Basic OpAmp Design and Compensation

Chapter 6
6.1 OpAmp applications

Typical applications of OpAmps in analog integrated circuits:
(a) Amplification and filtering
(b) Biasing and regulation
(c) Switched-capacitor circuits

Chapter 6 Figure 01
The classic Two-State OpAmp

The two-stage circuit architecture has historically been the most popular approach to OpAmp design.

It can provide high gain and high output swing.

It is an excellent example to illustrate many important design concepts that are also directly applicable to other designs.

The two-stage refers to the number of gain stages in the OpAmp. The output buffer is normally present only when resistive loads need to be driven. If the load is purely capacitive, it is not needed.

Chapter 6 Figure 02
The classic Two-State OpAmp

The load is assumed capacitive.

The first stage is a pMOS differential pair with nMOS current mirrors. Second stage is a common-source amplifier.

Shown in the diagram are reasonable widths in 0.18um technology (length all made 0.3um). Reasonable sizes for the lengths are usually 1.5 to 10 times of the minimum length (while digital circuits usually use the minimum).
6.1.1 OpAmp gain

For low-frequency applications, the gain is one of the most critical parameters. Note that compensation capacitorCc can be treated open at low frequency.

\[
A_{v1} = -g_{m1}(r_{ds2} \parallel dr_{ds4}) \quad g_{m1} = \sqrt{2\mu_p C_{ox} \left(\frac{W}{L}\right) I_{D1}} = \sqrt{2\mu_p C_{ox} \left(\frac{W}{L}\right) \frac{I_{bias}}{2}}
\]

The second gain stage is simply a common-source gain stage with a p-channel active load, Q6. Its gain is given by

\[
A_{v2} = -g_{m7}(r_{ds6} \parallel r_{ds7})
\]

(6.3)

Overall gain \[A_v = A_{v1} \times A_{v2}\]
Example 6.1 (page 244)

Find the gain of the opamp shown in Fig. 6.3. Assume the power supply is $V_{DD} = 1.8$ V and a purely capacitive load. Assume the process parameters for the 0.18-μm process in Table 1.5.

$I_{D8} = 20 \mu A$. \[ I_{D1} = I_{D2} = I_{D3} = I_{D4} = I_{DS}/2 = (W_5/2W_8)I_{D8} = 100 \mu A \]

$I_{D6} = I_{D7} = (W_6/W_5)I_{DS} = 300 \mu A$

$g_{m1}=g_{m2} = 1.30$ mA/V, and $g_{m7} = 3.12$ mA/V.

$r_{ds} = \frac{1}{\lambda I_{D-sat}} \approx \frac{1}{\lambda I_D}$ \[ r_{ds1} = r_{ds2} = r_{ds3} = r_{ds4} = 37.5 \text{ k}\Omega \quad r_{ds6} = r_{ds7} = 12.5 \text{ k}\Omega \]

$A_{v1} = -g_{m1}(r_{ds2} \parallel r_{ds4}) = -24.4 \text{ V/V}$ \[ A_{v2} = -g_{m7}(r_{ds6} \parallel r_{ds7}) = -19.5 \text{ V/V} \]

It should be noted again that the hand calculation using the approximate equations above is of only moderate accuracy, especially the output resistance calculation on $r_{ds}$. Therefore, later they should be verified by simulation by SPICE/SPECTRE.

However, the benefit of performing a hand calculation is to give an initial (hopefully good) design and also see what parameters affect the gain.
6.1.2 Frequency response: first order model

At frequencies where the comp. capacitor $C_c$ has caused the gain to decrease, but still at frequencies well below the unity-gain frequency of the OpAmp. This is typically referred to as Midband frequencies for many applications.

At these frequencies, we can make some simplifying assumptions. First, ignore all other capacitors except $C_c$, which typically dominates in these frequencies. Second, temporarily neglect $R_c$, which has an effect only around the unity-gain freq. of the OpAmp. The resulting simplified circuit is shown below.
6.1.2 Frequency response: first order model

The second stage introduces primarily a capacitive load on the first stage due to the compensation capacitor, $C_C$. Using Miller’s theorem (Section 4.2.3), one can show that the equivalent load capacitance, $C_{eq}$, at node $v_1$ is given by

$$C_{eq} = C_C(1 + A_{v_2}) \approx C_C A_{v_2} \quad (6.4)$$

The gain in the first stage can now be found using the small-signal model of Figure 4.37, resulting in

$$A_{v_1} = \frac{v_1}{v_{in}} = -g_{m1} Z_{out1} \quad (6.5)$$

$$Z_{out1} = r_{ds2} || r_{ds4} || \frac{1}{sC_{eq}}$$

For midband frequencies, the impedance of $C_{eq}$ dominates, and we can write

$$Z_{out1} \approx \frac{1}{sC_{eq}} \approx \frac{1}{sC_C A_{v_2}}$$

$$A_v(s) \equiv \frac{v_{out}}{v_{in}} = A_{v_2} A_{v_1} \approx A_{v_2} \frac{g_{m1}}{sC_C A_{v_2}} = \frac{g_{m1}}{sC_C}$$

Using the above equation, we can approximate the Unity-Gain frequency as follows:

$$\omega_{td} \approx \frac{g_{m1}}{C_C} = \frac{2I_{D1}}{V_{eff1} C_C} = \frac{I_{D5}}{V_{eff1} C_C}$$

For a fixed $w_{ta}$, power consumption is minimized by small $I_d$, therefore small $V_{eff1}$. 
6.1.2 Frequency response: second order model

