Chapter 4

STABILITY FREQUENCY AND STEP RESPONSE

In the previous chapter we enunciated the basic feedback concepts and described efficient techniques for analysing feedback amplifiers. Particularly, we defined the open-loop gain, the loop-gain (or return ratio), and other quantities, all as *DC values.* However, these quantities are in general a function of frequency and they should be better referred to as *transfer functions* instead of gains. Moreover, the feedback factor could also be frequency dependent (to this end, the best example is perhaps the well-known RC-active integrator made up of an op-amp and a feedback network constituted by a resistor and a capacitor). Thus, all these effects should be taken into account in the Rosenstark and Choma relationships, (3.8) and (3.16), which allow us to accurately obtain the closed-loop transfer function. In addition, for a first-order model, they should also be considered in (3.1). Similarly, in the Blackman equations, (3.23) and (3.24), we need to use the appropriate return ratio transfer function to obtain input and output impedances instead of resistances.

For the sake of simplicity, in this chapter we will assume that the feedback factor is constant, at least in the frequency range of interest. In addition, we will assume that the feedback network is designed so as to not introduce further poles in the loop gain. Such a condition is fortunately often verified in feedback amplifiers with a purely resistive feedback network.

It should be well known to the reader that an electronic circuit and system are said to be *stable* if all bounded excitations yield bounded responses. Otherwise, if bounded excitations produce an unbounded response the system is said to be *unstable.* Passive RLC circuits are by nature stable. Active networks contain internal energy sources that can combine with the

input excitation to cause the output to increase indefinitely or sustain oscillations. Note, however, that in practice the output of an unstable circuit cannot diverge *indefinitely,* since a limit is set by the power supply rails.

It should also be well known that stability is ensured if all the poles of a given circuit/system lie in the left-half of the *s*-plane. Thus, we could check the stability of a feedback amplifier by evaluating the closed-loop transfer function and determining the locations of its poles. This procedure, however, does not provide design insides and does not specify the margins by which stability is achieved. In fact, circuit components are affected by manufacturing tolerances, temperature and ageing phenomena, etc., which cause a parameter to deviate from its nominal value. Under this scenario, we need to introduce safety stability margins, which are the *phase margin* and *gain margin.* Moreover, even stable amplifiers, hence that have a bounded response, can take too much time to reach a steady state. For this purpose, the classical feedback circuit analysis technique derived from the well-known Bode disclosures can be utilised [B45].

In the following paragraphs we will examine the frequency response of transfer functions characterised by different combinations of poles (and zeros) that are found usually in real practice. Starting from this, useful definitions will be given which help designers to derive fundamental relations to ensure closed-loop stability with adequate margins. The closed-loop step response in the time domain, for each typology of transfer function, is also derived.

4.1 ONE-POLE FEEDBACK AMPLIFIERS

Among the feedback properties, the closed-loop bandwidth extension to the original open-loop amplifier is often included [G85], [SS91]. We will show that this property applies only to one-pole amplifiers, but is not effective in multi-pole amplifiers.

Let us consider an open-loop amplifier having the following transfer function including a single (negative) pole, whose angular frequency is p_1

$$A(s) = \frac{A_o}{1 + \dfrac{s}{p_1}} \tag{4.1}$$

Now connect the amplifier in feedback with a pure resistive network, whose feedback factor is *f,* as shown in Fig. 4.1.

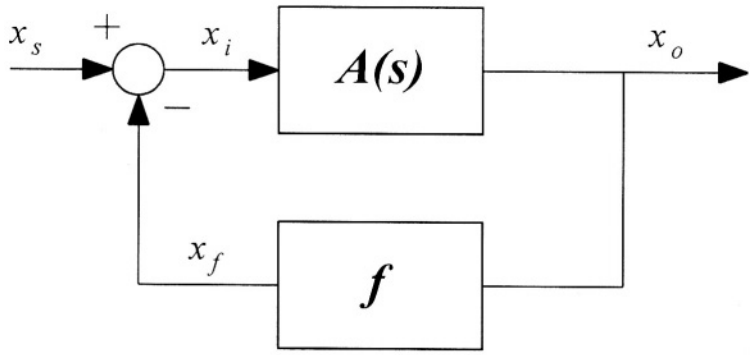

Fig. 4.1. Block diagram of a feedback amplifier.

Returning to (3.1a), the closed-loop transfer function results as

$$G_F(s) = \frac{G_{Fo}}{1 + \dfrac{s}{p_{F1}}} \tag{4.2}$$

where G_{Fo} is the DC closed-loop gain and p_{F1} is the closed-loop pole, each given by

$$G_{Fo} = \frac{A_o}{1 + fA_o} = \frac{1}{f}\frac{T_o}{1 + T_o} \approx \frac{1}{f} \tag{4.3a}$$

$$p_{F1} = (1 + fA_o)p_1 = (1 + T_o)p_1 \approx T_o p_1 \tag{4.3b}$$

The foregoing approximations hold for large loop gains (T_o>>1), which are required for an adequate desensitisation of the closed-loop response with respect to open-loop parameters. It is seen that increasing T_o from zero shifts the pole along the negative real axis, as illustrated in Fig. 4.2[1]. Since the pole is located in the negative s-plane for any value of f, the system is termed *absolutely* or *unconditionally* stable. This denotes an attractive condition indicating that a one-pole amplifier is stable under all input signal conditions

[1] This plot is called the root locus diagram. Its construction can become tedious for higher order systems and we do not make use of this tool to examine stability. The interested reader is referred to [SS91], [G85], or any feedback control text e.g. [FPE94],

and for all ranges of component values. Unfortunately, this is not a realistic case, since real amplifiers have more than one pole.

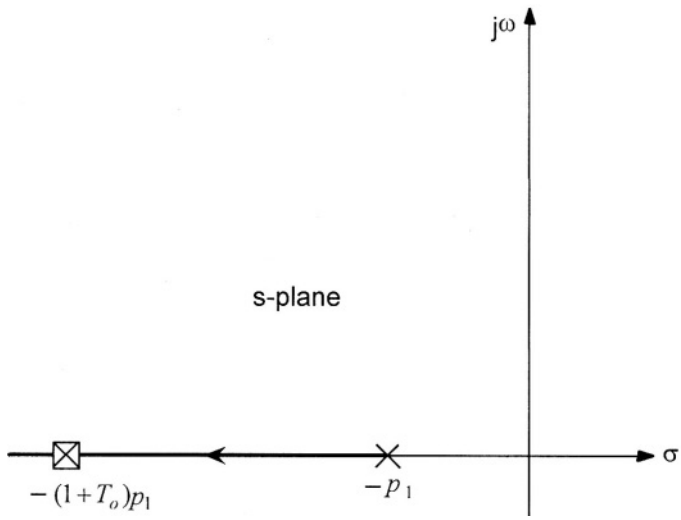

Fig. 4.2. Effect of feedback on the pole location for a single-pole amplifier.

Returning to (4.3b), we see that the closed-loop pole has been shifted to a higher frequency by a factor equal to $1+T_o$ (that is, approximately the DC loop gain), which is the same amount of reduction experienced by the closed-loop DC gain with respect to the open-loop gain. Thus, a gain bandwidth trade-off exists between the open- and closed-loop transfer functions, indicating that in a one-pole amplifier we can apply feedback to obtain higher bandwidth where amplifier gain reduction is allowed. This trade-off is represented by the *gain-bandwidth product,* ω_{GBW}, which is the product of the DC open-loop gain, A_o, and its –3-dB angular frequency p_1. Note also that the gain-bandwidth product of a single-pole function exactly equals its *unity-gain frequency,* ω_T (i.e., the frequency at which the module of the gain becomes unitary, for this reason it is also called the *transition* frequency). Moreover, ω_{GBW} is an invariant amplifier parameter, since its value is the same for the open-loop and closed-loop amplifier, as illustrated in Fig. 4.3, showing the open-loop, closed-loop and loop-gain transfer functions. Of course, the gain-bandwidth product of $A(s)$ is independent of the degree of feedback applied and is equal to the maximum bandwidth achieved with the unitary feedback factor, $f = 1$ (i.e., with the amplifier in unity gain feedback configuration). More interestingly, (4.3b) predicts that the gain-bandwidth product of the loop-gain transfer function will equal the closed-loop pole. Thus, when studying the stability of a feedback amplifier,

it is usually convenient to analyse the frequency response of the loop-gain rather than that of the open-loop transfer function. This is because the loop-gain transfer function gives information on most of the closed-loop properties.

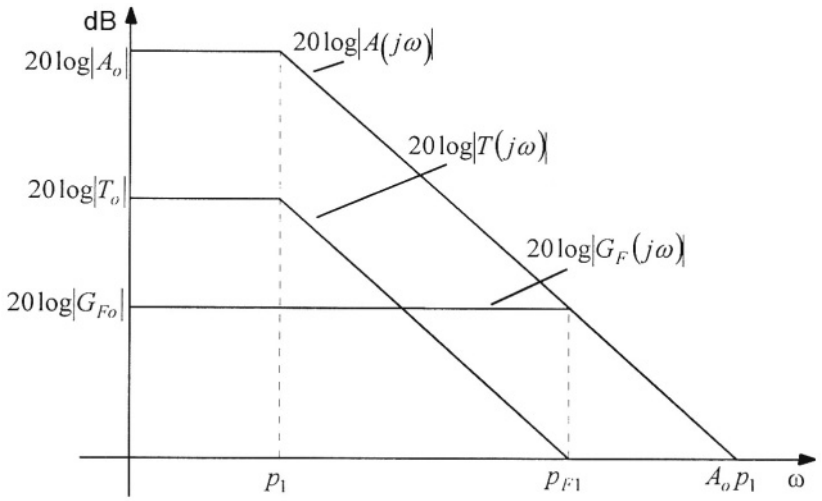

Fig. 4.3. Relation between the open-loop and closed-loop transfer function magnitudes.

The closed-loop characteristics of an amplifier can be also investigated in the time domain, by evaluating the response, $x_o(t)$, to a unitary input step *u*(t). The step response gives important specifications for applications (such as instrumentation, control, and sample data systems) sensitive to the amplifier's transient response.

Let us consider a closed-loop configuration comprising the single-pole amplifier with the transfer function shown in (4.1). The step response is easily found to be

$$x_o(t) = \left(1 - e^{-p_{F1}t}\right) G_{Fo} u(t) \tag{4.4}$$

The output steady state value is G_{Fo} and is reached exponentially with time constant $1/p_{F1}$.

The reader should know that the *settling time*, t_s, is the time interval required for the output response to settle to some specified percentage of the final value. For a single-pole amplifier the settling time is then proportional to p_{F1}. Usually, the settling time needed to reach 1% or 0.1% of the final value is considered. In these two cases, the settling time results $4.6/p_{F1}$ and $6.9/p_{F1}$.

4.2 TWO-POLE FEEDBACK AMPLIFIERS

The loop-gain transfer function of real amplifiers includes more than one single pole. In the absence of suitable compensation, this can cause instability phenomena even under negative feedback. To demonstrate these instability problems, and the related importance of a sufficiently large separation between the two lowest poles of the loop-gain transfer function, consider now an open-loop amplifier with two real negative poles. As already mentioned, and as will be further explained in the following chapters, it is more convenient to analyse the loop gain instead of the amplifier open-loop gain.

Assume that the amplifier operating within the given feedback network gives rise to the following two-pole loop-gain function

$$T(s) = \frac{T_o}{\left(1 + \frac{s}{p_1}\right)\left(1 + \frac{s}{p_2}\right)} \tag{4.5}$$

It should now be observed that for second-order and higher-order transfer functions, the gain-bandwidth product, $\omega_{GBW} = T_o p_1$, does not necessarily coincide with the unity-gain frequency ω_t. However, this is still a good approximation if the second pole is greater than ω_{GBW}. This observation is graphically explained in Fig. 4.4, where two loop-gain functions (with different pole separation) are plotted.

Note that hereinafter and unless differently indicated, we will denote the gain-bandwidth product of the loop-gain transfer function with ω_{GBW}

Using (3.la), the closed-loop transfer function becomes

$$G_F(s) = \frac{G_{Fo}}{1 + 2\frac{\xi}{\omega_o}s + \frac{s^2}{\omega_o^2}} = \frac{G_{Fo}}{\left(1 + \frac{s}{p_{F1}}\right)\left(1 + \frac{s}{p_{F2}}\right)} \tag{4.6}$$

where ω_O and ξ are the parameters called *pole frequency* and *damping factor,* respectively, and are expressed as functions of the loop-gain poles by [SS91]

$$\omega_o = \sqrt{p_1 p_2 (1 + T_o)} \tag{4.7}$$

$$\xi = \frac{p_1 + p_2}{2\omega_o} \approx \frac{1}{2\sqrt{T_o}}\left(\sqrt{\frac{p_1}{p_2}} + \sqrt{\frac{p_2}{p_1}}\right) \tag{4.8}$$

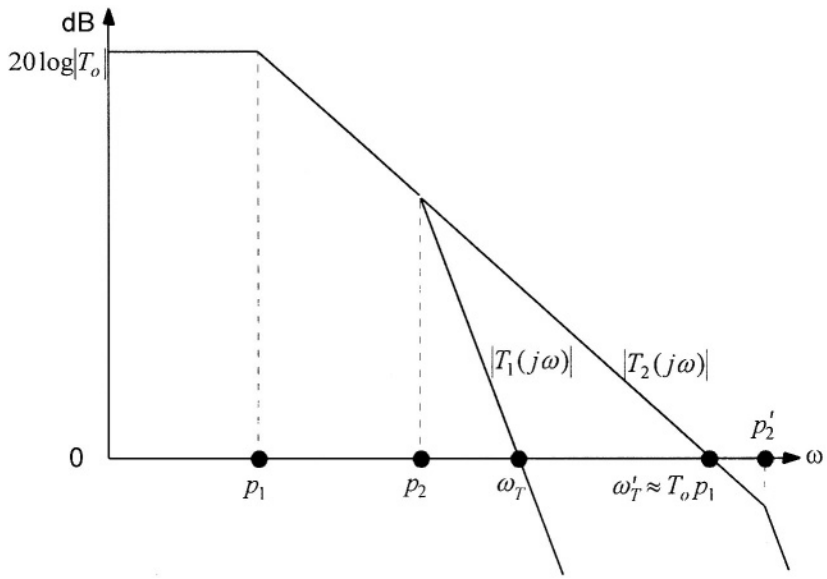

Fig. 4.4. Magnitude of two second-order loop gain transfer functions characterised by two different second poles p_2 and p_2'. Only in the second case does the gain-bandwidth product coincide with the transition frequency.

