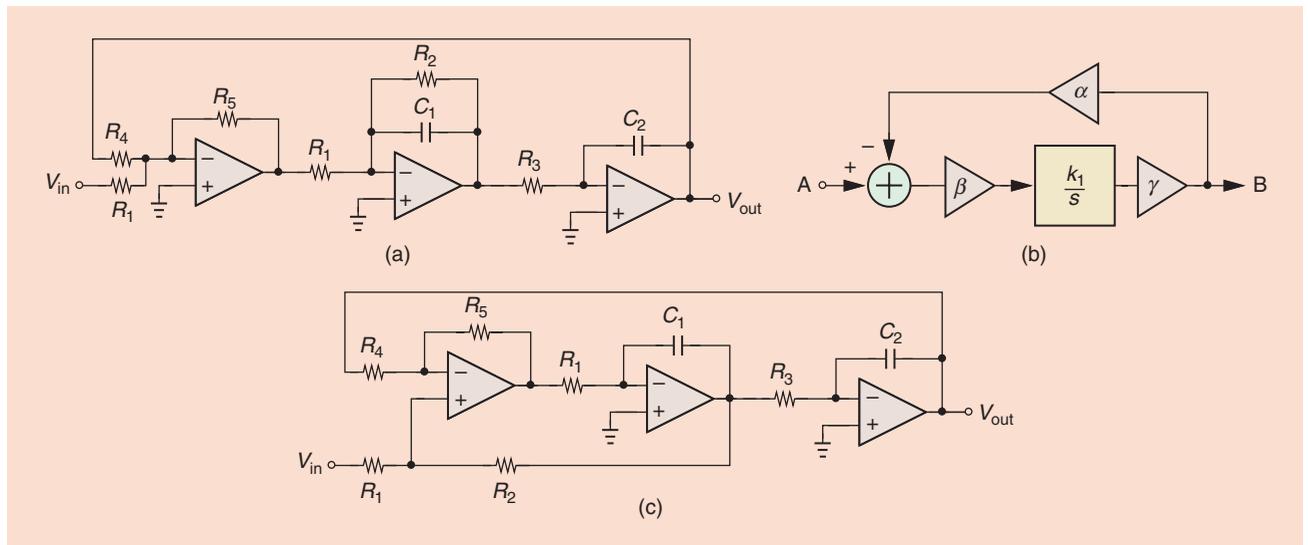
FIGURE 5: (a) The Tow–Thomas filter with the inverting amplifier moved to front end, (b) a lossy integrator with additional stages in the loop, and (c) a KHN biquad.

corresponding sensitivity is unity. For example, the sensitivity of the pole Q expressed by (12) is equal to one with respect to R_2 but lower than one with respect to R_3 or R_F.

It is shown that the sensitivities of the Tow–Thomas biquad with respect to the passive devices are equal to or less than unity [2]. The same holds for the KHN circuit [1].

Effect of Finite Op-Amp Bandwidth

The tradeoffs among gain, bandwidth, and power consumption of op-amps require that we design them for only as much bandwidth as necessary for the proper operation of a filter. We intuitively expect that the additional poles contributed by the op-amps within the feedback loop of Figure 4(c) degrade the phase margin, thus creating additional peaking in the frequency response. Called "Q enhancement" in the filter literature, this phenomenon imposes a lower bound on the op-amps' unity-gain bandwidth.

The analyses in [5] and [6] study this effect for high-Q designs. We describe a different approach that is applicable to any Q value. Our approximation is that the closed-loop filter transfer function is still in biquadratic form but with a greater Q.

Before delving into loop calculations, we must formulate the effect of the op-amps' bandwidth on the two integrators in Figure 6. Modeling the op-amp by a one-pole system, $A_0/(1 + s/\omega_0)$, we begin with the integrator in Figure 6(a) and write

$$\frac{V_{\text{out}}}{V_{\text{in}}}(s) = \frac{-A_0}{\frac{R_1 C_1}{\omega_0} s^2 + A_0 R_1 C_1 s + 1},$$