In the second-order model, it is assumed that any parasitic poles in the first stage are at frequencies much higher than the \( \omega_{ta} \) and can therefore be ignored (except at the node \( V_1 \)).

\[
R_1 = r_{ds4} \parallel r_{ds2}
\]
\[
C_1 = C_{db2} + C_{db4} + C_{gs7}
\]
\[
R_2 = r_{ds6} \parallel r_{ds7}
\]
\[
C_2 = C_{db7} + C_{db6} + C_{L2}
\]

\( C_{gd2} \) and \( C_{gd4} \) may be included

\( C_{gd6} \) may be included (\( C_{gd7} \) may be lumped to \( C_c \))

Assume \( R_c = 0 \) at first, then

\[
A_v(s) = \frac{V_{out}}{V_{in}} = \frac{g_{m1}g_{m7}R_1R_2(1 - \frac{sC_c}{g_{m7}})}{1 + sa + s^2b}
\]

\[
a = (C_2 + C_c)R_2 + (C_1 + C_c)R_1 + g_{m7}R_1R_2C_c
\]

\[
b = R_1R_2(C_1C_2 + C_1C_c + C_2C_c)
\]

Assume that the two poles are widely separated, then the denom. of \( A_v(s) \) is

\[
D(s) = \left( 1 + \frac{s}{\omega_{p1}} \right) \left( 1 + \frac{s}{\omega_{p2}} \right) \approx 1 + \frac{s}{\omega_{p1}} + \frac{s^2}{\omega_{p1}\omega_{p2}}
\]
From the two poles, increasing $g_{m7}$ is good to separate them more; also increasing $C_c$ makes $w_{p1}$ smaller. Both make the OpAmp more stable.

However, a problem arises from the zero, as it gives negative phase shift in the transfer function, which makes stability difficult. Making $C_c$ large does not help as $w_z$ will reduce too. Increasing $g_{m7}$ helps at the cost of power. $W_{ta}<0.5w_{p2}$ for 65 degrees of phase margin.
6.1.3 Slew rate

The maximum rate at which the output of an OpAmp can change is limited by the finite bias current.

When the inputs change too quickly the OpAmp’s output voltage changes at its maximum rate, called slew rate. In this case, the OpAmp’s response is nonlinear until it is able to resume linear operation without exceeding the slew rate.

Such transient behavior is common in switched-capacitor circuits, where the slew rate is a major factor determining the circuit’s setting time.

Chapter 6 Figure 06
Example 6.4 (page 249)

Consider a closed-loop feedback amplifier with a first-order linear settling time constant of $\tau = 0.2$ $\mu$s and a slew rate of 1 V/$\mu$s. What is the time required for the output to settle when generating to a 10-mV step output with 0.1 mV accuracy? What about a 1-V step output?

The amplifier will slew-rate limit whenever the slope demanded by linear settling exceeds the maximum imposed by slew-rate limiting. For a step height $V_{\text{step}}$, linear settling is exponential.

$$v_o = V_{\text{step}}(1 - e^{-t/\tau}) \quad (6.22)$$

The highest rate of change is observed right at the time of the step.

$$\left. \frac{dv_o}{dt} \right|_{\text{max}} = \left. \frac{dv_o}{dt} \right|_{t=0} = \frac{V_{\text{step}}}{\tau} \quad (6.23)$$

So long as this maximum slope is less than the slew rate, slew-rate limiting is avoided. Hence, the maximum step size that can be tolerated without slew-rate limiting is

$$V_{\text{step, max}} < SR \cdot \tau = 0.2V \quad (6.24)$$

![Diagram showing voltage (V) over time (t) with x and y-axis labels](image)

Case 1:

4.6$\tau = 0.92$ $\mu$s.

Case 2: note that linear settling starts when output $V_o$ reaches 0.8V. Initially slew rate for $(1-0.2)/SR=0.8\text{us}$, then it needs another $-\ln(0.1/200) = 7.6$ time constants. So total

$$0.8 \mu s + 7.6\tau = 2.32 \mu s.$$
6.1.3 Slew rate

When the opamp of Fig. 6.3 is limited by its slew rate because a large input signal is present, all of the bias current of Q₅ goes into either Q₁ or Q₂, depending on whether \( v_{in} \) is negative or positive. When \( v_{in} \) is a large positive voltage, the bias current, \( I_{DS} \), goes entirely through Q₁ and also goes into the current-mirror pair, Q₃, Q₄. Thus, the current coming out of the compensation capacitor, \( C_C \), (i.e., \( I_{D4} \)) is simply equal to \( I_{DS} \) since Q₂ is off. When \( v_{in} \) is a large negative voltage, the current-mirror pair Q₃ and Q₄ is shut off because Q₁ is off, and now the bias current, \( I_{DS} \), goes directly into \( C_C \). In either case, the maximum current entering or leaving \( C_C \) is simply the total bias current, \( I_{DS} \). \(^7\) (In fact, it requires the \( I_{DS} \))

\[
SR = \left. \frac{d v_{out}}{dt} \right|_{max} = I_{C_C} \left|_{max} = \frac{I_{DS}}{C_C} = \frac{2I_{D1}}{C_C}
\]

From first order model

\[
\omega_{ta} = \frac{g_{m1}}{C_C} \quad \Rightarrow \quad SR = \frac{2I_{D1}\omega_{ta}}{g_{m1}}
\]

\[
g_{m1} = \sqrt{2\mu_p C_{ox}(W/L)} I_{D1} \quad \Rightarrow \quad SR = \frac{2I_{D1}}{\sqrt{2\mu_p C_{ox}(W/L)} I_{D1}} \omega_{ta} = V_{eff1} \omega_{ta}
\]

\[
V_{eff1} = \sqrt{\frac{2I_{D1}}{\sqrt{\mu_p C_{ox}(W/L)} I_{D1}}}
\]

Since stability demands that \( \omega_{ta} \) be lower than \( \omega_{p2} \), the only ways of improving the slew rate for a properly-compensated two-stage CMOS opamp is to increase \( V_{eff1} \) or \( \omega_{p2} \).