The last identity in (4.6) is an alternative expression of $G_f(s)$, with p_{f1} and p_{f2} as the closed-loop amplifier poles given by

$$p_{f1,2} = \omega_o\left(\xi \pm \sqrt{\xi^2 - 1}\right) \tag{4.9}$$

These poles are either real or complex conjugate pairs, depending on the value of ξ (or parameter Q equal to $1/(2\xi)$, sometimes used instead of ξ and called the *pole Q factor*). The location of the closed-loop poles, as the DC loop-gain is increased from zero, is illustrated in Fig. 4.5 showing that a second-order feedback system is absolutely stable. However, the design of a second-order system having a specific and well-defined frequency and transient response requires careful consideration of where the poles are to be located.

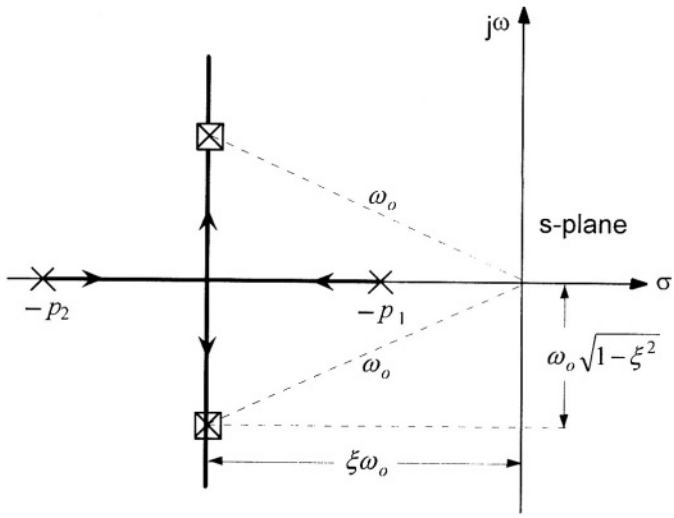

Fig. 4.5. Effect of feedback on the pole location for an amplifier with two real poles.

By normalising the module of the closed-loop transfer function and the angular frequency to G_{Fo} and ω_o, respectively, we obtain the frequency responses plotted in Fig. 4.6.

This figure helps to visualise that when ξ is lower than a critical value (namely $1/\sqrt{2}$), an overshoot in the frequency domain arises at a frequency, ω_{cp}, with the (peak) amplitude, G_{Fp}, both given below [MG91]

$$\omega_{cp} = \omega_o\sqrt{1-2\xi^2} \tag{4.10}$$

$$G_{Fp} = G_{Fo}\frac{1}{2\xi\sqrt{1-\xi^2}} \tag{4.11}$$

It is apparent that the relative amplitude of the overshoot depends only on the damping factor, ξ. For $\xi = 1/\sqrt{2}$ it can be shown that the module of the frequency response is maximally flat (which is often referred to as the *Butterworth* condition). Specifically, this condition yields the largest possible closed-loop 3-dB bandwidth within the constraint of a monotone decreasing frequency response.

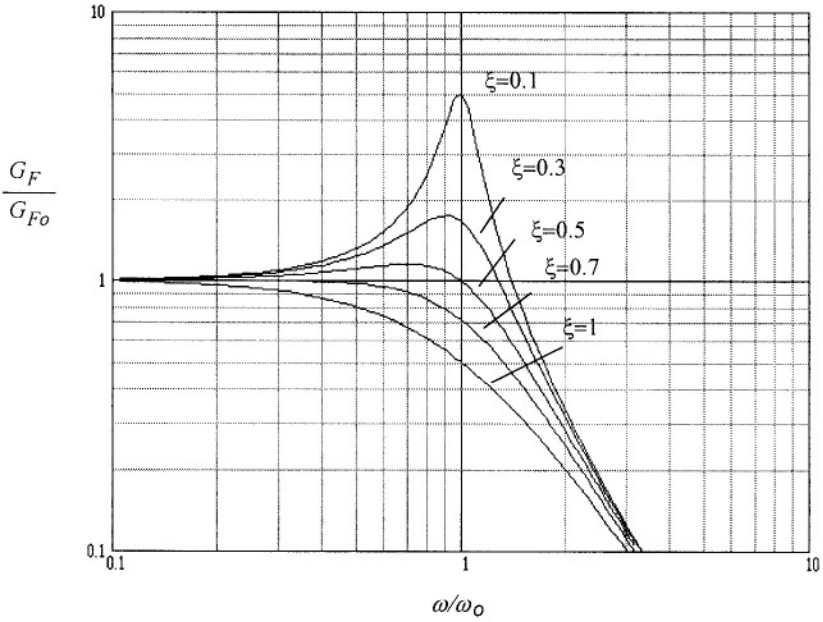

Fig. 4.6. Normalised module of the closed-loop frequency-response of a two-pole amplifier, for different values of the damping factor.

The expression of the step response of the closed-loop two-pole amplifier is

$$x_o(t) = \left[1 - \frac{1}{p_{F1} - p_{F2}}\left(p_{F2}\mathrm{e}^{-p_{F1}t} - p_{F1}\mathrm{e}^{-p_{F2}t}\right)\right] G_{Fo}u(t) \qquad (4.12)$$

If the closed-loop poles are complex conjugate –a condition which arises when the value of ξ is lower than 1– the step response exhibits an *underdamped* behavior (conversely, an *overdamped* closed-loop response requires $\xi > 1$). In such cases the step response is better expressed by the following relationship

$$x_o(t) = \left\{1 - \left[\frac{\xi}{\sqrt{1-\xi^2}}\sin\left(\sqrt{1-\xi^2}\,\omega_o t\right) + \cos\left(\sqrt{1-\xi^2}\,\omega_o t\right)\right]\mathrm{e}^{-\xi\omega_o t}\right\} G_{oF}u(t) =$$

$$= \left\{1 - \left[\sqrt{\frac{1}{1-\xi^2}}\cos\left(\sqrt{1-\xi^2}\,\omega_o t - \frac{\xi}{\sqrt{1-\xi^2}}\right)\right]\mathrm{e}^{-\xi\omega_o t}\right\} G_{oF}u(t) \qquad (4.13)$$

Underdamped amplifiers are not unstable systems, but nonetheless they are usually unacceptable, because overshoot arises in the time domain which is responsible for slow settling behaviour.

Normalising the step response to $u(t)$, we can draw the plots in Fig. 4.7, illustrating the step response of a two-pole feedback amplifier for different values of ξ, versus $\omega_o t$.

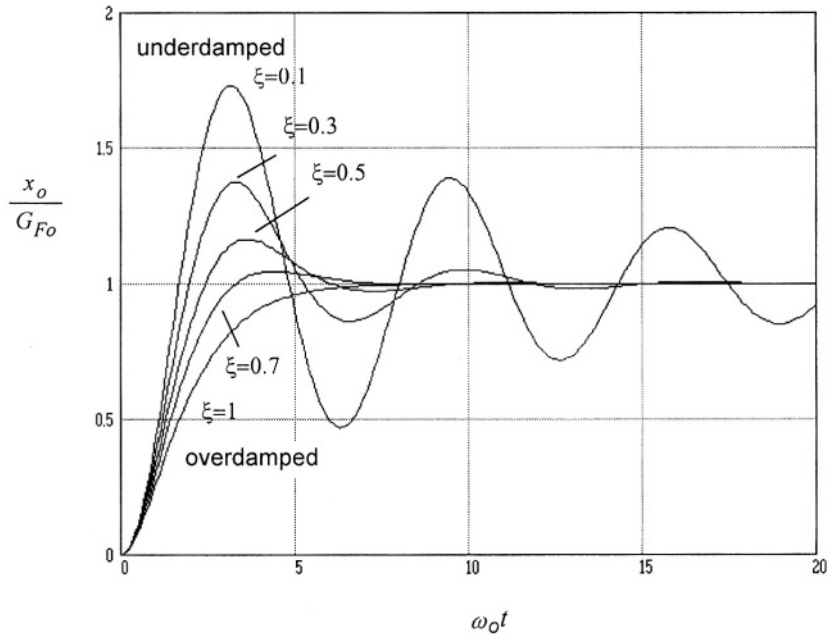

Fig. 4.7. Normalised step response of a two-pole feedback amplifier for different values of the damping factor.

To maintain peaking in both the frequency and step responses below a desired value, parameter ξ must be properly set. To this end, relationship (4.8) implicitly provides the required relation between the two (open-loop) poles for a given value of ξ and T_o. In order to avoid excessive underdamping, open-loop amplifiers must be designed with a dominant pole and a second pole at a frequency higher than the gain-bandwidth product of the loop gain (i.e., $p_2 > T_o p_1$). Thus, to analyse and design feedback amplifiers, it is useful to introduce a new parameter called the *separation factor, K,* which is the ratio between the second pole and the gain-bandwidth product of the return ratio $T(s)$

$$K = \frac{p_2}{\omega_{GBW}} = \frac{p_2}{T_o p_1} \tag{4.14}$$

The separation factor is strictly related to a parameter of the loop gain commonly used to measure the degree of stability of a feedback system namely, the *phase margin*[2], ϕ. Indeed, the phase margin is defined as 180° plus the phase of the return ratio evaluated at the transition frequency, ω_T. Figure 4.8 illustrates how the phase margin is determined on the Bode plots[3] of a second-order transfer function.

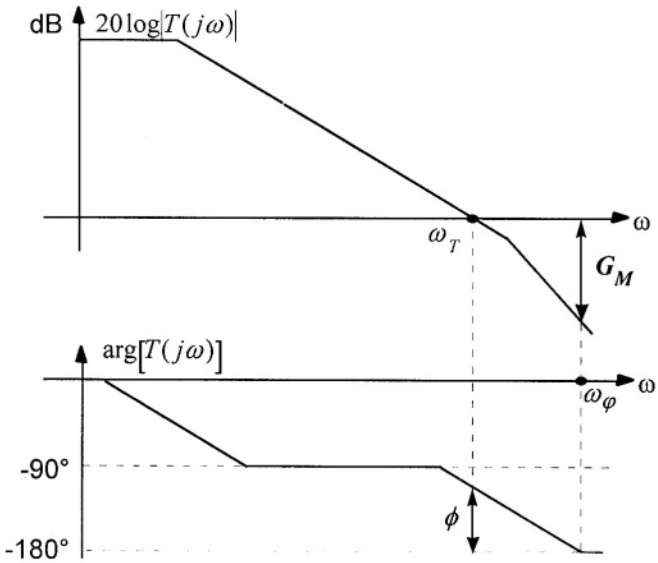

Fig. 4.8. Bode plots of a two-pole system and graphical definition of phase and gain margins.

For a second-order system with negative poles we have

$$\phi = 180° - \tan^{-1}\frac{\omega_T}{p_1} - \tan^{-1}\frac{\omega_T}{p_2} = \tan^{-1}\frac{p_1}{\omega_T} + \tan^{-1}\frac{p_2}{\omega_T} \qquad (4.15)$$

In a well-designed amplifier T_o is larger than unity and the condition $p_2 > T_o p_1$ must also hold. Thus, ω_{-3dB} is about equal to the transition frequency, ω_T, and $\tan^{-1}(p_1/\omega_T) \approx 0$. Then (4.14) is reduced to

---

[2] Another parameter, less frequently utilised by electronic designers, is the *gain margin*, G_M, defined as the difference between the gain 20 log $|T(\omega_\varphi)|$ and 0 dB, where ω_φ is the frequency at which the phase equals –180°.

[3] We assume that the reader is familiar with the Bode plot technique. For a review of this method see for instance [SS91], [G85].

$$K \approx \tan\phi \tag{4.16}$$

indicating that in a two-pole amplifier K is almost equal to the trigonometric tangent of the phase margin. In other words, for a target phase margin, we obtain through (4.16) the value of the separation factor required during the compensation design step.

From the above it derives that to design and analyse feedback amplifiers it is more convenient to represent the closed-loop transfer function, $G_F(s)$, as a function of the gain-bandwidth product, ω_{GBW}, and the separation factor, K [PP982]. Indeed, the conventional parameters, ω_o and ξ (or parameter Q) traditionally used in feedback systems, have been found very useful in designing and analysing filters, but are less effective in the context of feedback amplifiers. This because, unlike ω_{GBW} and K, which are parameters of the loop gain, ω_o and ξ are parameters related to the closed-loop amplifier. But designer effort is mainly (if not exclusively) focused on properly setting the open-loop amplifier parameters in order to achieve the closed-loop specifications. In addition, the new representation provides a simple vehicle for characterising feedback systems. Indeed, the pole frequency and the damping factor can be expressed as

$$\omega_o = p_1\sqrt{KT_o(1+T_o)} \approx \omega_{GBW}\sqrt{K} \tag{4.17}$$

$$\xi = \frac{1+KT_o}{2T_o\sqrt{K}} \approx \frac{\sqrt{K}}{2} \tag{4.18}$$

Upon inserting (4.17) and (4.18) into (4.6), the closed-loop transfer function becomes

$$G_F(s) = \frac{G_{Fo}}{1+\dfrac{s}{\omega_{GBW}}+\dfrac{s^2}{\omega_{GBW}^2 K}} \tag{4.19a}$$

or equivalently

$$G_F(s') = \frac{G_{Fo}}{1+s'+\dfrac{s'^2}{K}} \tag{4.19b}$$

where the complex frequency s' is the complex frequency s normalised to ω_{GBW}.

The normalised overshoot frequency and correspondent peak (as functions of ω_{GBW} and K) are now determined to be

$$\omega_{cp} = \omega_{GBW}\sqrt{K(1-K)} \tag{4.20a}$$

$$\frac{G_{Fp}}{G_{Fo}} = \frac{2}{\sqrt{K(4-K)}} \tag{4.20b}$$

The magnitude of the frequency response normalised to G_{Fo} versus ω/ω_{GBW} for different values of K, is plotted in Fig. 4.9. It can be noted that condition $K = 2$ i.e., $p_2 = 2\omega_{GBW}$, means a maximally flat frequency response. Moreover, for a given ω_{GBW}, the bandwidth diminishes for decreasing values of K.