(20)

where it is assumed that $A_0 \gg 1$ and $A_0 \omega_0 \gg 1/(R_1 C_1)$. Note that $A_0 \omega_0$ is the unity-gain bandwidth of the op-amp. If the two poles of this transfer function, ω_{p1} and ω_{p2}, are widely spaced, i.e., if $\omega_{p1} \ll \omega_{p2}$, then we can apply the dominant-pole approximation and write $(1 + s/\omega_{p1})(1 + s/\omega_{p2}) = s^2/(\omega_{p1}\omega_{p2}) + (1/\omega_{p1} + 1/\omega_{p2}) s + 1 \approx s^2/(\omega_{p1}\omega_{p2}) + (1/\omega_{p1}) s + 1$. That is,

$$\omega_{p1} \approx \frac{1}{A_0 R_1 C_1}$$

(21)

$$\omega_{p2} \approx A_0 \omega_0.$$

(22)

Interestingly, the dominant pole is negligibly affected by the op-amp, and the nondominant pole is simply given by its unity-gain bandwidth.

We repeat the foregoing computation for the lossy integrator in Figure 6(b), arriving at

$$\frac{V_{\text{out}}}{V_{\text{in}}}(s) = \frac{-R_2/R_1}{\frac{R_2 C_1}{A_0 \omega_0} s^2 + R_2 C_1 s + 1},$$

(23)

where it is assumed that $R_2/R_1 + 1 \ll A_0 \gg 1$, and $A_0 R_2 C_1$ is much greater than $1/\omega_0$ and $R_2/(R_1 \omega_0)$. The dominant-pole approximation yields

$$\omega_{p1} \approx \frac{1}{R_2 C_1}$$

(24)

$$\omega_{p2} \approx A_0 \omega_0,$$

(25)

which follows the same observations made above. The key point here is that each op-amp contributes an additional phase equal to $-\tan^{-1}[\omega/(A_0 \omega_0)]$ at a frequency ω.

Let us return to the Tow–Thomas biquad of Figure 4(c) and summarize our findings. Figure 7 sketches the magnitude and phase of the biquad's open-loop transfer function, $G(s)$. The magnitude begins at $(R_2/R_1)A_0$ (the product of the two integrators'
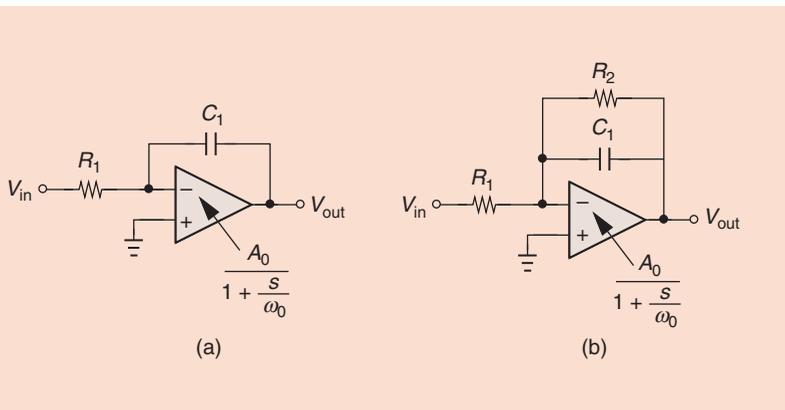
FIGURE 6: (a) Lossless and (b) lossy integrators with a one-pole op-amp model.
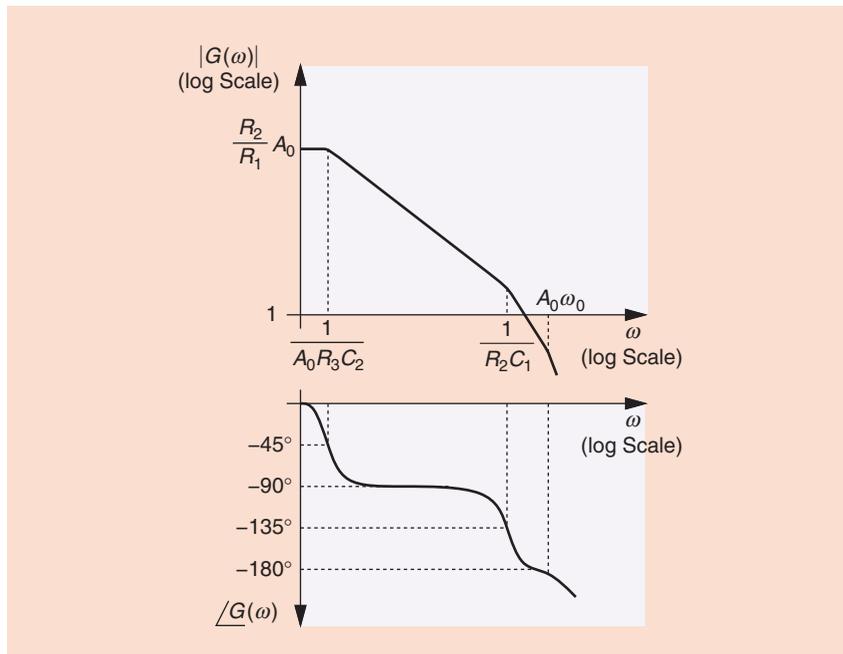
FIGURE 7: The open-loop response of the Tow–Thomas biquad.