Assuming a fixed power consumption, and hence fixed bias currents, increasing \( V_{eff1} \) improves the slew rate (6.30) and helps to minimize distortion, but also lowers the transconductance of the input stage which decreases the dc gain (6.1), and increases the equivalent input thermal noise (see Chapter 9).
6.1.4 nMOS or pMOS input stage?

The choice depends on a number of tradeoffs.

First, the gain does not seem to be affected much to first order.

Second, have pMOS input stage allows the second stage be nMOS common-source amplifier to that its $g_m$ can be maximized when high frequency operation is important, as both $w_p$ and $w_t$ are proportional to $g_m$. ($g_m$ of nMOS is larger under the same current and size).

Third, if the third stage of source follower is needed, then an nMOS version is preferable as this will have less voltage drop. (but it is not used when there is only capacitive load).

Fourth, noise is a concern. Typically, pMOS helps reduce the noise.

In summary, when using a two-stage OpAmp, the pMOS input stage is preferred to optimize $w_t$ and minimize noise.
6.1.5 Systematic offset voltage

When designing two-stage OpAmp, the sizes of transistor has to be carefully set to avoid inherent or systematic input offset voltage. When input differential voltage is 0, $V_{GS7}$ should be what is required to make $I_{D7}$ equal to $I_{D6}$.

$$V_{GS7} = \sqrt{\frac{2I_{D6}}{\mu_n C_{ox}(W/L)_{7}}} + V_{tn}$$

Also, note that $V_{GS7} = V_{DS3} = V_{GS4}$

$$V_{GS4} = \sqrt{\frac{2I_{D4}}{\mu_n C_{ox}(W/L)_{4}}} + V_{tn}$$

$$\sqrt{\frac{2I_{D4}}{\mu_n C_{ox}(W/L)_{4}}} = \sqrt{\frac{2I_{D6}}{\mu_n C_{ox}(W/L)_{7}}} \quad \Rightarrow \quad \frac{I_{D4}}{(W/L)_{4}} = \frac{I_{D6}}{(W/L)_{7}}$$

Also

$$\frac{I_{D6}}{I_{D4}} = \frac{I_{D6}}{I_{D5}/2} = \frac{(W/L)_6}{(W/L)_5/2}$$

Finally, we see that the necessary condition to ensure that no input-offset voltage is present is

$$\frac{(W/L)_7}{(W/L)_4} = 2\frac{(W/L)_6}{(W/L)_5}$$

By meeting these constraints, one can achieve a smaller offset voltage (it may still exist due to mis-match of transistors).
6.2 OpAmp compensation

Optimal compensation of OpAmps may be one of the most difficult parts of design. Here a systematic approach that may result in near optimal designs are introduced that applies to many other OpAmps.

Two most popular approaches are dominant-pole compensation and lead compensation.

A further increase in phase margin is obtained by lead compensation which introduces a left half plane zero at a frequency slightly greater than the unity gain frequency \( w_t \). If done properly, this has minimal effect on \( w_t \) but gives an additional 20-30 degrees of phase margin.
6.2.2. Dominant pole compensation

The capacitor, $C_C$, controls the dominant first pole, (i.e., $\omega_{p1}$), and thereby the loop’s unity-gain frequency, $\omega_u$.

$$\omega_t = L_0 \omega_{p1} = \beta g_m / C_C$$

(6.42)

Hence, by properly selecting the value of $C_C$ dominant-pole compensation can be achieved.

Especially if the load capacitor $C_L$ dominants so that the second pole $\omega_{p2}$ is relatively constant when $C_c$ changes (see slide 10).
6.2.2. Lead compensation

This results in a number of design opportunities:

1. One could take \( R_c = 1/g_m \)

2. One can make \( R_c \) larger so that \( \omega_z \) cancels the non-dominant pole (pole-zero canceling), this requires:

\[
R_c = \frac{1}{g_m} \left( 1 + \frac{C_1 + C_2}{C_C} \right)
\]

Unfortunately, \( C_2 \) is often not known a priori, especially when no output stage is present.

3. The third way is to take \( R_c \) even larger so that it is slightly larger than the unity gain frequency that would results if the lead resistor were not present. For example, if the new \( \omega_z \) is 70% higher than \( \omega_t \) \( \omega_z = 1.7 \omega_t \) it will introduce a phase lead of \( \tan^{-1}(1/1.7) = 30^\circ \).