The poles of the closed-loop amplifier can be also expressed as

$$p_{F1,2} = \omega_{GBW}\left[\frac{K}{2} \pm \sqrt{\left(\frac{K}{4}-1\right)K}\right] \tag{4.21}$$

and the response to an input unitary step (assuming underdamped behaviour) is

$$x_o(t) = \left\{1 - \left[\sqrt{\frac{K}{4-K}}\sin\left(\sqrt{K-\frac{K^2}{4}}\omega_{GBW}t\right) + \cos\left(\sqrt{K-\frac{K^2}{4}}\omega_{GBW}t\right)\right]e^{-\frac{K}{2}\omega_{GBW}t}\right\}G_{oF}u(t) =$$

$$= \left\{1 - \left[\sqrt{\frac{4}{4-K}}\cos\left(\sqrt{K-\frac{K^2}{4}}\omega_{GBW}t - \sqrt{\frac{K}{4-K}}\right)\right]e^{-\frac{K}{2}\omega_{GBW}t}\right\}G_{oF}u(t) \tag{4.22}$$

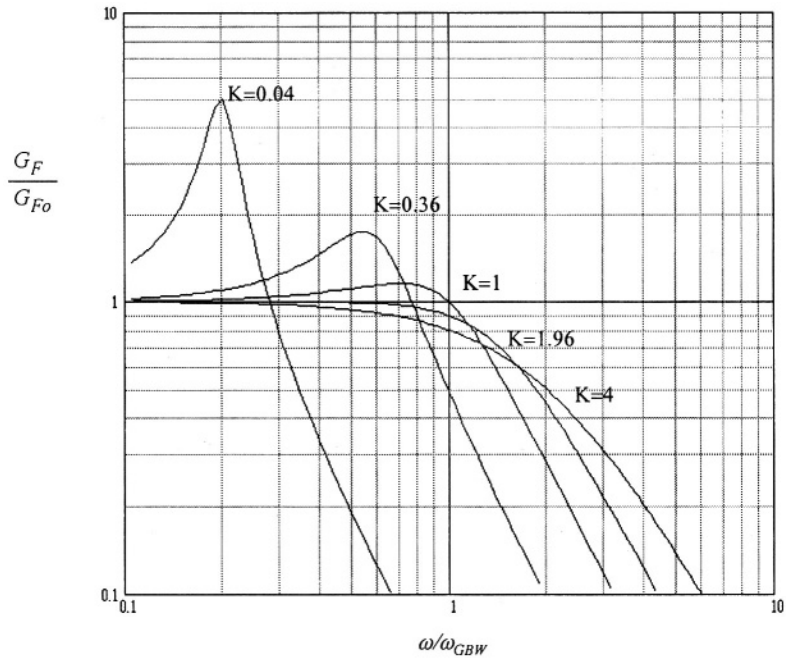

Fig. 4.9. Normalised module of the closed-loop frequency response for a two-pole feedback amplifier versus normalised frequency, for different values of K.

The step response versus time normalised to ω_{iBW} for different values of K, is plotted in Fig. 4.10.

To optimise the closed-loop amplifier step response, useful information for the designer are the time, t_p, when the *first* peak occurs (i.e., the time at which the first time derivative of $x_o(t)$ becomes zero) and its overshoot, D, [YA90], that are given by

$$t_p = \frac{2\pi}{\omega_{GBW}\sqrt{4K - K^2}} \tag{4.23}$$

$$D = e^{-\pi\sqrt{\frac{K}{4-K}}} \tag{4.24}$$

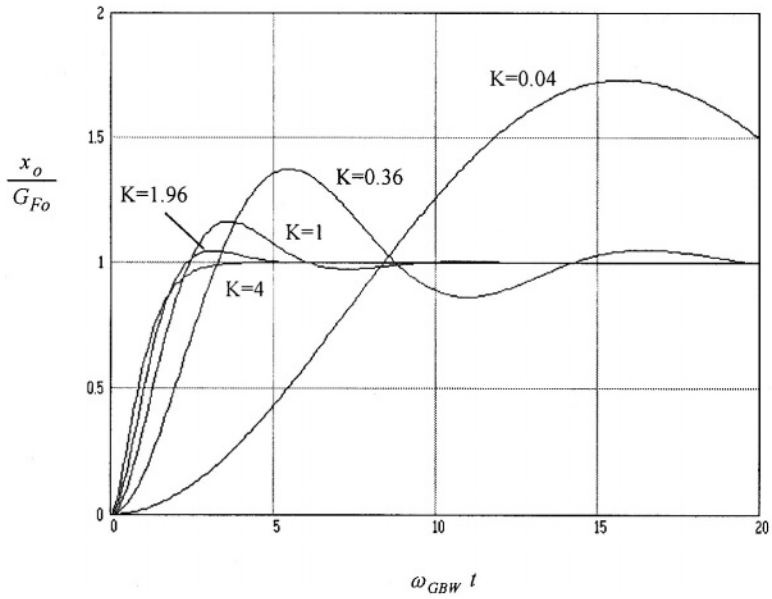

Fig. 4.10. Step response for a two-pole feedback amplifier versus $\omega_{GBW}t$ for different values of parameter K.

Like for the peaking amplitude in the frequency domain, the overshoot amplitude in the time domain depends only on the value of K. Relationships (4.23) and (4.24) are useful for optimising design in the time domain. Equation (4.24) gives the value of K for a specified settling error, and from (4.23) we determine the gain-bandwidth product needed by the settling time required. For instance, obtaining a step response to within 1% means K = 2.73. From (4.16) this value corresponds to a phase margin of about 70°. Then, if 1% settling is to be achieved within a time period not greater than 100 ns, the required gain-bandwidth product is $2\pi \cdot 5.37$ Mrad/s.

It should now be pointed out that in real amplifiers the second pole is generally fixed by design and topology constraints. Subsequently, the requirement on parameter K (or equivalently on the phase margin) indicates the gain-bandwidth we must provide to the loop-gain transfer function to ensure an adequate stability margin. To this end, as shall be discussed in detail in the next chapter, we have to properly reduce the dominant pole of the open-loop amplifier. This mandatory operation drastically reduces the high-frequency capability of the feedback amplifier, which, if operated in open-loop conditions, is characterised by a high-sensitive gain, but has its maximum bandwidth potential limited by the frequency of the second pole. As a consequence, the bandwidth improvement caused by the feedback is

effectively achieved only in one-pole amplifiers. However, these are somewhat an abstraction, since real architectures –even single-stage ones– exhibit multiple poles. Bandwidth extension is, therefore, not such a general and effective property as commonly reported. Actually, amplifiers with the highest frequency performance (e.g., RF amplifiers) all adopt open-loop topologies.

4.3 TWO-POLE FEEDBACK AMPLIFIERS WITH A POLE-ZERO DOUBLET

The loop gain of real amplifiers can include a pole-zero doublet beside two significant poles. Usually, a doublet arises from imperfect pole-zero or feed-forward compensation due to process tolerances [KM74], [BAR80], [PP95], or is caused by the frequency limitation of current mirrors when they are used to provide a differential-to-single conversion [GPP99].

The degradation in the settling performance of a one-pole amplifier with a pole-zero doublet was first discussed in [KMG74]. The effect of the doublet in a class AB one-pole amplifier was then analysed for both the settling and slewing time periods in [S91], [SY94]. However, extending the results in [KMG74] to two-pole amplifiers is not as straightforward as sometimes reported [GM74], [LS94], [EH95].

A simpler representation of a two-pole amplifier with a pole-zero doublet was proposed in [PP992]. The approach is based on the consideration that, in practice, the pole and the zero forming the doublet are often very close. In addition, they are usually located at a frequency around or greater than ω_{GBW}. Thus, such a doublet leaves the loop-gain unity-gain frequency almost unchanged, but can considerably alter the phase margin. We now demonstrate that a two-pole amplifier with such a pole-zero doublet can be modeled by an *equivalent* pure two-pole amplifier with a modified second pole.

Consider an amplifier whose loop-gain transfer function includes two poles and a pole-zero doublet (p_d and z_d), as given below

$$T(s) = T_o \frac{1}{\left(1+\frac{s}{p_1}\right)\left(1+\frac{s}{p_2}\right)} \cdot \frac{1+\frac{s}{z_d}}{1+\frac{s}{p_d}} \tag{4.25}$$

Without loss generality, assume p_1 to be the lowest frequency pole (remember that a dominant-pole behavior is mandatory to achieve stability). Phase margin evaluation of (4.25) gives

$$\phi = \tan^{-1}\frac{p_2}{\omega_T} + \tan^{-1}\frac{p_d}{\omega_T} - \tan^{-1}\frac{z_d}{\omega_T} \tag{4.26}$$

From the considerations regarding the location of p_d and z_d made above, the transition frequency, ω_T, can be assumed to be equal to the gain-bandwidth product, ω_{GBW}, (which formally represents the unity-gain frequency of a one-pole amplifier) so that (4.26) can be rewritten as

$$\phi = \tan^{-1}\frac{p_2}{\omega_{GBW}} + \tan^{-1}\frac{p_d}{\omega_{GBW}} - \tan^{-1}\frac{z_d}{\omega_{GBW}} \tag{4.27}$$

Nevertheless, if we want to accurately evaluate the deviation of ω_T from ω_{GBW} we can use the following results.

By using parameter K defined in (4.14), and the trigonometric identity

$$\tan(a+b+c) = \frac{\tan a + \tan b + \tan c - \tan a \tan b \tan c}{1 - \tan a \tan b - \tan a \tan c - \tan b \tan c} \tag{4.28}$$

from relationship (4.27) we get

$$\tan(\phi) = K\left[1 + \frac{\left(K + \frac{1}{K}\right)\frac{p_d}{\omega_{GBW}}\Delta}{1 + \left(\frac{\omega_d}{\omega_{GBW}}\right)^2 - K\frac{p_d}{\omega_{GBW}}\Delta}\right] \tag{4.29}$$

where parameter Δ is the spacing of the doublet normalised to its pole and ω_d is the doublet average frequency (evaluated as the geometric mean between p_d and z_d), respectively, defined by

$$\Delta = \left(1 - \frac{z_d}{p_d}\right) \tag{4.30}$$

$$\omega_d = \sqrt{p_d z_d} \tag{4.31}$$

Normalising the doublet frequency, ω_d to ω_{GBW}

$$\omega_{dn} = \frac{\omega_d}{\omega_{GBW}} \tag{4.32}$$

relationship (4.29) can be written

$$\tan(\phi) = K\left[1 + \frac{\left(K + \frac{1}{K}\right)\omega_{dn}\Delta}{\left(1+\omega_{dn}^2\right)\sqrt{1-\Delta} - K\omega_{dn}\Delta}\right] \tag{4.33}$$

By inspection of (4.33) we note that the second term in square brackets represents the change caused by the doublet in the tangent of the phase margin. Of course, if $\Delta = 0$ pole p_d perfectly matches zero z_d and (4.33) simplifies to (4.16). Moreover if z_d is greater (lower) than p_d, Δ is negative (positive), and the effect of the doublet is to decrease (increase) the phase margin compared to the same two-pole system without the doublet.

From the above it derives that we can model the two-pole amplifier with a pole-zero doublet by using an equivalent pure two-pole amplifier having the same gain-bandwidth product (i.e., gain and dominant pole) and a second pole, p_{2eq}, which guarantees the same phase margin given by (4.33). Hence, (4.25) is approximated by

$$T(s) \approx T_o \frac{1}{\left(1 + \frac{s}{p_1}\right)\left(1 + \frac{s}{p_{2eq}}\right)} \tag{4.34}$$

where

$$p_{2eq} = K\left[1 + \frac{\left(K + \frac{1}{K}\right)\omega_{dn}\Delta}{\left(1+\omega_{dn}^2\right)\sqrt{1-\Delta} - K\omega_{dn}\Delta}\right]\omega_{GBW} \tag{4.35}$$

The second term within square brackets in (4.35) gives the deviation of the equivalent second pole with respect to the actual second pole, which depends on both ω_{dn} and Δ. It can be easily shown that the deviation is at a maximum when $\omega_{dn} = 1$, for a fixed value of Δ (and K), that is when the average doublet frequency is equal to the gain-bandwidth product. In contrast, with ω_{dn} and K given, the influence of the doublet is highest in

correspondence to the value of Δ which nullifies the denominator of (4.35), that is

$$\Delta = \frac{1}{2}\left(\frac{1+\omega_{dn}^2}{K\omega_{dn}}\right)^2\left(\sqrt{1+4\left(\frac{K\omega_{dn}}{1+\omega_{dn}^2}\right)^2}-1\right) \tag{4.36}$$

For instance, assuming $K = 2$ and $\omega_{dn} = 2$, Δ is 0.693.

In [PP992] the model was validated for values of Δ in a range from –1 to 0.5, meaning a doublet with its pole and zero spaced by a factor of two.

The time-domain closed-loop step response can be also approximately represented through that of a pure two-pole amplifier. To evaluate the effect of a pole-zero doublet, we calculate the relative deviation of t_p and D in a two-pole amplifier given by (4.23) and (4.24)

$$\frac{\Delta t_p}{t_p} = -\frac{(K-2)(K^2+1)}{K(K-4)}\frac{\omega_{dn}\Delta}{(1+\omega_{dn}^2)\sqrt{1-\Delta}-K\omega_{dn}\Delta} \tag{4.37}$$

$$\frac{\Delta D}{D} = -2\pi\frac{K(K^2+1)\sqrt{K(4-K)}}{(4-K)^2}\frac{\omega_{dn}\Delta}{(1+\omega_{dn}^2)\sqrt{1-\Delta}-K\omega_{dn}\Delta} \tag{4.38}$$

For those cases in which Δ is small, such as when a doublet arises from process tolerances in a pole-zero compensation [BAR80], [PP95], relationships (4.37) and (4.38) can be approximated to

$$\frac{\Delta t_p}{t_p} \approx -\frac{(K-2)(K^2+1)}{K(K-4)}\left(\frac{\omega_{dn}}{1+\omega_{dn}^2}\right)\Delta \tag{4.39}$$

$$\frac{\Delta D}{D} \approx -2\pi\frac{K(K^2+1)\sqrt{K(4-K)}}{(4-K)^2}\left(\frac{\omega_{dn}}{1+\omega_{dn}^2}\right)\Delta \tag{4.40}$$

The approximate relationships show that the relative variations of t_p and D are linearly related to the spacing of the doublet. It can be seen that for phase margins greater than 50° (i.e. $K > 1.2$) the variation in D is much greater than that of t_p. Besides, when the zero is lower than the pole, the doublet has the effect of reducing overshoot (both in the frequency and in the time domain).