low-frequency gains), deflects at the first pole $1/(A_0 R_3 C_2)$ (arising from the second integrator), falls at 20 dB/dec up to the second pole $1/(R_2 C_1)$, and then declines at 40 dB/dec. The two op-amp-induced poles at $A_0 \omega_0$ lie beyond the unity-gain bandwidth, ω_u. The phase reaches $-135°$ at the second pole, leading to an inadequate phase margin of 45° if $1/(R_2 C_1) < \omega_u$. In addition, the two poles at $A_0 \omega_0$ must be high enough not to affect the phase margin significantly.

Now, suppose we design a low-pass biquad that exhibits a Q of one (i.e., a peaking of 1.15 dB at $\omega_a = 0.71 \omega_n$) with ideal op-amps. We wish to determine the minimum tolerable op-amp unity-gain bandwidth that raises the peaking by, at most, 1 dB.

We first observe that a negative-feedback system such as that in Figure 4(a) has an open-loop transfer function of the form $[\omega_n/(s + 2\zeta\omega_n)](1/s)$, with a phase margin given by [7]

$$PM = \tan^{-1} \frac{2\zeta}{\sqrt{\sqrt{4\zeta^4 + 1} - 2\zeta^2}} \quad (26)$$

$$= \tan^{-1} \sqrt{\frac{2}{\sqrt{4Q^4 + 1} - 1}}. \quad (27)$$

For example, $Q = 1$ gives PM = 51.8°.

We then note from Figure 1 that a peaking of $1.15\,dB + 1\,dB = 2.15\,dB$ translates to $Q = 1.16$, which, from (27), yields a phase margin of 46.1°. That is, the additional allowable 1-dB peaking is equivalent to a phase margin degradation of $51.8° - 46.1° = 5.7°$. Thus, each op-amp must not contribute more than 2.85° of phase shift at ω_a in Figure 1:

$$\tan^{-1} \frac{\omega_a}{A_0 \omega_0} \leq 2.85°. \quad (28)$$

It follows that

$$A_0 \omega_0 \geq 20\omega_a \quad (29)$$
$$\geq 14\omega_n. \quad (30)$$

This simple rule of thumb proves useful in the design of the op-amps and can be readily revised for other Q or peaking values.

Equation (30) translates to a high-power consumption for the op-amps, but it is possible to select the filter val-

ues for a Q lower than desired and permit the op-amps to raise it. For example, if we begin with $Q = 0.5\,(PM = 76°)$ and choose $A_0 \omega_0 \approx 4.7\omega_a \approx 3.3\omega_n$, then the equivalent Q is about unity and the peaking about 1.15 dB. Such a strategy must, nonetheless, account for the other op-amp poles and the variation of Q and $A_0 \omega_0$ with the process, supply voltage, and temperature (PVT).

Noise Considerations

In many applications, a filter's noise contribution can limit the dynamic range. For example, the baseband channel-select filters in RF receivers must provide proper noise and linearity levels to reject unwanted channels while negligibly corrupting the desired signal. For this reason, the noise performance of the biquad is of interest.