Assuming \( R_c \gg (1/g_m) \), then \( \omega_z \approx 1/(R_c C_C) \). Recall \( \omega_t = \beta g_m/C_C \), choose \( R_c \) according to

\[
R_c \approx \frac{1}{1.7\beta g_m}
\]
6.2.2. Lead compensation

Finally, the lead compensation resistor $R_c$ may be replaced by a transistor operating in the triode region, as illustrated in Fig. 6.10. Transistor $Q_9$ has $V_{DS9} = 0$ since no dc bias current flows through it, and therefore $Q_9$ is deep in the triode region. Thus, this transistor operates as a resistor, $R_c$, with a value given by

$$R_c = r_{ds9} = \frac{1}{\mu_n C_{ox} \left(\frac{W}{L}\right)_9 V_{eff9}}$$

(6.48)

It should be noted here that $r_{ds}$ indicates the drain-source resistance of $Q_9$ when it is in the triode region as opposed to the finite-output impedance of $Q_9$ when it is in the active mode. The same notation, $r_{ds}$, is used to indicate the drain-source resistance in both cases—whether the transistor is in the active or the triode region. One simply has to check which region a transistor is operating in to ensure that the correct equation is used to determine $r_{ds}$. 
6.2.2. Summary of Lead compensation

This approach leads to the following design procedure for compensation of a two-stage CMOS opamp:

\[ \omega_1 = L_0 \omega_{p1} = \beta g_{m1}/C_C \]

\[ \omega_{p2} \approx \frac{g_m^7 C_C}{C_1 C_2 + C_2 C_C + C_1 C_C} \]

\[ \approx \frac{g_m^7}{C_1 + C_2} \]

1. Start by choosing, somewhat arbitrarily, \( C'_C \approx (\beta g_{m1}/g_m^7) C_L \). This initially places the loop’s unity gain frequency (6.42) approximately at the frequency of the second pole (6.20), where it has been assumed that the load capacitance \( C_L \) is dominant.

2. Using SPICE, find the frequency at which a \(-125^\circ\) phase shift exists. Let the gain at this frequency be denoted \( A' \). Also, let the frequency be denoted \( \omega_1 \). This is the frequency that we would like to become the unity-gain frequency of the loop gain.

3. Choose a new \( C_C \) so that \( \omega_1 \) becomes the unity-gain frequency of the loop gain, thus resulting in a 55\(^\circ\) phase margin. (Obtaining this phase margin is the reason we chose \(-125^\circ\) in step 2.) This can be achieved by taking \( C_C \) according to the equation

\[ C_C = C'_C A' \]

This make \( \omega_1 \) smaller while \( \omega_{p2} \) relatively constant if \( C_L \) dominates (6.49)

It might be necessary to iterate on \( C_C \) a couple of times using SPICE.

4. Choose \( R_C \) according to

\[ R_C = \frac{1}{1.7 \omega_1 C_C} \]

(6.50)

This choice will increase the phase margin approximately 30\(^\circ\) resulting in a total phase margin of approximately 85\(^\circ\). It allows a margin of 5\(^\circ\) to account for processing variations without the poles of the closed-loop response becoming real. *This choice is also almost optimum lead compensation for almost any opamp when a resistor is placed in series with the compensation capacitor.* It might be necessary to iterate on \( R_C \)
5. If, after step 4, the phase margin is not adequate, then increase $C_C$ while leaving $R_C$ constant. This will move both $\omega_1$ and the lead zero to lower frequencies while keeping their ratio approximately constant, thus minimizing the effects of higher-frequency poles and zeros which, hopefully, do not also move to lower frequencies. In most cases, the higher-frequency poles and zeros (except for the lead zero) will not move to significantly lower frequencies when $C_C$ is increased.

6. The final step is to replace $R_C$ by a transistor. The size of the transistor can be chosen using equation (6.48), which is repeated here for convenience:

$$R_C = r_{ds9} = \frac{1}{\mu_n C_{ox} \left( \frac{W}{L} \right)_9 V_{eff9}} \quad (6.51)$$

Finally, SPICE can be used again to fine-tune the device dimensions to optimize the phase margin to that obtained in steps 4 and 5.
An opamp has an open-loop transfer function given by

$$A(s) = \frac{A_0(1 + s/\omega_z)}{(1 + s/\omega_{p1})(1 + s/\omega_2)}$$  \hspace{1cm} (6.52)

Here, $A_0$ is the dc gain of the opamp and $\omega_z$, $\omega_{p1}$, and $\omega_2$ are the frequencies of a zero, the dominant pole, and the equivalent second pole, respectively. Assume that $\omega_2 = 2\pi \times 50$ MHz and that $A_0 = 10^4$. The opamp is to be used in a unity-gain configuration so that $\beta = 1$ and $L(s) = A(s)$.

a. Assuming $\omega_z \to \infty$, find $\omega_{p1}$ and the unity-gain frequency, $\omega'_i$, so that the opamp has a unity-gain phase margin of 55°.

b. Assuming $\omega_z = 1.7 \omega'_i$ (where $\omega'_i$ is as found in part (a)), what is the new unity-gain frequency, $\omega'_i$? Also, find the new phase margin.

$$\omega >> \omega_{p1} \quad L(s) = A(s) \approx \frac{A_0(1 + s/\omega_z)}{(s/\omega_{p1})(1 + s/\omega_2)}$$

(a) For $\omega_z \to \infty$

$$\angle A(j\omega'_i) = -90° - \tan^{-1}(\omega'_i/\omega_2) = -125°$$

$$\tan^{-1}(\omega'_i/\omega_2) = 35° \quad \Rightarrow \omega'_i = 2.2 \times 10^8 \text{ rad/s} = 2\pi \times 35 \text{ MHz}$$

$$\frac{A_0}{(\omega'_i/\omega_{p1})\sqrt{1 + (\omega'_i/\omega_2)^2}} = 1 \quad \Rightarrow \omega_{p1} = \frac{\omega'_i\sqrt{1 + (\omega'_i/\omega_2)^2}}{A_0} = 2\pi \times 4.28 \text{ kHz}$$