As discussed in section 4.3, the second pole is often already defined and the compensation task requires setting the dominant pole or, better, the gain-bandwidth product, ω_{GBW}. Relationship (4.29) can be written as

$$\tan(\phi) = \frac{1}{\omega_{GBW}} \frac{\omega_{GBW}^2 (p_2 + p_d \Delta) + p_2 \omega_{dn}^2}{\omega_{GBW}^2 + \omega_d^2 - p_2 p_d \Delta} \tag{4.41}$$

hence, the required ω_{GBW} implies having to solve the following third-order equation

$$\tan(\phi)\omega_{GBW}^3 - (p_2 + p_d \Delta)\omega_{GBW}^2 + \tan(\phi)(\omega_{dn}^2 - p_2 p_d \Delta)\omega_{GBW} - p_2 \omega_{dn}^2 = 0 \tag{4.42}$$

As particular cases, first consider the one where the pole-zero doublet is derived from differential-to-single conversion. In this event doublet spacing, Δ, is exactly equal to –1. By developing (4.42) in Taylor series around the point $\omega_{GBW} = p_2/\tan(\phi)$ truncated to the second term, we get

$$(2p_2 + p_d)\omega_{GBW}^2 + \left[\tan(\phi)(p_2 + p_d)p_d - \frac{3p_2^2}{\tan(\phi)}\right]\omega_{GBW} + \left(\frac{p_2^2}{\tan^2(\phi)} - 2p_d^2\right)p_2 = 0 \tag{4.43}$$

which is sufficiently simple to be solved with pencil-and-paper.

In contrast, when the second pole can be moved to guarantee stability, such as in the design strategy for cascode amplifiers proposed in [MN89], from (4.41) noting that

$$p_d = \frac{\omega_{dn}}{\sqrt{1-\Delta}} \tag{4.44}$$

we have to set p_2 according to

$$p_2 = \omega_{GBW} \tan(\phi) \left(\frac{1 + \omega_{dn}^2 - \frac{\Delta}{\sqrt{1-\Delta}} \omega_{dn} \tan(\phi)}{1 + \omega_{dn}^2 + \frac{\Delta}{\sqrt{1-\Delta}} \omega_{dn} \tan(\phi)} \right) \tag{4.45}$$

4.4 THREE-POLE FEEDBACK AMPLIFIERS WITH REAL POLES

Some amplifier architectures have three separate poles [P99], one of which must be dominant to allow stability. Consider then the third-order loop-gain transfer function given below

$$T(s) = \frac{T_o}{\left(1 + \frac{s}{p_1}\right)\left(1 + \frac{s}{p_2}\right)\left(1 + \frac{s}{p_3}\right)} \tag{4.46}$$

The phase margin of the loop gain is equal to (approximating ω_t with ω_{GBW})

$$\phi = \tan^{-1}\frac{p_2}{\omega_{GBW}} + \tan^{-1}\frac{p_3}{\omega_{GBW}} - 90° \tag{4.47}$$

Since

$$\tan(a+b) = \frac{\tan(a) + \tan(b)}{1 - \tan(a)\tan(b)} \tag{4.48}$$

and

$$\tan(a + 90°) = -\frac{1}{\tan(a)} \tag{4.49}$$

relationship (4.47) can be rewritten as

$$\frac{1}{\omega_{GBW}^2} - \tan(\phi)\left(\frac{1}{p_2} + \frac{1}{p_3}\right)\frac{1}{\omega_{GBW}} - \frac{1}{p_2 p_3} = 0 \tag{4.50}$$

Thus, assuming the non-dominant poles to be definitely set, the required gain-bandwidth for a given phase margin can be achieved from (4.50). In particular, we get

$$\frac{1}{\omega_{GBW}} = \frac{\tan(\phi)}{2}\left(\frac{1}{p_2}+\frac{1}{p_3}\right)\left[1+\sqrt{1+\frac{4}{\tan^2(\phi)}\frac{p_2 p_3}{(p_2+p_3)^2}}\right] \approx \tan(\phi)\left(\frac{1}{p_2}+\frac{1}{p_3}\right) \tag{4.51}$$

It is worth noting that the frequency compensation of a three-pole amplifier can be performed following the same procedure as for an equivalent two-pole amplifier with a loop gain given by

$$T(s) \approx T_o \frac{1}{\left(1+\frac{s}{p_1}\right)\left(1+\frac{s}{p_{2eq}}\right)} \tag{4.52}$$

The time constant of the equivalent second pole equals the sum of the second and third pole time constants of the three-pole amplifier. In other words, the equivalent pole is

$$p_{2eq} = \frac{p_2 p_3}{p_2 + p_3} \tag{4.53}$$

and the frequency and time-domain behaviour of the closed loop amplifier can be approximated with those developed in section 4.2.

4.5 THREE-POLE FEEDBACK AMPLIFIERS WITH A PAIR OF COMPLEX AND CONJUGATE POLES

Another common situation for a three-pole amplifier is when a dominant pole occurs in conjunction with a pair of complex conjugate poles. We use the symbolism introduced in section 4.3 to express such a loop-gain transfer function

$$T(s) = \frac{T_o}{\left(1+\frac{s}{p_1}\right)\left(1+\frac{s}{\omega_{GBWi}}+\frac{s^2}{\omega_{GBWi}^2 K_i}\right)} \tag{4.54}$$

This representation can be particularly useful when the complex poles derive from two indented feedback loops (such as in three-stage amplifiers with nested-Miller[4] compensation). In this case, term ω_{iBWi} is the gain-bandwidth product of the inner loop-gain and K_i is the ratio between the second pole and the gain-bandwidth product in this inner loop. The phase margin of the whole amplifier is given by

$$\phi = 90° - \tan^{-1}\frac{\frac{\omega_{GBW}}{\omega_{GBWi}}}{1-\frac{\omega_{GBW}^2}{\omega_{GBWi}^2 K_i}} = \tan^{-1}\left(\frac{1-\frac{\omega_{GBW}^2}{\omega_{GBWi}^2 K_i}}{\frac{\omega_{GBW}}{\omega_{GBWi}}}\right) \tag{4.55}$$

hence from (4.55) we get

$$\left(\frac{\omega_{GBWi}}{\omega_{GBW}}\right)^2 - \tan(\phi)\frac{\omega_{GBWi}}{\omega_{GBW}} - \frac{1}{K_i} = 0 \tag{4.56}$$

and we can determine the gain-bandwidth required for a fixed phase margin when the higher poles are fixed

$$\frac{1}{\omega_{GBW}} = \frac{\tan(\phi)}{2}\left[1+\sqrt{1+\frac{4}{K_i \tan^2(\phi)}}\right]\frac{1}{\omega_{GBWi}} \tag{4.57}$$

Like the case of three separate poles, now we can define an equivalent second pole and, if the quantity within the square roots is close to one, which means

$$K_i \tan^2(\phi) > 4 \tag{4.58}$$

the equivalent second pole is approximated by ω_{iBWi}. The frequency and time domain behaviour of the closed loop amplifier are hence equal to those of the closed loop amplifier whose open loop transfer function is

$$T(s) \approx T_o \frac{1}{\left(1+\frac{s}{p_1}\right)\left(1+\frac{s}{\omega_{GBWi}}\right)} \tag{4.59}$$

[4] See Sec. 5.5 of this book.

4.6 TWO-POLE FEEDBACK AMPLIFIERS WITH A ZERO

Often the loop gain of a feedback amplifier has a zero which can heavily affect the transfer function of the closed-loop amplifier. Indeed, if the loop gain is

$$T(s) = T_o \frac{1 + \frac{s}{z}}{\left(1 + \frac{s}{p_1}\right)\left(1 + \frac{s}{p_2}\right)} \tag{4.60}$$

the closed-loop transfer function exhibits the same zero and is given by

$$G_F(s) = G_{Fo} \frac{1 + \frac{s}{z}}{1 + 2\frac{\xi}{\omega_o} s + \frac{s^2}{\omega_o^2}} \tag{4.61}$$

where in the denominator ω_o is still given by (4.7) and the damping factor is modified with respect to (4.8) according to

$$\xi = \frac{1}{2\sqrt{1 + T_o}} \left(\sqrt{\frac{p_1}{p_2}} + \sqrt{\frac{p_2}{p_1}} + T \frac{\sqrt{p_1 p_2}}{z} \right) \tag{4.62}$$

The phase margin of the feedback amplifier, under the assumption of a dominant-pole behaviour whose pole and zero is higher than the transition frequency, ω_T, is given by

$$\phi = 90° + \text{arctg}\frac{p_2}{\omega_{GBW}} - \text{arctg}\frac{z}{\omega_{GBW}} \tag{4.63}$$

which shows that a negative zero helps stability, but a positive zero can drastically reduce the phase margin. Therefore, during compensation particular care must be taken to avoid or minimise the effect of positive zeros.

Although it is seldom used, the step response of a feedback amplifier with the closed-loop transfer function in (4.61) is

$$x_o(t) = \left\{1 - \frac{1}{p_{F1} - p_{F2}}\left[\left(1 - \frac{p_{F1}}{z}\right)p_{F2}\mathrm{e}^{-p_{F1}t} - \left(1 - \frac{p_{F2}}{z}\right)p_{F1}\mathrm{e}^{-p_{F2}t}\right]\right\}G_{olF}u(t) \tag{4.64}$$

where p_{F1} and p_{F2} are the poles given in (4.9). For an underdamped amplifier (4.64) can be expressed more profitably as

$$x_o(t) = \left\{1 - \left[H\sin\left(\sqrt{1-\xi^2}\,\omega_o t\right) + \cos\left(\sqrt{1-\xi^2}\,\omega_o t\right)\right]\mathrm{e}^{-\xi\omega_o t}\right\}G_{olF}u(t) \tag{4.65}$$

where $H = \left[\frac{\sqrt{1-\xi^2}\,\omega_o}{z} + \frac{\xi}{\sqrt{1-\xi^2}}\left(\frac{\xi\omega_o}{z} - 1\right)\right]$.

This page intentionally left blank

Chapter 5

FREQUENCY COMPENSATION TECHNIQUES

In the previous chapter we demonstrated the necessity, in a feedback network, to achieve an open loop dominant-pole frequency response whose a phase margin is greater than 45° (or $K > 1$). Indeed, this condition not only ensures closed-loop stability but also avoids unacceptably underdamped closed-loop responses. Unfortunately, many amplifiers, and particularly broadbanded amplifiers, earmarked for use as open-loop cells are not characterised by dominant-pole frequency responses. The loop-gain frequency response of these amplifiers must be therefore properly optimised in accordance with standard design practices known as *frequency compensation techniques* [SS91], [GM93], [LS94]. These methods imply the inclusion of compensation *RC* networks in the uncompensated circuit to introduce additional poles or to modify the original loop-gain poles so as to provide a given phase margin.

Referring to Fig. 4.8, it is easily understood that the simplest way to achieve stability is to reduce the loop gain. If the frequency of the poles remain unchanged, the unity-gain frequency is decreased by the same amount as the loop gain reduction and consequently the ratio between the second pole and the gain-bandwidth product is increased. The loop gain can be reduced via the feedback factor f or by decreasing the amplifier open-loop gain. However, neither are practical design choices because changing the loop gain may conflict with closed-loop performance such as gain, accuracy, etc. Moreover, it is worthwhile noting that compensation must be ensured for all the possible feedback configurations. If the feedback factor is not specified, compensation should be performed in the worst-case condition, that corresponds to the unitary feedback (i.e., with the highest loop gain and gain-bandwidth product, $f = 1$ and $T_o = A_o$).

In the following three paragraphs we will study the engineering methods and related tradeoffs underlying the key issue of the frequency compensation for a two-pole open loop transfer function. Of course, the discussion is easily extended to multi-pole functions with two dominant poles. The last two paragraphs deal with the frequency compensation of three-stage amplifiers.

5.1 DOMINANT-POLE COMPENSATION

Let us consider the two-pole amplifier in Fig. 5.1 whose open-loop transfer function is

$$A(s) = \frac{A_o}{\left(1 + \frac{s}{p_1}\right)\left(1 + \frac{s}{p_2}\right)} \tag{5.1}$$

where

$$A_o = G_{m1} G_{m2} R_1 R_2 \tag{5.2}$$

$$p_1 = \frac{1}{R_1 C_1} \tag{5.3}$$

$$p_2 = \frac{1}{R_2 C_2} \tag{5.4}$$

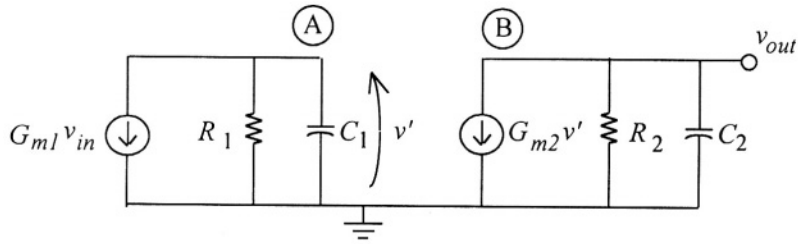

Fig.5.1 Small-signal model of a two-pole amplifier.

The two poles are determined by the parasitic capacitances associated with node A and B. Assuming these poles are widely separated with $p_1 << p_2$, the tangent of the phase margin becomes

$$\tan(\phi) = \frac{p_2}{|A_o|p_1} = \frac{1}{|A_o|}\left(\frac{R_2C_2}{R_1C_1}\right) \tag{5.5a}$$

Conversely, if $p_2 << p_1$ we would have

$$\tan(\phi) = \frac{p_1}{|A_o|p_2} = \frac{1}{|A_o|}\left(\frac{R_1C_1}{R_2C_2}\right) \tag{5.5b}$$

To guarantee a phase margin greater than 45°, $\tan(\phi)$ must be greater than unity. Hence, from (5.5a) and (5.5b), we must ensure that the ratio between the two time constants is in the order of the DC gain. For example, assuming the two equivalent resistances to be equal and a typical gain of 30 one of the capacitances should be more than 30 times the other, to guarantee stability within proper margins.

At this point, the most intuitive way to provide stability is to *add* a capacitance in parallel to C_1 (or C_2), thus setting the dominant pole at the input or the output. If we adopt this strategy, the choice of where to insert the compensation capacitor depends on convenience in terms of lower added capacitance. This simple compensation approach is called dominant-pole compensation, which is rarely used, except in single-stage (cascode) amplifiers, because it requires large compensation capacitors and leads to feedback amplifiers with very low bandwidth. To show the reduction in bandwidth, without loss of generality consider the amplifier as being in unitary feedback and set the dominant pole at the input by adding the compensation capacitor $C_{C'}$ to C_1. Thus (5.5b) turns out to be

$$\tan(\phi) \approx \frac{1}{|A_o|}\frac{R_1(C_1 + C_{C'})}{R_2C_2} \tag{5.6}$$

and the dominant pole after compensation, $p_{1C'}$, which defines the open-loop bandwidth

$$p_{1C'} \approx \frac{1}{R_1(C_1 + C_{C'})} \tag{5.7}$$

must be lower than the second pole (which remains unchanged to p_2) reduced by the DC gain times the tangent of the phase margin (always higher than 1)

$$p_{1C} < \frac{p_2}{\tan(\phi)|A_o|} \tag{5.8}$$

To conclude, the bandwidth of the dominant-pole compensated amplifier is defined by the *second pole and the DC open-loop gain.* As we shall see in the next paragraph this condition does not hold for the Miller compensation strategy.