In the Tow–Thomas topology of Figure 4(c), resistors R_1, R_2, and R_F contribute low-frequency input-referred noise in the amount of $2(4kTR_1 + 4kTR_1^2/R_2 + 4kTR_1^2/R_F)$, where the factor of two accounts for the resistors in the upper and lower paths. To minimize this noise we wish to reduce R_1, but at the cost of loading the stage preceding the biquad. For low-bandwidth applications, a small R_1 also translates to a large value for C_1 and hence a significant area penalty. The noise of R_3 is less critical as it is divided by the first integrator's gain if referred to the main input, a fortunate situation because this resistor loads the first op-amp and should preferably have a high value.

The op-amp noise proves problematic, too. It can be shown that the noise of the first op-amp in Figure 4(c), V_{n1}, travels to V_{out} according to the following transfer function:

$$\frac{V_{out}}{V_{n1}}(s) = \frac{R_2 C_1 s + 1 + \dfrac{R_2}{R_F \| R_1}}{R_3 R_2 C_1 C_2 s^2 + R_3 C_2 s + \dfrac{R_2}{R_F}}, \quad (31)$$

which reduces to $[1 + R_2/(R_F \| R_1)]/(R_2/R_F)$ at low frequencies. Dividing this result by the passband gain, R_F/R_1, gives an input-referred noise of $(1 + R_1/R_2 + R_1/R_F)V_{n1}$. Thus,
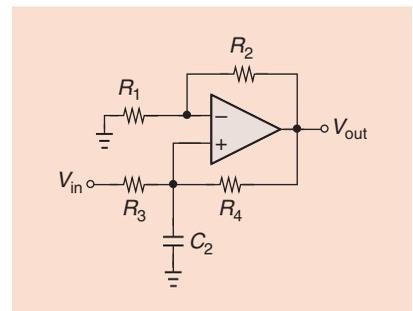
FIGURE 8: A noninverting integrator.

$R_1/R_2 + R_1/R_F$ should be chosen well below unity, i.e., we wish to select relatively large low-frequency gains for the first integrator (R_2/R_1) and for the overall biquad (R_F/R_1). The second op-amp's noise is less problematic as it is preceded by the first integrator's gain.

Tuning

The PVT dependence of resistor and capacitor values requires that filters be tuned before they operate in a system. Equations (11) and (12) suggest that tuning can proceed as follows: 1) we set ω_n by adjusting R_3, C_1, or C_2, and 2) we set Q by adjusting R_2 [the bandwidth is then defined by (6)]. The passband gain, R_F/R_1, can also be programmed by adjusting R_F, a useful property if automatic gain control is necessary, but such an adjustment must occur before the two steps outlined above. Each programmable device typically consists of a constant, coarse component in parallel or series with a variable, fine section.

Questions for the Reader

1) Figure 8 shows a noninverting integrator [1]. Derive the condition for the elements so that the circuit acts as an ideal integrator. What is the principal difficulty with this topology?
2) Suppose the Tow–Thomas biquad of Figure 4(c) senses a large, narrowband undesired channel at $\omega = \omega_{-3dB}$. Which of the two integrators produces greater voltage swings and hence experiences more nonlinearity?

in Engineering Award, the IEEE MTT Special Award, and the IEEE Region 8 Student Paper Contest Award, were given to papers with the best practice implementation. All participants had the opportunity to present posters and demonstrations of their projects. The full-day event ended with the presentation of the awards for best project. The conference was very well received by

A cake celebrating the 10th anniversary of the IEEESTEC Conference.

the attendees of about 350 high school and academic students.

More information about the IEEESTEC Conference can be found at http://ieee .elfak.ni.ac.rs.

—*Dr. Danijel Danković*
Chapter Chair
ED/SSC University of Niš Student Branch
SSC