Or we can simply estimate $\omega_{p1}$ equal to $\omega'_i/A_0 = 3.5 \text{ kHz}$
Example 6.7 (page 258)

An opamp has an open-loop transfer function given by

\[ A(s) = \frac{A_0 (1 + s/\omega_z)}{(1 + s/\omega_p)(1 + s/\omega_2)} \]  \hspace{1cm} (6.52)

Here, \( A_0 \) is the dc gain of the opamp and \( \omega_z, \omega_p, \) and \( \omega_2 \) are the frequencies of a zero, the dominant pole, and the equivalent second pole, respectively. Assume that \( \omega_2 = 2\pi \times 50 \text{ MHz} \) and that \( A_0 = 10^4 \). The opamp is to be used in a unity-gain configuration so that \( \beta = 1 \) and \( L(s) = A(s) \).

a. Assuming \( \omega_z \to \infty \), find \( \omega_p \) and the unity-gain frequency, \( \omega'_1 \), so that the opamp has a unity-gain phase margin of \( 55^\circ \).

b. Assuming \( \omega_z = 1.7 \omega'_1 \) (where \( \omega'_1 \) is as found in part (a)), what is the new unity-gain frequency, \( \omega_1 \)? Also, find the new phase margin.

\[ \omega >> \omega_p, \quad L(s) = A(s) \approx \frac{A_0 (1 + s/\omega_z)}{(s/\omega_p)(1 + s/\omega_2)} \]

(b) First, we set

\[ \omega_z = 1.7 \omega'_1 = 2\pi \times 59.5 \text{ MHz} \]

To find the new unity-gain frequency, setting \( |A(j\omega_1)| = 1 \)

\[ \frac{A_0 \sqrt{1 + (\omega_1/\omega_z)^2}}{(\omega_1/\omega_p) \sqrt{1 + (\omega_1/\omega_2)^2}} = 1 \quad \Rightarrow \omega_1 = \frac{A_0 \omega_p \sqrt{1 + (\omega_1/\omega_z)^2}}{\sqrt{1 + (\omega_1/\omega_2)^2}} \]

\[ \omega_1 = 2\pi \times 39.8 \text{ MHz} \]

\[ \angle A(j\omega_1) = -90^\circ + \tan^{-1}(\omega_1/\omega_z) - \tan^{-1}(\omega_1/\omega_2) = -95^\circ \]

a phase margin of \( 85^\circ \)
6.2.3 Making compensation independent of process and temperature

In a typical process, the ratios of all $g_m$s remain relatively constant over process and temperature variation since the $g_m$s are all determined by the same biasing network. ($\mu_n/\mu_p$ is relatively constant tool)

Also, mostly the capacitors also track each other or remain relatively constant.

The lead zero is at a frequency given by

$$\omega_z = \frac{-1}{C_C(1/g_m - R_C)}$$

Thus, if $R_C$ can also be made to track the inverse of transconductances, and in particular $1/g_m$, then the lead zero will also be proportional to the transconductance of $Q_7$. As a result, the lead zero will remain at the same relative frequency with respect to $\omega_1$ and $\omega_{p2}$, as well as all other high-frequency poles and zeros. In other words, the lead compensation will be mostly independent of process and temperature variations.

Recall that $R_C$ is actually realized by $Q_9$, and therefore we have

$$R_C = r_{ds9} = \frac{1}{\mu_n C_{ox}(W/L)_9 V_{eff9}}$$

$$g_m7 = \mu_n C_{ox}(W/L)_7 V_{eff7}$$

Thus, the product $R_C g_m7$, which we want to be a constant, is given by

$$R_C g_m7 = \frac{(W/L)_7 V_{eff7}}{(W/L)_9 V_{eff9}}$$

So then we need to make sure that $V_{eff9}/V_{eff7}$ is independent of process and temperature variations. It may be made constant by deriving $V_{GS9}$ from the same biasing network used to derive $V_{GS7}$. (see the circuit in next slide)
First, we need to make $V_a = V_b$, which is possible is $V_{\text{eff}13} = V_{\text{eff}7}$, i.e.

\[
\sqrt{\frac{2I_{D7}}{\mu_n C_{\text{ox}}(W/L)_7}} = \sqrt{\frac{2I_{D13}}{\mu_n C_{\text{ox}}(W/L)_{13}}} \quad \rightarrow \quad \frac{I_{D7}}{I_{D13}} = \frac{(W/L)_7}{(W/L)_{13}}
\]

Also note the ratio $I_{D7}/I_{D13}$ is set from the current mirror pair $Q_6$, $Q_{11}$, resulting in

\[
\frac{I_{D7}}{I_{D13}} = \frac{(W/L)_6}{(W/L)_{11}}
\]

Thus, to make $V_{\text{eff}13} = V_{\text{eff}7}$, we need to satisfy the following relationship:

\[
\frac{(W/L)_6}{(W/L)_7} = \frac{(W/L)_{11}}{(W/L)_{13}}
\]

The note that once $V_a = V_b$, then $V_{GS12} = V_{GS9}$, which mean $V_{\text{eff}12} = V_{\text{eff}9}$,

\[
\frac{V_{\text{eff}7}}{V_{\text{eff}9}} = \frac{V_{\text{eff}13}}{V_{\text{eff}12}} = \sqrt{\frac{2I_{D13}}{\mu_n C_{\text{ox}}(W/L)_{13}}} = \sqrt{\frac{(W/L)_{12}}{(W/L)_{13}}}
\]

So finally, we have

\[
R_c g_m = \frac{(W/L)_7 V_{\text{eff}7}}{(W/L)_9 V_{\text{eff}9}} = \frac{(W/L)_7}{(W/L)_9} \sqrt{\frac{(W/L)_{12}}{(W/L)_{13}}}
\]
6.3 Advanced current mirrors: wide-swing

As MOS technologies migrate to shorter lengths, it becomes difficult to achieve large as \( r_{ds} \) is smaller due to short channel length.