5.2 MILLER (POLE-SPLITTING) COMPENSATION

The well-known *Miller* effect can be efficiently exploited to perform frequency compensation that for this reason is called Miller compensation or pole splitting compensation. To understand its properties and design issues consider the small-signal model in Fig. 5.2, which but for the presence of the interstage capacitance C_r coupling the first and second stage, is equal to the one in Fig. 5.1.

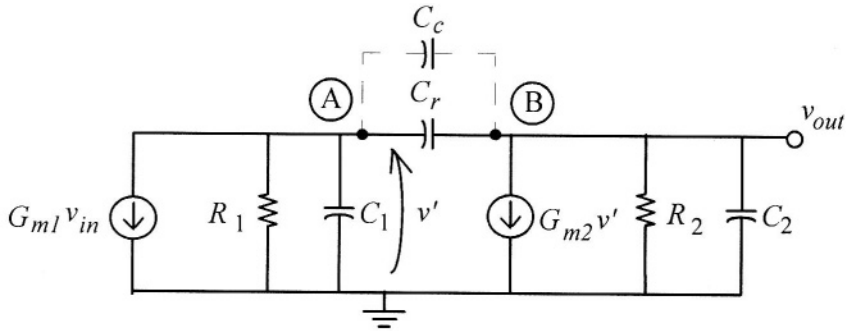

Fig. 5.2. Small-signal model of a two-pole amplifier with interstage capacitance.

Neglecting for the moment capacitor C_c depicted in dashed lines, the subject open-loop transfer function is

$$A(s) = A_o \frac{1 + b_1 s}{1 + a_1 s + a_2 s^2} = A_o \frac{1 + \frac{s}{z}}{1 + \left(\frac{1}{p_1} + \frac{1}{p_2}\right)s + \frac{s^2}{p_1 p_2}} \tag{5.9}$$

where the DC gain A_o is still given by (5.2) whose the coefficients are

$$a_1 = R_1C_1 + R_2C_2 + (R_1 + R_2 + R_1G_{m2}R_2)C_r \tag{5.10}$$

$$a_2 = R_1R_2C_2\left[C_1 + \left(1 + \frac{C_1}{C_2}\right)C_r\right] \tag{5.11}$$

$$b_1 = -\frac{C_r}{G_{m2}} \tag{5.12}$$

Thus, assuming the poles are widely separated their approximate expressions become:

$$p_1 \approx \frac{1}{R_2(C_2 + C_r) + R_1[C_1 + (1 + G_{m2}R_2)C_r]} \tag{5.13}$$

$$p_2 \approx \frac{R_2(C_2 + C_r) + R_1[C_1 + (1 + G_{m2}R_2)C_r]}{R_1R_2C_2\left[C_1 + \left(1 + \frac{C_1}{C_2}\right)C_r\right]} \tag{5.14}$$

Capacitor C_r provides a path for feedback and for feedforward. The feedforward leakage produces a real zero in the right-half plane (RHP) given by

$$z = -\frac{G_{m2}}{C_r} \tag{5.15}$$

The effect of this zero is neglected here for simplicity (the zero may be either at a very-high frequency or be compensated with one of the methods described in the next paragraph).

In (5.13) and (5.14), term $(1 + G_{m2}R_2)C_r$ accounts for the Miller effect [MG87]. In practical cases it is the dominant term because capacitance C_r is multiplied by a factor as high as a stage gain, $G_{m2}R_2$. In such cases the expressions of the two poles (5.13) and (5.14) simplify to

$$p_1 \approx \frac{1}{R_1[C_1 + (1 + G_{m2}R_2)C_r]} \tag{5.16}$$

$$p_2 \approx \frac{C_1 + (1 + G_{m2} R_2) C_r}{R_2 C_2 \left[C_1 + \left(1 + \frac{C_1}{C_2} \right) C_r \right]} \tag{5.17}$$

From equations (5.15) to (5.17) the pole splitting due to Miller effect becomes apparent. In fact, an increase in the internal feedback capacitance, C_r, shifts the dominant pole and the second pole to a lower and higher frequency, respectively (and also decreases the RHP zero). For this purpose, to improve the separation of the two poles it is very efficient to multiply C_r. Thus, pole-splitting compensation entails connecting a capacitor between two phase inverting nodes of the open-loop amplifier. With reference to the equivalent circuit in Fig. 5.2, the electrical impact of this additional element is the replacement of the internal interstage capacitance, C_r, by the capacitance sum, $C_C + C_r$.

Letting $C_p = C_C + C_r$, (5.16) and (5.17) can be further simplified to

$$p_{1C} \approx \frac{1}{R_1 G_{m2} R_2 C_p} \tag{5.18}$$

$$p_{2C} \approx \frac{G_{m2}}{(C_2 + C_1)} \tag{5.19}$$

where capacitance C_p is usually significantly larger than C_r and has also been assumed to be greater than either C_1 or C_2. Note that the value of the compensated second pole given by (5.19) encounters an intuitive justification. In fact, at the frequency at which it occurs (i.e. after the transition frequency or equivalently the gain-bandwidth product), capacitance C_p can be considered as short-circuited. Hence, the input and the output of the voltage-controlled current-source are shorted, C_2 and C_1 are in parallel and the equivalent resistance seen at their terminals is approximately $1/G_{m2}$.

In contrast, the expression of the zero (5.15) becomes

$$z_C = -\frac{G_{m2}}{C_r + C_C} \tag{5.20}$$

Although z_C can exert a significant influence on the high-frequency response of the compensated amplifier, the following discussion presumes tacitly that $z_C > p_{2C}$. Hence, the gain-bandwidth product and the phase margin are

$$\omega_{GBW} = \frac{G_{m1}}{C_r + C_C} \tag{5.21}$$

$$\tan(\phi) = \frac{G_{m2}}{G_{m1}} \frac{C_r + C_C}{C_2 + C_1} \tag{5.22}$$

and the required compensation capacitance must be set according to

$$C_C = \frac{G_{m1}}{G_{m2}} \tan(\phi)(C_2 + C_1) - C_r \tag{5.23}$$

For a fixed phase margin C_C is proportional to the ratio between the transconductance of the first and second stage. Moreover, it is proportional to the total (input-output) capacitance. Note also that for a given DC gain and phase margin, the gain-bandwidth product is set by the frequency of second pole. Consequently, to compare the Miller and dominant-pole compensations we can compare only the second poles, and it is apparent that the second pole resulting from the Miller compensation is much higher (due to pole-splitting) than that of a dominant-pole compensated amplifier. In addition, Miller magnification allows us to use lower capacitance values.

For these reasons the Miller compensation technique is extensively used to design IC amplifiers. Compensation capacitor C_C can be fabricated as a part of the amplifier (in this case the amplifier is said to be *internally compensated*) or can be externally applied to pins reserved for this purpose on the (*uncompensated*) opamp package.

By comparing (5.20) and (5.21), we find that we can neglect the right-half plane zero when the transconductance gain of the second stage is much higher than that of the first stage. This condition is seldom satisfied in CMOS transconductance amplifiers and especially when low-power dissipation is required, so that a specific strategy to compensate the zero must be applied.

5.3 COMPENSATION OF THE MILLER RHP ZERO

In the previous paragraph we showed that the pole-splitting technique is a convenient vehicle for achieving the desired pole separation in an open-loop phase-inverting amplifier. Unfortunately, (5.20) indicates that the larger the C_C ,the lower the RHP zero. In bipolar technologies the transconductance G_{m2} is invariably large enough to ensure that the frequency of the zero is greater than the compensated unity-gain frequency, thereby rendering the

impact of $z_{c\cdot}$ on the compensated frequency response inconsequential. But for MOS and CMOS technologies, the transconductance is small, and as a result, the effects of the RHP zero evidenced in the forward transfer function of a phase inverting amplifier may not be negligible. When the transmission zero is significant, its primary effect is to incur excess phase lag (phase lag in addition to that produced by the two open-loop poles), while prohibiting a uniform 20 dB-per-decade frequency response roll-off rate at high frequencies. The stability problems caused by the resultant deterioration in phase margin justifies the implementation of compensation techniques that neutralise the effects of the RHP zero.

Various compensation schemes have been proposed for two-stage MOS opamps. They are based on the concept of breaking the forward path through the compensation capacitor by using active or passive components. The first of these was applied in a NMOS opamp [TG76] and then in a CMOS opamp [SHG78]. It breaks the forward path by introducing a voltage buffer in the compensation branch. Next, a compensation technique was proposed which uses a nulling resistor in series with the compensation capacitor [A83]. Another solution works like the former but uses a current buffer to break the forward path [A83]. Finally, both current and voltage buffers can be adopted to compensate the right half-plane zero [MT90].

5.3.1 Nulling Resistor

The most widely used compensation technique is the one based on the nulling resistor. It entails the incorporation of a resistor, R_C, in series with the Miller compensation capacitor as shown in Fig. 5.3.

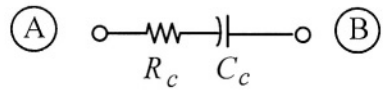

Fig. 5.3. Compensation network using a nulling resistor.

The popularity of this scheme stems from the fact that it can be implemented monolithically with a MOS transistor biased in its triode regime (which approximates a linear resistor). Moreover, its highpass nature does not reduce the low-frequency dynamic range of the imcompensated configuration. By using this compensation branch in the equivalent circuit in Fig. 5.2, and neglecting capacitance C_r (usually much lower than $C_{c\cdot}$), the zero is now at frequency

$$z_r = -\frac{1}{\left(\frac{1}{G_{m2}} - R_C\right)C_C} \tag{5.24}$$

and is moved to infinite frequency by setting R_C equal to $1/G_{m2}$. Thus, the RHP zero originally ignored in the process of arriving at the pole-splitting results has effectively been eliminated.

If R_C is greater than $1/G_{m2}$, a left-hand zero is created because z_r becomes positive. Ideally, this zero can be exploited to offset or even cancel the effects of the second compensated pole, thereby leading to an open-loop amplifier with an increased gain-bandwidth product as first proposed in [BAR80].

By imposing the condition

$$\frac{G_{m2}}{C_2} = \frac{G_{m2}}{(G_{m2}R_C - 1)C_C} \tag{5.25}$$

a new second pole arises which is given by $1/(R_C C_C)$, as can be found by directly analysing the equivalent circuit. This pole does not depend on the load capacitance. However, this optimised approach has a quite worse ω_{GBW} than the other optimised compensation strategies for equal power consumption and area of the amplifier including the compensation network (i.e., global transconductance in the amplifier) as demonstrated in [PP95] and [PP97].

5.3.2 Voltage Buffer

Figure 5.4 shows the compensation branch with a voltage buffer. It eliminates feedforward through the compensation capacitance C_C.

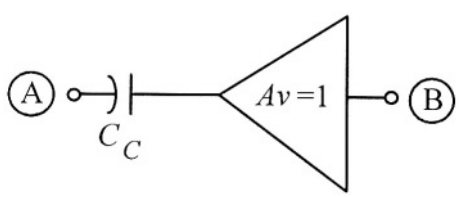

Fig. 5.4. Compensation network with an ideal voltage buffer.

Unfortunately, unlike the passive compensation strategy discussed above, the buffer utilised attenuates the achievable output swing of the amplifier. The adoption of an ideal voltage buffer (i.e., with infinitely large input impedance, zero output impedance, and unitary gain) gives the same dominant pole as in (5.18) and the same second pole as in (5.19) without depending on capacitance C_1. But by eliminating capacitive feedforward, the troublesome RHP zero incurred by the internal interstage capacitance, C_r, is not decreased by the compensation element C_C. In other words, the effective feedback capacitance is C_p, while the feedforward capacitance is C_r.

The foregoing discussion presumes an ideal voltage buffer. Practical buffers have small, but not zero, output impedance and large, but not infinite, input impedance (see Fig. 5.5).

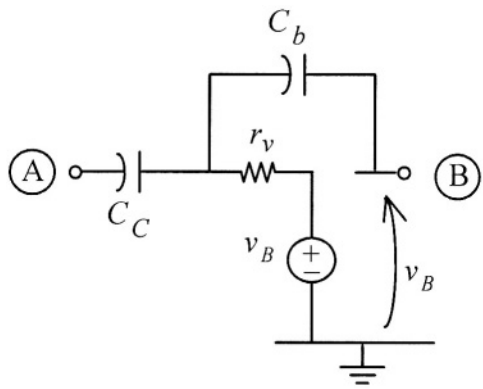

Fig. 5.5. Compensation network with a real voltage buffer.

The resistive component of the buffer output impedance, r_v, establishes a left-half plane zero with capacitance C_C. As for the case with the nulling resistor, this zero can be exploited to increase the amplifier gain-bandwidth [PP95]. Following this last compensation strategy and with some approximations, the poles and zeros become [PP95]

$$p_{1C} \approx \frac{1}{R_1 G_{m2} R_2 C_C} \tag{5.26}$$

$$p_{2C} \approx \frac{G_{m2}}{C_2 + C_b G_{m2} r_v} \tag{5.27}$$

$$p_{3C} \approx \frac{1}{(C_1 + C_b) r_v} \tag{5.28}$$

$$z_{1C'} = \frac{1}{C_{C'} r_v} \tag{5.29}$$

$$z_{2C'} = -\frac{G_{m2}}{C_b} \tag{5.30}$$

Now the right-half plane zero, $z_{2C'}$, is placed at a very high frequency and can be neglected. Moreover, as proposed in [AH87] and developed in [PP95], a pole-zero compensation can be performed to increase the gain bandwidth product. In particular, we can properly design the voltage buffer to ensure the output resistance is equal to

$$r_v = \frac{C_2}{G_{m2}(C_{C'} - C_b)} \approx \frac{C_2}{G_{m2} C_{C'}} \tag{5.31}$$

which sets $z_{1C'} = p_{2C'}$. The new second pole is now the old third pole in (5.28) which by using (5.31) becomes

$$p_{2C'N} = p_{3C'} = \frac{G_{m2}}{C_2} \frac{C_{C'}}{C_1 + C_b} \tag{5.32}$$

The phase margin is given by

$$\tan(\phi) = \frac{G_{m2}}{G_{m1}} \frac{C_{C'}^2}{C_2 (C_1 + C_b)} \tag{5.33}$$

which yields the required compensation capacitance

$$C_{C'} = \sqrt{\tan(\phi) \frac{G_{m1}}{G_{m2}} (C_1 + C_b) C_2} \tag{5.34}$$

After substituting the value found in the gain bandwidth product we get

$$\omega_{GBW} = \frac{G_{m1}}{\sqrt{\tan(\phi) \dfrac{G_{m1}}{G_{m2}} (C_1 + C_b) C_2}} \tag{5.35}$$

The resulting ω_{GBW} has a higher value than that given by (5.21), and is inversely dependent on the geometric media of $C_1 + C_b$ and C_2.