A CIRCUIT FOR ALL SEASONS (continued from p. 15)
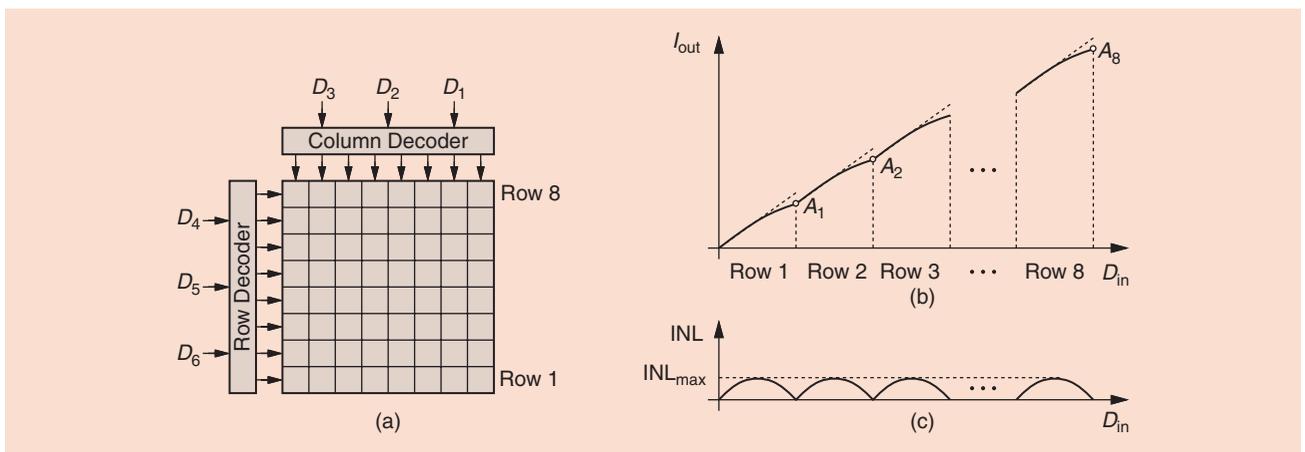
FIGURE 9: (a) A DAC matrix architecture, (b) its input–output characteristic in the presence of a gradient, and (c) the resulting INL profile.

Answers to Last Issue's Questions

1) By what factor is the integral nonlinearity (INL) of a differential current-steering digital-to-analog converter (DAC) lower than that of a single-ended topology if only the finite output impedance of the current sources is considered?

It can be proven that the normalized INLs of the single-ended and differential topologies are, respectively, given by $NR_L/(4r_O)$ and $(NR_L/r_O)^2/(12\sqrt{3})$, where N is the number of current sources, R_L the load resistance, and r_O the single-ended output resistance of each current cell. The key point here is that the differential topology's INL is proportional to the *square* of NR_L/r_O and, hence, much less than that of the single-ended counterpart.

2) In the matrix architecture of Figure 9(a), each row experiences the same gradient from left to right. If each cell current is higher than the one to its left by ΔI, what is the maximum INL?

We begin with row 1 and note that its current cells can be expressed as $I, (I+\Delta I), \ldots, (I+7\Delta I)$. Thus, if only the entire first row is activated, the output current is equal to $I + (I+\Delta I) + \cdots + (I+7\Delta I) = 8I + 28\Delta I$. The next row has the same behavior. We can therefore plot the input–output characteristic as shown in Figure 9(b), where ΔI is assumed negative. We observe that I_{out} is equal to $m(8I + 28\Delta I)$ if m rows are activated. That is, points $A_1, A_2, \ldots, A_8$ lie on a straight line, and the INL is simply the difference between the actual characteristic and this line [Figure 9(c)]. The maximum INL occurs in the middle of each row and is given by $8\Delta I$. This periodic INL behavior often manifests itself in mea-

surements and signifies the existence of gradients on the chip.

References

[1] W. J. Kerwin, L. P. Huelsman, and R. W. Newcomb, "State-variable synthesis for insensitive integrated circuit transfer functions," *IEEE J. Solid-State Circuits*, vol. 2, pp. 87–92, Sept. 1967.
[2] J. Tow, "Active RC filters: A state-space realization," *Proc. IEEE*, vol. 56, pp. 1137–1139, June 1968.
[3] J. Tow, "A step-by-step active filter design," *IEEE Spectr.*, vol. 6, pp. 64–68, Dec. 1969.
[4] N. Balabanian, *Network Synthesis*. Englewood Cliffs, NJ: Prentice-Hall, 1958.
[5] L. C. Thomas, "The biquad. Part I: Some practical design considerations," *IEEE Trans. Circuit Theory*, vol. 18, pp. 350–357, May 1971.
[6] D. Ackerberg and K. Mossberg, "A versatile active RC building block with inherent compensation for the finite bandwidth of the amplifier," *IEEE Trans. Circuits Syst.*, vol. 21, pp. 75–78, Jan. 1974.
[7] G. F. Franklin, J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*. Upper Saddle River, NJ: Prentice Hall, 2002.

SSC