One way to cope with that is to use cascode current mirrors to have a large impedance, but conventional ones reduce the signal swing, which may not be acceptable in low-voltage applications.

One circuit proposed is the wide-swing cascode current mirror that does not limit the signal swing as much as the conventional one. The basic idea is to bias the drain source voltages of transistor Q1 and Q3 to be close to the minimum possible without going to triode region.
6.3 Advanced current mirrors: wide-swing

Q3 and Q4 acts like a single-diode connected transistor to create the gate source voltage for Q3. Including Q4 helps lower the Vds3 so that it matches Vds2. Other than that, Q4 has little effect on the circuit’s operation.

\[ V_{\text{eff}} = V_{\text{eff2}} = V_{\text{eff3}} = \sqrt{\frac{2I_D}{\mu_n C_{ox}(W/L)}} \]

Assume \( I_{D2} = I_{D3} = I_{D5} \)

since Q5 has the same drain current but is \((n + 1)^2\) times smaller, \( V_{\text{eff5}} = (n + 1)V_{\text{eff}} \)

\[ V_{\text{eff1}} = V_{\text{eff4}} = nV_{\text{eff}} \]

Thus, \( V_{G5} = V_{G4} = V_{G1} = (n + 1)V_{\text{eff}} + V_{tn} \)

Furthermore, \( V_{DS2} = V_{DS3} = V_{G5} - V_{GS1} = V_{G5} - (nV_{\text{eff}} + V_{tn}) = V_{\text{eff}} \)

\( V_{GS1} = V_{GS4} \)

This drain-source voltage puts both Q2 and Q3 right at the edge of the triode region.

Thus, the minimum allowable output voltage is now

\[ V_{\text{out}} > V_{\text{eff1}} + V_{\text{eff2}} = (n + 1)V_{\text{eff}} \]

\[ V_{\text{out}} > 2V_{\text{eff}} \] If \( n=1 \)

Also we need \( V_{DS4} > V_{\text{eff4}} = nV_{\text{eff}} \)

\[ V_{DS4} = V_{G3} - V_{DS3} = (V_{\text{eff}} + V_{tn}) - V_{\text{eff}} = V_{tn} \]

not a difficult requirement.
6.3 Advanced current mirrors: wide-swing

In most applications, it is desirable to make \((W/L)5\) smaller than that given in the Figure so that Q2 and Q3 can be biased with a slightly larger Vds. This would help counter the body effect of Q1 and Q4, which have their Vt increased.

To save power consumption, Ibias and Q5 size can be scaled down a little bit while keeping the same gate voltage.

Also, it may be wise to make the length of Q3 and Q2 larger than the minimum and that of Q1 and Q4 even larger since Q1 often sees a larger voltage Vout. This helps reduce short-channel effects.
6.3.2 Enhanced output impedance CM and Gain boosting

The basic idea is to use a feedback amplifier to keep the drain-source voltage across Q2 as stable as possible, irrespective of the output voltage.

From small-signal analysis, \( I_x = g_m v_{gs} + \frac{(V_x - V_s)}{r_{ds1}} \), \( V_{gs} + V_s = A(0 - V_s) \), \( V_x = I_x * r_{ds2} \)

\[
R_{out} \approx g_m r_{ds1} r_{ds2} (1 + A)
\]

Note that the stability of the feedback loop comprised of A and Q1 must be verified.
6.3.2 Enhanced output impedance CM and Gain boosting

This technique can also be applied to increase the Rout of a cascode gain stage (the small signal current \(-g_{m2}v_{in}\) must go through \(R_{out}\) and \(C_L\)).

\[
A_V(s) = \frac{V_{out}(s)}{V_{in}(s)} = -g_{m2}\left(\frac{1}{sC_L} + R_{out}(s)\right) \quad R_{out}(s) = g_{m1}r_{ds1}r_{ds2}(1 + A(s))
\]

Comparing the DC gain only, it can be seen that it is a factor of \((1+A)\) larger than the conventional cascode amplifier discussed in Chapter 3.

To realize this gain, note that the \(I_{bias}\) current source must be similarly enhanced to achieve comparable output impedance as \(R_{out}\).
6.3.2 Sackinger’s design

The feedback amplifier in this case is realized by transistor Q3 and Q1. Note that Q3 is a CS amplifier, therefore the gain is \( g_m r_{ds3} / 2 \) if \( I_{B1} \) has an output impedance of \( r_{ds3} \).

So the total output impedance from the drain of Q1 is:

\[
r_{out} \approx \frac{g_m g_m r_{ds1} r_{ds2} r_{ds3}}{2}
\]

The circuit consisting of Q4, Q5 and Q6, \( I_{in} \) and \( I_{B2} \) operates like a diode-connected transistor, but its main purpose is to match those transistors in the output circuitry so that all transistors are biased accurately and \( I_{out} = I_{in} \).