5.3.3 Current Buffer

Consider now the ideal compensation branch using the current buffer depicted in Fig. 5.6. This solution is very efficient both for the gain-bandwidth [C93], [RK95] and PSRR performance [A83], [RC84], [SS90], [SGG91]. Moreover, it does not have the drawback exhibited by the voltage buffer of reducing the amplifier output swing.

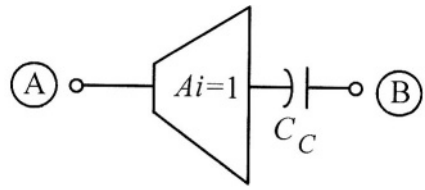

Fig. 5.6. Compensation network with an ideal current buffer.

With an ideal current buffer, the second pole is given by

$$p_{2c} = \frac{G_{m2}}{C_1\left(1+\frac{C_2}{C_C}\right)} \tag{5.36}$$

and the phase margin is by

$$\tan(\phi) = \frac{G_{m2}}{G_{m1}} \frac{C_C^2}{C_1(C_C + C_2)} \tag{5.37}$$

By solving for the compensation capacitance we found

$$C_C = \tan(\phi)\frac{G_{m1}}{2G_{m2}}\left(1+\sqrt{1+\frac{4G_{m2}}{\tan(\phi)G_{m1}}\frac{C_2}{C_1}}\right)C_1 \approx$$

$$\approx \tan(\phi)\frac{G_{m1}}{2G_{m2}}C_1 + \sqrt{\tan(\phi)\frac{G_{m1}}{G_{m2}}C_1C_2} \tag{5.38}$$

Generally the output capacitance, C_2, is much higher than the inner capacitance, C_1, and relationship (5.38) can be further simplified to

$$C_C \approx \sqrt{\tan(\phi)\frac{G_{m1}}{G_{m2}}C_1C_2} \tag{5.39}$$

Hence, for a given phase margin, the required compensation capacitance is slightly lower than the value required by the optimised compensation with voltage buffer in (5.34), while the resulting gain bandwidth product is slightly higher.

However, compensation with a real current buffer (and specifically, with finite input resistance) is not as straightforward as the other compensation approaches. As shown in reference [PP97], to achieve compensation we have to guarantee that the input resistance of the current buffer, r_c, is equal to or lower than half $1/G_{m1}$. Moreover, the condition

$$r_c = \frac{1}{2G_{m1}} \tag{5.40}$$

represents an optimum to maximise the gain-bandwidth product. Under condition (5.40) the required compensation capacitor is

$$C_C \approx h\frac{G_{m1}}{2G_{m2}}C_1 + \sqrt{\left(h+\frac{1}{2}\right)\frac{G_{m1}}{G_{m2}}C_1C_2} \tag{5.41}$$

where

$$h = \frac{2\tan(\phi)-1}{2+\tan(\phi)} \tag{5.42}$$

Usually, relationship (5.41) can be further approximated

$$C_C \approx \sqrt{\left(h+\frac{1}{2}\right)\frac{G_{m1}}{G_{m2}}C_1C_2} \tag{5.43}$$

and the gain bandwidth product results as

$$\omega_{GBW} \approx \frac{G_{m1}}{\sqrt{\left(h+\frac{1}{2}\right)\frac{G_{m1}}{G_{m2}}C_1C_2}} \tag{5.44}$$

For practical phase margin values, the gain bandwidth product in (5.44) is even higher than that obtained with a ideal current buffer. It is also higher than the one obtained using a real voltage buffer. However, compensation with a real current buffer is a less efficient strategy because, as demonstrated in [PP97], it needs more area and/or power for equal gain bandwidth product than compensation based on a real voltage buffer.

5.4 NESTED MILLER COMPENSATION

The compensation of multistage amplifiers (i.e., with a number of gain stages higher than two) requires iteration of the simple Miller compensation described previously [C78], [C821], [C96], [HL85], [EH95]. Typically, three- and even four-stage amplifiers are found in CMOS implementations including an output power stage for driving heavy off-chip loads [C822], [OA90], [PD90], [TGC90], [CN91], [PNC93]. Moreover, given the decrease in supply voltages, *cascoding* is not a suitable technique for IC applications demanding both high gain and swing. Hence, *cascading* three or more simple stages is the only viable option. Consequently, multistage amplifiers and their frequency compensation issues have become increasingly important in modern microelectronics [FH91], [EH92], [NG93], [PPS99], [GPP00]. In the following we will discuss in detail compensation of three-stage amplifiers, which are the most common architectures, but the results obtained can also be adapted (although often not very directly) to architectures with a higher number of stages.

5.4.1 General Features

Among the possible ways of exploiting Miller compensation in multistage amplifiers, the so-called *nested Miller* (NM) compensation is one of the most widely used. It can be utilised when only the final gain stage is voltage-inverting.

The small-signal equivalent circuit of a three-stage amplifier including nested Miller compensation is depicted in Fig. 5.7. Parameters G_{mi} and R_i are the i-th stage transconductance and output resistance, respectively. Capacitors C_i represent the equivalent capacitance at the output of each stage, C_{Ci} are the compensation capacitors, and C_L is the equivalent load capacitor.

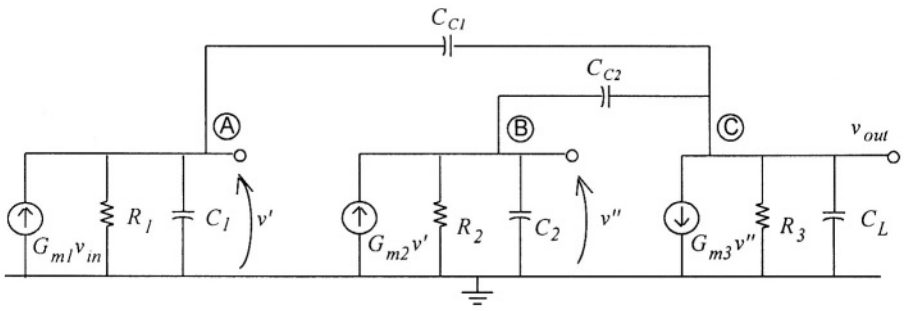

Fig. 5.7. Small-signal model of a three-stage NM-compensated amplifier.

In the following we neglect the effects of the parasitic capacitances since they are generally one order of magnitude lower than the compensation capacitances. Neglecting second-order terms, the open-loop transfer function of the circuit in Fig. 5.7 is expressed by

$$A(s) = A_o \frac{1 - \dfrac{C_{C2}}{G_{m3}} s - \dfrac{C_{C1}C_{C2}}{G_{m2}G_{m3}} s^2}{\left(1 + \dfrac{s}{p_1}\right)\left[1 + \left(\dfrac{1}{G_{m2}} - \dfrac{1}{G_{m3}}\right)C_{C2}s + \dfrac{C_{C2}C_L}{G_{m2}G_{m3}} s^2\right]} \tag{5.45}$$

where A_o is the DC open-loop gain equal to

$$A_o = G_{m1}G_{m2}G_{m3}R_1R_2R_3 \tag{5.46}$$

and p_1 is the frequency of the dominant pole

$$p_1 \approx \frac{1}{R_1G_{m2}R_2G_{m3}R_3C_{C1}} \tag{5.47}$$

Hence, the gain-bandwidth product, ω_{GBW}, of the amplifier is equal to

$$\omega_{GBW} = \frac{G_{m1}}{C_{C1}} \tag{5.48}$$

Equation (5.45), in addition to a dominant pole, also includes two other higher poles and two zeros. Moreover, since the coefficients of the s and s^2 terms in the numerator are both negative, a RHP zero is created, which is located at a lower frequency than the other LHP zero. In analogy to the discussion of the previous paragraph, using voltage followers or current followers can nominally eliminate both zeros. Another solution is the *multipath* Miller approach proposed in [YES97] that, according to Fig. 5.8, provides a zero cancellation due to the effect described in [EH95]. In brief, the forward path contribution is ideally nullified by setting G_{mc} equal to G_{m1}

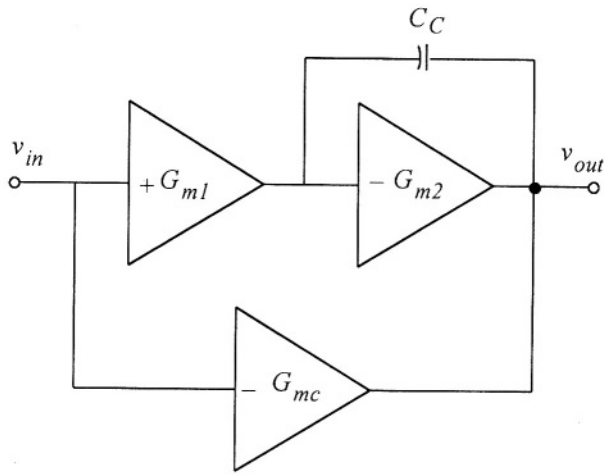

Fig. 5.8. Basic module for *multipath* nested-Miller zero cancellation.

When using any of these techniques, or in the case of a very large G_{m3}, such as in power amplifiers (whose output stage is biased with large quiescent currents and is realised with large devices), relationship (5.45) simplifies to

$$A(s) = A_o \frac{1}{\left(1+\frac{s}{p_1}\right)\left(1+\frac{C_{C2}}{G_{m2}}s+\frac{C_{C2}C_L}{G_{m2}G_{m3}}s^2\right)} \tag{5.49}$$

Equation (5.49) allows an interesting interpretation of the compensation process. We will show that assuming a dominant-pole frequency response, the second and third stage can be considered as closed in a unity-gain feedback configuration by capacitor C_{C1}, acting as a short circuit for frequencies above ω_{GBW}.

Consider now the open-loop gain of the second and third stage alone (which we also refer to as the *inner* amplifier), $A_i(s)$, its DC gain, the dominant pole p_{1i} due to the Miller effect on C_{C2}, and the second pole p_{2i} at the output terminal. They are given by

$$A_{oi} = G_{m2} G_{m3} R_2 R_3 \tag{5.50}$$

$$p_{1i} \approx \frac{1}{R_2 G_{m3} R_3 C_{C1}} \tag{5.51}$$

$$p_{2i} \approx \frac{G_{m3}}{C_L} \tag{5.52}$$

If now we assume $A_i(s)$ in unity-gain feedback connection, the resulting closed-loop transfer function is characterised by exactly the same second-order polynomial as in the denominator of (5.49). This consideration justifies the representation utilised in equation (4.54) and allows, in turn, the straightforward compensation technique discussed below.

For a well designed (i.e., with appropriate stability margins) inner amplifier, the second pole p_{2i} must be located well beyond the unity-gain frequency ω_{Ti}, which, under the dominant-pole behaviour assumption, is approximately equal to ω_{GBWi} and given by

$$\omega_{GBWi} \approx \frac{G_{m2}}{C_{C2}} \tag{5.53}$$

In order to avoid overshoot in the module of the inner amplifier frequency response, a proper ratio, K_i, between p_{2i} and ω_{GBWi} has to be set as described in paragraph 4.5. A fairly optimum value of K_i is 2 (i.e. an inner phase margin of about 64°) which is the minimum value guaranteeing monotonic behaviour in the frequency response module. This leads to the expression of capacitor C_{C2},

$$C_{C2} = 2 \frac{G_{m2}}{G_{m3}} C_L \tag{5.54}$$

In other words, we have an external feedback loop through C_{C1} and an inner one through C_{C2}. The stability of the inner loop must first be established so that we can proceed to the external one. Any design attempt not providing a proper phase margin for the inner loop would inevitably require an extremely high value of C_{C1} or even not achieve stability at all.

Now we return to the frequency response of the whole open-loop amplifier, which can be rewritten as in (4.54) and is here reported for clarity

$$A_o(s) = A_o \frac{1}{1+\frac{s}{p_1}} \frac{1}{1+\frac{s}{\omega_{GBWi}}+\frac{s^2}{2\omega_{GBWi}^2}} \tag{5.55}$$

Evaluation of the phase margin yields (see (4.55))

$$\phi = 90° - \tan^{-1}\frac{\frac{\omega_{GBW}}{\omega_{GBWi}}}{1-\frac{\omega_{GBW}^2}{2\omega_{GBWi}^2}} = \tan^{-1}\left(\frac{1-\frac{\omega_{GBW}^2}{2\omega_{GBWi}^2}}{\frac{\omega_{GBW}}{\omega_{GBWi}}}\right) \tag{5.56}$$

Solving (5.56) for $\omega_{\!/}/\omega_{\!/i}$ and combining with (5.48) and (5.53) gives the expression of capacitance C_{C1} as a function of the required phase margin

$$C_{C1} = \left[\tan(\phi) + \sqrt{\tan^2(\phi)+2}\right]\frac{G_{m1}}{G_{m3}}C_L \tag{5.57}$$

Equations (5.54) and (5.57), are very similar to those in [EH95], where a third-order Butterworth frequency response in unity-gain configuration is assumed. However, (5.57) is more general because allows to set compensation capacitor C_{C1} for the desired phase margin.

5.4.2 RHP Cancellation with Nulling Resistors

Now we extend the considerations on the nulling resistor network reported in 5.3.1, to the three-stage nested-Miller compensated amplifier. Figure 5.9 illustrates the *RC* compensation network which includes two nulling resistors R_{C1} and R_{C2}, to be used in the amplifier of Fig. 5.2.

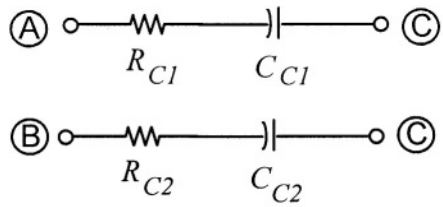

Fig. 5.9. Conventional compensation network using two nulling resistors.

With the introduction of these two resistors the open-loop gain given in (5.45) changes to

$$A(s) = A_o \frac{1 + \left[R_{C1}C_{C1} + \left(R_{C2} - \frac{1}{G_{m3}} \right) C_{C2} \right] s - \frac{1 + (1 - G_{m3}R_{C2})G_{m2}R_{C1}}{G_{m2}G_{m3}} C_{C1}C_{C2}s^2}{\left(1 + \frac{s}{p_1} \right) \left[1 + \left(R_{C2} + \frac{1}{G_{m2}} - \frac{1}{G_{m3}} \right) C_{C2}s + \frac{C_{C2}C_L}{G_{m2}G_{m3}} s^2 \right]} \tag{5.58}$$

Observe that only R_{C2} modifies the denominator because R_{C2} changes the zero of the inner amplifier. It is also clear that the numerator of (5.58) is greatly different from that of (5.45) and now depends on R_{C1} and R_{C2}. By inspection of (5.58), it is possible to nullify the s^2 term and to make the s term positive by choosing

$$R_{C2} = \frac{1}{G_{m2}G_{m3}R_{C1}} + \frac{1}{G_{m3}} \tag{5.59}$$

In this manner, the residual LHP zero can be exploited to increase the phase margin. However, we shall not further develop this approach because better ones have been elaborated.

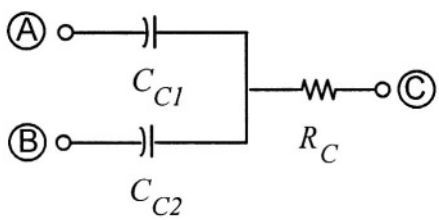

Fig. 5.10. Compensation network using a single nulling resistor.

A simpler technique based on a single nulling resistor and illustrated in Fig. 5.10 was proposed in [LM99]. When applied to the amplifier in Fig. 5.7, it gives the following loop-gain expression

$$A(s) = A_o \frac{1 + \left[R_C C_{C1} + \left(R_C - \frac{1}{G_{m3}} \right) C_{C2} \right] s + \frac{G_{m3} R_C - 1}{G_{m2} G_{m3}} C_{C1} C_{C2} s^2}{\left(1 + \frac{s}{p_1} \right) \left[1 + \left(\frac{1}{G_{m2}} - \frac{1}{G_{m3}} \right) C_{C2} s + \frac{1 - G_{m2} R_C}{G_{m2} G_{m3}} C_{C2} C_L s^2 \right]} \tag{5.60}$$

In this case the s^2 term in the numerator can be simply set equal to zero by choosing

$$R_C = \frac{1}{G_{m3}} \tag{5.61}$$

and the loop-gain only has a negative zero which can be used to increase the phase margin.

Now equation (5.54) cannot be used, but the same procedure can still be adopted to achieve simple new equations for C_{C1} and C_{C2}, for a given value of phase margin. After having substituted (5.61) in (5.60) and assuming $K_i = 2$, we can set C_{C2} and evaluate the phase margin

$$C_{C2} = 2 \frac{G_{m2}}{G_{m3} - G_{m2}} C_L \tag{5.62}$$

$$\phi = 90° - \tan^{-1} \frac{\frac{\omega_{GBW}}{\omega_{GBWi}}}{1 - \frac{\omega_{GBW}^2}{2\omega_{GBWi}^2}} + \tan^{-1} \frac{\omega_{GBW}}{z} =$$

$$= \tan^{-1} \left\{ \frac{2\omega_{GBWi} \left(\frac{\omega_{GBW}}{\omega_{GBWi}} \right)^2 + z \left[2 - \left(\frac{\omega_{GBW}}{\omega_{GBWi}} \right)^2 \right]}{\frac{\omega_{GBW}}{\omega_{GBWi}} \left\{ \omega_{GBWi} \left[\left(\frac{\omega_{GBW}}{\omega_{GBWi}} \right)^2 - 2 \right] + 2z \right\}} \right\} \tag{5.63}$$

where z is the zero $1/R_C C_{C1}$ and ω_{GBW1}, comparing (5.55) with (5.60) and using (5.62), is $½G_{m3}/C_L$. Solving (5.63) for $\omega_{GBW1}/\omega_{GBW}$ and combining with $\omega_{GBW} = G_{m1}/C_{C1}$ and (5.61) gives the value of capacitance C_{C1}

$$C_{C1} = \frac{\left[\sqrt{\chi^2\left(2\tan^2(\phi)+1\right)+2\chi\tan(\phi)+2+\tan^2(\phi)}+\tan(\phi)-\chi\right]}{1+\chi\tan(\phi)}\chi C_L \tag{5.64}$$

where

$$\chi = \frac{G_{m1}}{G_{m3}} \tag{5.65}$$

By considering that χ is lower than $\tan(\phi)$ for the phase margin of interest (i.e., for $\phi \geq 60°$), the above equation can be approximated as

$$C_{C1} \approx \frac{\tan(\phi)+\sqrt{2+\tan^2(\phi)+2\chi\tan(\phi)}}{1+\chi\tan(\phi)}\chi C_L \tag{5.66}$$

which is more suitable for pencil and paper design, and provides the same results as (5.57) for $\chi << 1$. Compared to (5.57), relationship (5.66) gives lower values of C_{C1} for the same phase margin.

It is interesting to note that we assumed no constraint for transconductances except $G_{m3} > G_{m2}$, otherwise C_{C2} in (5.62) would be negative. This allows the power consumption to be optimised since low quiescent currents can be used and, perhaps more importantly, we are free to choose the input and output transconductances G_{m1} and G_{m3}. Unfortunately, like for classic NM compensation, this method still requires large compensation capacitors for heavy capacitive loads. For instance, if $g_{m3} = 4g_{m1}$ and for a target phase margin of 70°, the required compensation capacitor C_{C1} equals $0.9C_L$.

An alternative and efficient compensation technique is based on the compensation network shown in Fig. 5.11 [PP02]. The previous single-resistor compensation network is here modified by adding another resistor, R_{C2}, in series with capacitor C_{C2}. Although this change may appear of marginal significance, it turns out to be very attractive because it allows pole-zero compensation to be achieved by using reduced compensation capacitor values. This in turn leads to an improvement in terms of gain-bandwidth product, slew-rate and settling time.

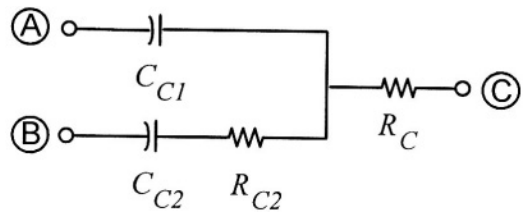

Fig. 5.11. Alternative two-resistor compensation network.

The transfer function of the amplifier in Fig. 5.7, using the compensation network in Fig. 5.11 becomes

$$A(s) = A_o \frac{1 + \left[R_C C_{C1} + \left(R_{C2} + R_C - \frac{1}{G_{m3}} \right) C_{C2} \right] s + \frac{(1 + G_{m2} R_{C2}) G_{m3} R_C - 1}{G_{m2} G_{m3}} C_{C1} C_{C2} s^2}{\left(1 + \frac{s}{p_1} \right) \left[1 + \left(R_{C2} + \frac{1}{G_{m2}} - \frac{1}{G_{m3}} \right) C_{C2} s + \frac{1 - G_{m2} R_C}{G_{m2} G_{m3}} C_{C2} C_L s^2 \right]} \quad (5.67)$$

The above shows that the zeros can both be made negative and their values adjusted to exactly cancel the two higher poles. Hence, by setting

$$R_C = \frac{1}{G_{m3}} \quad (5.68)$$

and equating the coefficients of the second-order polynomials we get

$$C_{C1} = \left(\frac{G_{m3}}{G_{m2}} - 1 \right) C_{C2} \quad (5.69)$$

$$R_{C2} = \frac{1}{G_{m3}} \frac{C_L}{C_{C2}} \quad (5.70)$$

The transfer function of the amplifier in (5.67) now has a single pole. This means that a suitable value of C_{C1} can be chosen to maximise the gain-bandwidth product, allowing it to reach the same order of magnitude as an optimised two-stage Miller-compensated amplifier. Again G_{m3} must be higher than G_{m2}, so that the compensation elements will be positive. Moreover, it is worth noting that relations (5.68)-(5.70) are independent of G_{m1} and, ideally, the compensation capacitors are also independent of the load capacitor.

Of all the possible solutions that reduce (5.67) to a single-pole function, the one chosen also has the property of providing an inherent pole-zero cancellation for the (open-loop) transfer function of the amplifier containing only the second and third stage. Indeed, by denoting their second pole and (negative) zero as $\widetilde{p}_2$ and $\widetilde{z}$, respectively, these are given by

$$\widetilde{p}_2 = \frac{G_{m3}}{C_L} \tag{5.71}$$

$$\widetilde{z} = \frac{1}{\left(R_{C2} + R_C - \frac{1}{G_{m3}}\right)C_{C2}} \tag{5.72}$$

whose expressions perfectly match if equations (5.68)-(5.70) are used. However, note that the *inner* amplifier, which is closed in the feedback loop by capacitor C_{C1}, is now comprised between the input of the second stage and the common node of R_{C2} and R_C. Therefore, according to our design methodology, we firstly have to check the stability of this feedback loop. The open-loop transfer function of the inner amplifier is

$$A_i(s) = A_{io}\frac{1+\left(R_{C2} - \frac{1}{G_{m3}}\right)C_{C2}s - \frac{R_C C_{C2} C_L}{G_{m3}}s^2}{1+G_{m3}R_2R_3C_{C2}s + R_2R_3C_{C2}C_L s^2} \tag{5.73}$$

where A_{io} is given by (5.50). From (5.73) and using (5.68) and (5.70) we derive the expressions of the unity-gain frequency and those of the second pole and zeros of the inner amplifier

$$\omega_{GBWi} = \frac{G_{m2}}{C_{C2}} \tag{5.74}$$

$$p_{2i} \approx \frac{G_{m3}}{C_L} \tag{5.75}$$

$$z_{1i} \approx \frac{G_{m3}}{C_L - C_{C2}} \tag{5.76}$$

$$z_{2i} \approx -\left(\frac{C_L}{C_{C2}} - 1\right)\frac{G_{m3}}{C_L} \tag{5.77}$$

indicating that the second pole and the first zero remain very close provided that $C_{C2} << C_L$. In this case, the second (RHP) zero tends to G_{m3}/C_{C2} and must be higher than the unity-gain frequency given in (5.74). to ensure stability. Setting the inner phase margin greater than 64° yields $G_{m3} > 2G_{m2}$. The above relation establishes a lower limit for the ratio between G_{m3} and G_{m2}. Under this condition, any value of $C_{C2} << C_L$ ideally ensures the stability of the inner amplifier. A minimum usable value for C_{C2} exists in reality. Compensation capacitors must be greater than the parasitic capacitances at the high-impedance nodes to be valid for development. Besides, and usually more importantly, slew-rate considerations posit the fundamental limit for the minimum value of C_{C2} [PP02], [PPP01].

5.5 REVERSED NESTED MILLER COMPENSATION

When the amplifier is made up of three gain stages and the inner stage is the only inverting one, reversed nested Miller compensation (RNMC) becomes the most suitable technique [EH95].

5.5.1 General Features

Figure 5.12 shows a three-stage amplifier small-signal circuit including reversed nested Miller compensation performed by capacitors C_{C1} and C_{C2}. As usual, parameters G_{mi} and R_i are the i-th stage transconductance and output resistance, respectively. Capacitors C_i represent the equivalent capacitance at the output of each stage, while C_L is the equivalent load capacitor. Since capacitor C_{C2} has no connection with the load capacitor (but only with parasitic capacitor C_2), inner loop stability is virtually achieved for all practical C_{C2} values, and will not be examined. For the same reason, this technique has an inherent bandwidth advantage over other multistage compensation approaches based on the Miller effect.

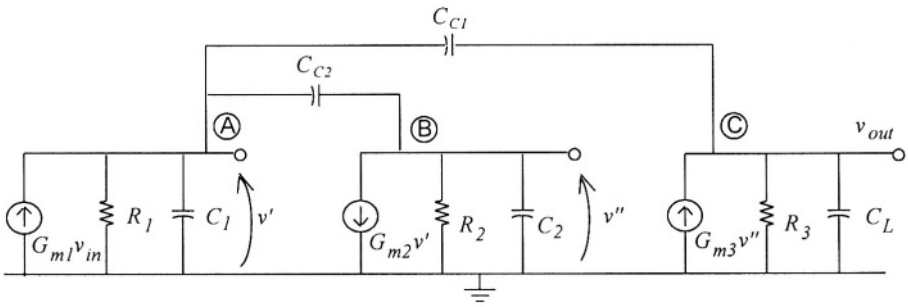

Fig. 5.12. Small-signal model of a three-stage amplifier with reversed nested Miller compensation.

Neglecting second-order terms, the open-loop gain of the circuit in Fig. 5.12 is given by

$$A(s) = A_o \frac{1-\left(\frac{C_{C2}}{G_{m2}} + \frac{C_{C1}}{G_{m2}G_{m3}R_2}\right)s - \frac{C_{C1}C_{C2}}{G_{m2}G_{m3}}s^2}{\left(1+\frac{s}{p_1}\right)\left[1+\left(\frac{C_{C2}C_L}{G_{m3}C_{C1}} - \frac{C_{C2}}{G_{m2}} + \frac{C_{C2}}{G_{m3}}\right)s + \frac{C_{C2}C_L}{G_{m2}G_{m3}}s^2\right]} \quad (5.78)$$

where A_o is the DC open-loop gain equal to $G_{m1}R_1G_{m2}R_2G_{m3}R_3$ and p_1 is the dominant pole due to compensation capacitor C_{C1}. Therefore, the dominant pole and the gain-bandwidth product are equal to those of the nested Miller compensation in (5.47) and (5.48), respectively. Moreover, again as for the NMC, we have two other higher poles (usually complex and conjugates) and two zeros, the lower one on the right-half plane and the other on the left-half plane.

Unlike in the NMC, large values of G_{m3} do not facilitate the task of compensation. If G_{m3} is much higher than G_{m2}, except when considering parasitic capacitances, a pole-zero cancellation occurs which modifies (5.78) into a single pole transfer function. But the pole and the zero involved in this compensation are positive, a condition which is critical for stability. Therefore, we must provide viable compensation procedures also for large values of G_{m3}.

First observe that both zeros can be eliminated by using two voltage or current buffers in series with compensation capacitors to break the forward paths, as illustrated in Fig. 5.13 and 5.14, respectively.

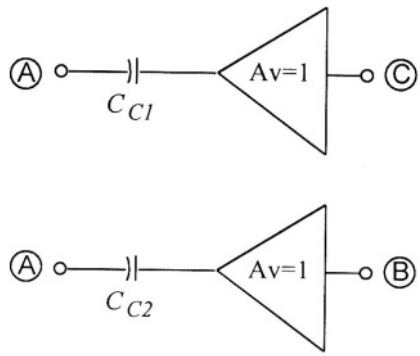

Fig. 5.13. Compensation network with voltage buffers.