One major limitation is that the signal swing is significantly reduced due to Q2 and Q5 being biased to have drain-source voltages much larger ( \( V_{DS2} = V_{DS5} = V_{eff3} + V_{tn} \) ).
6.3.3. Wide-swing current mirror with enhanced output impedance

Such a circuit is very similar to the Sachinger’s design, except that diode-connected transistors used as level shifters Q4 have been added in front of the CS amplifiers.

The current density of most transistors (except Q3 and Q7) are about the same, $V_{eff}$, and that of Q3, Q7, $2V_{eff}$. So

$$V_{G3} = 2V_{eff} + V_{In}$$
$$V_{DS2} = V_{S4} = V_{G3} - V_{GS4} = (2V_{eff} + V_{In}) - (V_{eff} + V_{In}) = V_{eff}$$
$$V_{out} > V_{DS2} + V_{eff1} = 2V_{eff}$$

Two issues with this circuit: 1. power consumption may be large, 2. additional poles introduced by the enhanced circuitry may be at lower frequencies.

![Diagram](image)
6.3.3. Wide-swing current mirror with enhanced output impedance

A variation of the previous circuit is shown below. It reduces the power, but matching is poorer.

Note that Q2 in previous circuit is split to Q2 and Q5 in this circuit.

It is predicted that this current may be more used when power supply voltage is smaller or larger gains are desired.
6.3.4 Summary of improved current mirrors

When using the OpAmp-enhanced current mirrors, it may be necessary to add local compensation capacitors to the enhancement loops to prevent ringing during transients.

Also, the settling time may be increased (to tradeoff with large gain).

Many other current mirrors exist, each having its own advantages and disadvantages. Which one to use depends on the requirements of the specific application.

OpAmps may be designed using any of the current mirrors, therefore we can use the following symbol without showing the specific implementation of the current mirror.

Just one specific implementation of the current mirror in (a)
6.4 Folded-cascode OpAmp

Many modern OpAmps are designed to drive only capacitive loads. In this case, it is not necessary to use a voltage buffer to obtain a low output impedance. So it is possible to realize OpAmps with higher speeds and larger signal swings than those that drive resistive loads.

These OpAmps are possible by having only a single high-impedance node at the output. The admittance seen at all other nodes in these OpAmps are on the order of $1/g_m$, and in this way the speed of OpAmp is maximized.

With these OpAmps, compensation is usually achieved by the load capacitance $C_L$. As $C_L$ gets larger, these OpAmps gets more stable but also slower.

One of the most important parameters of these modern OpAmps is $g_m$ (ratio of output current over input voltage), therefore they are sometimes referred to as Operational Transconductance Amplifiers (OTA).

A simple first order small-signal model for an OTA may be shown below:
Operational Transconductance Amplifier (OTA)

Chapter 6 Figure 19
Folded-cascode OpAmp

A differential-input single-ended output folded-cascode OpAmp is shown below. The current mirror in the output side is a wide-swing cascode one, which increases the gain.

The basic idea of the FC-OpAmp is to apply cascode transistors to the input differential pair but using transistors opposite in type from those used in the input stage. (i.e. Q1, Q2 nMOS and Q5, Q6 pMOS). This arrangement allows the output to be the same as the input bias voltage.

The gain could be large due to large output impedance. If even larger gain is desired, one can use gain-enhancement techniques to Q5-Q8 as described in 6.3.2.
Folded-cascode OpAmp

The single-ended output FC-OpAmp can be converted to a fully-differential one (to be detailed later).

A biasing circuit can be included to replace $I_{bias1}$, $I_{bias2}$ and connect to $V_{B1}$ and $V_{B2}$.

The two extra transistors Q12 and Q13 can increase slew rate performance and prevent the drain voltages of Q1 and Q2 from having large transients thus allowing the OpAmp to recover faster following a slew rate condition.

The compensation is realized by the load capacitor $C_L$ (dominant pole compensation). When $C_L$ is small, it may be necessary to add additional capacitor in parallel with the load. If lead compensation is to be used, then a resistor is in series with $C_L$.

DC biasing: note

$I_{D3/4} = I_{D1/2} + I_{D5/6}$
6.4.1 Small-signal analysis

In small-signal analysis, the small-signal current from Q1 goes directly from source to drain and to $C_L$, while that of Q2 indirectly through Q5 and current mirror of Q7-Q10 to $C_L$. (assuming $1/gm5/6$ much larger than $rds3$ and $rds4$).

Note that these small-signal currents go through different path to the output, therefore their transfer function are different (due to the pole/zero caused by the current mirror for small-signal current of Q2). However, usually, these pole/zero are much larger than the unity-gain frequency of OpAmp and may be ignored.

So an approximate gain transfer function is:

$$A_V = \frac{V_{out}(s)}{V_{in}(s)} = g_{m1} Z_L(s)$$

$$A_V = \frac{g_{m1} r_{out}}{1 + s r_{out} C_L}$$

$Z_L$ is the parallel of impedance at drain of Q6, Q8, and $C_L$.

At high frequencies, $A_v$ is approximated as

$$A_V \approx \frac{g_{m1}}{s C_L}$$

unity-gain frequency of the opamp is found to be

$$\omega_{ta} = \frac{g_{m1}}{C_L}$$
6.4.1 Small-signal analysis

The first-order model shows close to 90 degrees of phase margin.

To maximize bandwidth, it is desirable to increase $g_m$ by using nMOS transistors, which means larger DC current on Q1/2 (Having large $g_m$ for Q1/2 also help reduce noise). Smaller currents on Q5/6 helps increase $r_{out}$, which increases the DC gain. (the current ratio between them has a practical limit of 4 to 5.)