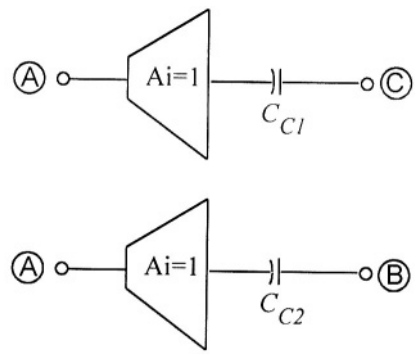

Fig. 5.14. Compensation network with current buffers.

In these two cases (5.78) respectively becomes

$$A(s) = A_o \frac{1}{\left(1 + \frac{s}{p_1}\right)\left(1 + \frac{C_{C2}C_L}{G_{m3}C_{C1}}s\right)} \tag{5.79}$$

$$A(s) = A_o \frac{1}{\left(1 + \frac{s}{p_1}\right)\left[1 + \frac{C_{C2}(C_{C1} + C_L)}{G_{m3}C_{C1}}s\right]} \tag{5.80}$$

It is worth noting that the above expressions have exactly two poles thanks to the action of the ideal buffers. Both second poles are also negative. The second pole in (5.79) can be simply interpreted by analysing the circuit in Fig. 5.15, where the inner amplifier, A_2, through the capacitive network, acts as a G_{m3} multiplier by a factor equal in module to C_{C1}/C_{C2}. The same

considerations hold for the second pole of (5.80). The only difference is that C_{C1} and C_L are now in parallel.

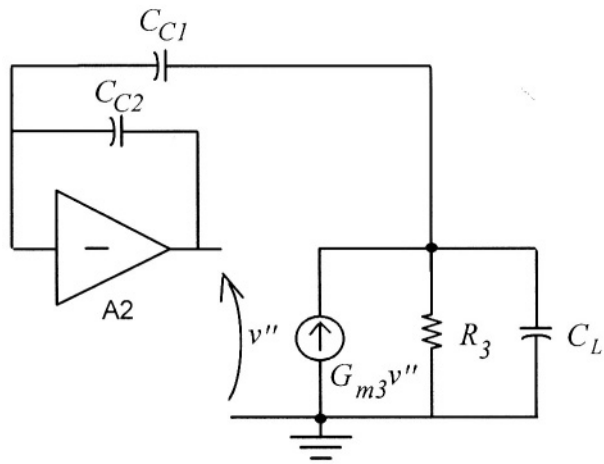

Fig. 5.15. Circuit model for the evaluation of the second pole.

The specified phase margin for the two cases respectively is given by

$$\tan(\phi) = \frac{G_{m3}}{G_{m1}} \frac{C_{C1}^2}{C_{C2}C_L} \tag{5.81}$$

$$\tan(\phi) = \frac{G_{m3}}{G_{m1}} \frac{C_{C1}^2}{C_{C2}(C_{C1} + C_L)} \tag{5.82}$$

Since C_{C1} is set by the required unity-gain bandwidth, and assuming G_{m1}, G_{m3} and C_L to be already set, (5.81) and (5.82) give the needed value of C_{C2}. Although it has been demonstrated here that ideal buffers provide a conceptually simple vehicle for the cancellation of the zeros, we will not stress these approaches any further because of the second-order effects of *real* buffers. In fact, the two approaches, as described above, prove to be inefficient especially in a low-voltage low-power context. Actually, the use of real voltage buffers unacceptably limits the output swing, while real current buffers –matching the requirement of very low input resistance– are expensive in terms of area and power consumption. Fortunately, we will show in 5.5.3 and 5.5.4 that both approaches can be simply modified so as to become suitable for practical applications.

For the sake of completeness, we shall first deal with the nulling resistor technique, which unfortunately is rather difficult to accomplish in a RNMC.

5.5.2 RHP Cancellation with Nulling Resistors

Figure 5.16 shows the compensation network including two nulling resistors, as customarily employed.

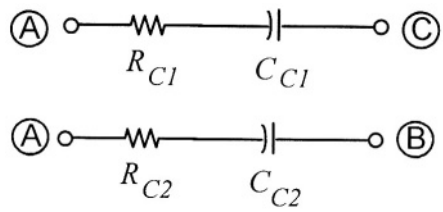

Fig. 5.16. Compensation network with two nulling resistors.

By using this network in the circuit in Fig. 5.7, the numerator of the open-loop gain in (5.78) becomes

$$N(s) = 1 + \left(R_{C1}C_{C1} + R_{C2}C_{C2} - \frac{C_{C2}}{G_{m2}} \right) s - C_{C1}C_{C2} \left(R_{C1}R_{C2} - \frac{1}{G_{m2}G_{m3}} - \frac{R_{C1}}{G_{m2}} \right) s^2 \tag{5.83}$$

in which, as usual, only dominant terms are considered.

It can be shown that (5.83) provides real and negative zeros only with complex matching between R_{C1} and R_{C2} (by setting one of the two resistances equal to zero, it also is easy to verify that the RHP zero cannot be eliminated as the s^2 coefficient is always negative).

A more effective solution is that shown in Fig. 5.17, which uses only one resistor and leads to the following expression of $N(s)$

$$N(s) = 1 + \left[R_C \left(C_{C1} + C_{C2} \right) - \frac{C_{C2}}{G_{m2}} \right] s - \frac{C_{C1}C_{C2}}{G_{m3}} \left(\frac{1}{G_{m2}} - R_C \right) s^2 \tag{5.84}$$

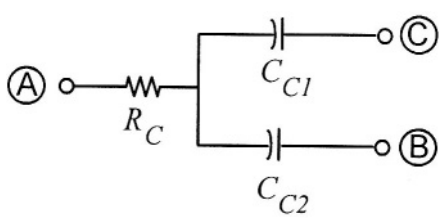

Fig. 5.17. Compensation network with one nulling resistor.

Setting $R_C = 1/G_{m2}$, (5.84) becomes

$$N(s) = 1 + \frac{C_{C1}}{G_{m2}} s \tag{5.85}$$

yielding only one negative zero. Of course, the denominator of the open-loop gain is still the same as in (5.78). In this case, it is convenient to have $G_{m2}=G_{m3}$. As we shall show, this choice allows a pole-zero cancellation to be achieved. Indeed, assuming that also

$$C_{C2}C_L > 4C_{C1}^2 \tag{5.86}$$

meaning that the determinant of the second order factor of (5.78) is positive, it follows that all poles are real and thus (5.78) becomes

$$A(s) = A_o \frac{1}{\left(1 + \frac{s}{p_1}\right)\left(1 + \frac{C_{C2}C_L}{G_{m3}C_{C1}} s\right)} \tag{5.87}$$

For a given phase margin we get

$$C_{C2} = \frac{C_{C1}^2 G_{m3}}{\tan(\phi) C_L G_{m1}} \tag{5.88}$$

Now, by substituting (5.88) in (5.86), condition (5.86) is satisfied if $G_{m2,3}/G_{m1} > 4\tan(\phi)$. Since a phase margin of about 60° is generally required, it follows that the transconductance of both the second and third stage must be at least seven times greater than the transconductance of the first stage. Since transconductances are usually set by other kinds of specifications, the application of this technique is implicitly limited.

5.5.3 RHP Cancellation with One Real Voltage Buffer

To achieve RHP cancellation, we can efficiently make use of only one voltage buffer in the inner loop, as shown in Figure 5.18.

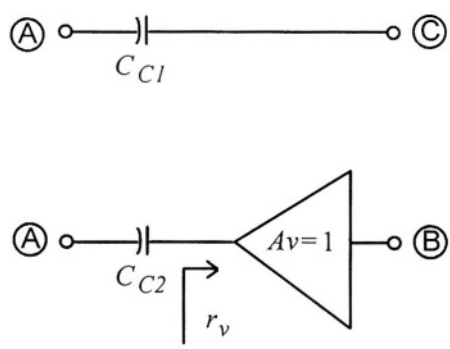

Fig. 5.18. Compensation network with one voltage follower.

By adopting this compensation network the output swing turns out to be completely preserved. In addition, we shall exploit the finite output resistance of the voltage follower to perform some simplifications as described below. Denoting this output resistance as r_v, the loop-gain transfer function is modified to

$$A(s)=\frac{A_o}{\left(1+\frac{s}{p_1}\right)}\cdot\frac{1+\left(r_vC_{C2}-\frac{C_{C1}}{G_{m2}G_{m3}r_v}\right)s-\frac{C_{C1}C_{C2}r_v}{G_{m2}G_{m3}R_2}s^2}{1+\frac{C_{C2}\left(C_L+C_{C1}+G_{m3}r_vC_{C1}\right)}{G_{m3}C_{C1}}s+\frac{r_vC_{C2}C_L}{G_{m2}G_{m3}r_v}s^2}\approx$$

$$\approx\frac{A_o}{\left(1+\frac{s}{p_1}\right)}\cdot\frac{\left(1+r_vC_{C2}s\right)\left(1-\frac{C_{C1}}{G_{m2}G_{m3}R_2}s\right)}{\left[1+\frac{C_{C2}\left(C_L+C_{C1}+G_{m3}r_vC_{C1}\right)}{G_{m3}C_{C1}}s\right]\left[1+\frac{C_{C1}C_Lr_v}{G_{m2}R_2\left(C_L+C_{C1}+G_{m3}r_vC_{C1}\right)}s\right]} \quad (5.89)$$

Relationship (5.89) includes one dominant LHP zero and a RHP zero that is now shifted to a very high frequency (since it is multiplied by the stage gain $G_{m2}R_2$). Moreover, there are two non-dominant poles which are real and negative under the condition (in practice usually met) $G_{m2}R_2C_{C2}>2\ C_L$. These two poles are well approximated by the terms inside the square brackets in the second expression of (5.89). We can use the output resistance of the voltage follower to obtain some forms of simplification, and among the possible alternatives, we can set the value of this resistance equal to the transconductance of the last stage

$$r_v = \frac{1}{G_{m3}} \tag{5.90}$$

Substituting (5.90) in (5.89) the non-dominant poles and the two zeros result as

$$p_2 = \frac{C_{C1}}{(C_L + 2C_{C1})} \frac{G_{m3}}{C_{C2}} \tag{5.91}$$

$$p_3 = \frac{(C_L + 2C_{C1})}{C_L} \frac{G_{m3} G_{m2} R_2}{C_{C1}} \tag{5.92}$$

$$z_1 = \frac{G_{m3}}{C_{C2}} = \frac{C_L + 2C_{C1}}{C_{C1}} p_2 \tag{5.93}$$

$$z_2 = -\frac{G_{m3} G_{m2} R_2}{C_{C1}} \tag{5.94}$$

It is apparent that p_3 and z_2 are at a very high frequency (with $p_3 > |z_2|$) and their contribution to the phase margin can be neglected. Moreover $z_1 > 2p_2$ assures a monotonic behaviour for the loop gain module. The phase margin is then given by

$$\phi = \tan^{-1} \frac{p_2}{\omega_{GBW}} + \tan^{-1} \frac{\omega_{GBW}}{z_1} \tag{5.95}$$

from which we get

$$C_{C2} = C_{C1} \frac{G_{m3}}{G_{m1}} \frac{2\alpha}{(1-\alpha)\tan(\phi) + \sqrt{(1-\alpha)^2 \tan(\phi)^2 - 4\alpha}} \tag{5.96}$$

where parameter α is equal to

$$\alpha = \frac{C_{C1}}{C_L + 2C_{C1}} \tag{5.97}$$

Equation (5.96) can be simplified by observing that for the auspicious condition $C_{C1} < C_L$, parameter α is lower than 1/3 and for practical phase margin values around 60°-70° we have

$$(1-\alpha)^2 \tan(\phi)^2 > 4\alpha \tag{5.98}$$

indicating that the zero can also be neglected when evaluating the phase margin. In conclusion, (5.96) can be approximated by (5.82).

5.5.4 RHP Cancellation with One Real Current Buffer

As can be deduced by returning to Fig. 5.14, the overall feedback current is given by the sum of the currents flowing into the two compensation capacitors. Thus, we need to use only *one* current buffer in the loop to break the forward path, as shown in Fig. 5.19, simplifying design and reducing power and area consumption. Since the overall feedback current is still the same, the loop-gain transfer function is again given by equation (5.80) and (5.82) still holds.

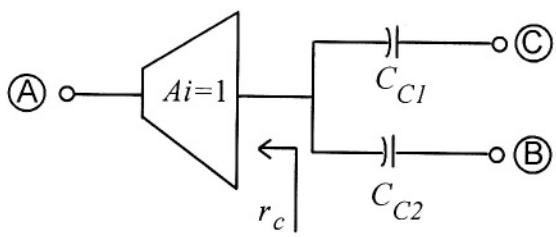

Fig. 5.19. Compensation network with one current follower.

If we consider the finite input resistance of the current follower, r_c, the loop gain will include another pole and two zeros, as shown below

$$A(s) = -A_0 \frac{1 + (C_{C1} + C_{C2}) r_c s + \frac{C_{C1} C_{C2} r_c}{G_{m3}} s^2}{\left(1 + \frac{s}{p_1}\right)\left(1 + \frac{C_{C2}(C_{C1} + C_L)}{G_{m3} C_{C1}} s\right)\left[1 + \frac{C_L C_{C1} r_c}{(C_L + C_{C1}) G_{m2} R_1} s\right]} \tag{5.99}$$

in which the two zeros have a negative real part. Besides, they are real if

$$r_c \le 4 \frac{C_{C1} C_{C2}}{G_{m3} (C_{C1} + C_{C2})^2} \tag{5.100}$$

that gives a higher limit for the current follower input resistance. If this condition is met, the expressions of the two zeros become

$$z_1 \approx \frac{1}{(C_{C1}+C_{C2})r_c} \tag{5.101}$$

$$z_2 \approx \frac{G_{m3}(C_{C1}+C_{C2})}{C_{C1}C_{C2}} \tag{5.102}$$

Choosing the highest value of r_c defined by equality in (5.100), it can be shown that z_2 is four times greater than z_1. Moreover, we have that $p_2 << p_3$ if $C_{C1}/C_{C2} << G_{m2}R_1/4$, which is a condition easily met in practice. Thus the second zero and the third pole in (5.99) are allocated well above the second pole, and do not appreciably modify the phase margin.

Finally, if $C_{C1} > C_{C2}$ we have that $p_2 < z_1$. This means that the first zero does not modify ω_{GBW}, but must be considered when evaluating the phase margin which is

$$\phi = \tan^{-1}\frac{G_{m3}}{G_{m1}}\frac{C_{C1}^{\ 2}}{C_{C2}(C_{C1}+C_L)} + \tan^{-1}\frac{\omega_{GBW}}{z_1} \tag{5.103}$$

and is reduced to (5.82) if $z_1 >> \omega_{GBW}$.

Thanks to the action of the negative zero, the approach adopting a current buffer is preferable to the one using a voltage buffer.