For more detailed analysis, the second pole is associated with the time constants at the source terminals of Q5/Q6. At high frequencies, the impedance is on the order of $1/gm5/6$, which in this case is relatively large due to smaller current. (so one can have larger currents in order to push this pole away and minimizing the capacitance is important too).

In summary, if the phase margin of a folded-cascode opamp with feedback is insufficient, one has two choices:

a. Add an additional capacitance in parallel with the load to decrease the dominant pole. This improves phase margin without additional power consumption and without effecting the dc gain, but decreases amplifier bandwidth and increases area.

b. Increase the current and device widths in the output stage, $I_{DS}$ and $I_{DS}$. This will increase the second pole frequency and improve phase margin, but sacrifices dc gain and increases power consumption.

C. Lead compensation

\[
A_V = \frac{g_m1}{\frac{1}{r_{out}} + \frac{1}{R_C + 1/sC_L}} \approx \frac{g_m1(1 + sR_CC_L)}{sC_L}
\]

$R_C$ can be chosen to place a zero at 1.7 times the unity-gain frequency.
6.4.2 Slew rate

Diode-connected transistors Q12/13 are turned off during normal operation (as $V_{gd3/4} < V_{tp}$) and have almost no effect on the OpAmp. However, they improve the operation during slew rate limiting.

If they are not present, then when slew rate occurs, all bias current of Q4 go to Q5 and out of $C_L$ through the mirror (at the same time Q6 conducts zero current in most cases). The slew rate $SR = \frac{I_{D4}}{C_L}$.

At this time, since all $I_{bias2}$ is diverted through Q1 and it is usually larger than $I_{D3}$, both Q1 and $I_{bias2}$ go into triode region, causing $I_{bias2}$ to decrease until it is equal to $I_{D3}$. As a result, the drain voltage of Q1 approaches ground. When OpAmp is back to normal operation, drain voltage of Q1 must slew back to the original biasing voltage, and this additional slewing increases distortion and transient delay.

If Q12/13 were included, then when slew rate occurs (as the above case), Q12 conducts extra current from Q11 and also the current on Q3/4 increases, which eventually makes the sum of $I_{D12}$ and $I_{D3}$ equal to $I_{bias2}$. On the other hand, $I_{D3/4}$ increment also make the slew rate larger.
Example 6.9 (page 272)

Find reasonable transistor sizes for the folded-cascode opamp shown in Fig. 6.20 to satisfy the following design parameters. Also find the opamp’s unity-gain frequency (without feedback) and slew rate, both without and with the clamp transistors.

- Assume the process parameters for the 0.18-μm process in Table 1.5, a single 1.8-V power supply, and limit the current dissipation of the opamp to 0.4 mA.
- Set the ratio of the current in the input transistors to that of the cascode transistors to be 4:1. Also, set the bias current of Q_{11} to be 1/10th that of Q_3 (or Q_4) such that its current can be ignored in the power dissipation calculation.
- The maximum transistor width should be 180 μm and channel lengths of 0.4 μm should be used in all transistors.
- All transistors should have effective gate-source voltages of around 0.24 V except for the input transistors, whose widths should be set to the maximum value of 180 μm. Also, round all transistor widths to the closest multiple of 2 μm, keeping in mind that if a larger transistor is to be matched to a smaller one, the larger transistor should be built as a parallel combination of smaller transistors.
- Finally, assume the load capacitance is given by C_L = 2.5 pF.

\[
I_{\text{total}} = I_{D_3} + I_{D_4} = 2(I_{D_1} + I_{D_6}) = 2(4I_B + I_B) = 10I_B
\]

\[
I_{D_3} = I_{D_4} = 5I_{D_5} = 200 \mu A \quad I_{D_1} = I_{D_2} = 4I_{D_5} = 160 \mu A
\]

\[
\text{all transistor channel lengths be 0.4 μm}
\]

This choice allows us to immediately determine the sizes of most transistors using

\[
\frac{W}{L} = \frac{2I_{D_i}}{\mu_i C_{ox} V_{eff}^2}
\]

<table>
<thead>
<tr>
<th>Q_1</th>
<th>180/0.4</th>
<th>Q_6</th>
<th>8/0.4</th>
<th>Q_{11}</th>
<th>4/0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q_2</td>
<td>180/0.4</td>
<td>Q_7</td>
<td>2/0.4</td>
<td>Q_{12}</td>
<td>4/0.4</td>
</tr>
<tr>
<td>Q_3</td>
<td>40/0.4</td>
<td>Q_8</td>
<td>2/0.4</td>
<td>Q_{13}</td>
<td>4/0.4</td>
</tr>
<tr>
<td>Q_4</td>
<td>40/0.4</td>
<td>Q_9</td>
<td>2/0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q_5</td>
<td>8/0.4</td>
<td>Q_{10}</td>
<td>2/0.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pre-set to maximum in order to maximize gm

Derived from Q3/4

Arbitrarily set equal to Q11/12
the transconductance of the input transistors would be given by

$$\sqrt{2I_{D1}}\mu_nC_{ox}(W/L)_{1} = 6.24 \text{ mA/V}$$

$$\omega_{ta} = \frac{g_{m1}}{C_L} = 1.6 \times 10^9 \text{ rad/s} \Rightarrow f_{ta} = 255 \text{ MHz}$$

The slew rate without the clamp transistors is given by

$$\text{SR} = \frac{I_{D4}}{C_L} = 80 \text{ V/\mu s}$$

When the clamp transistors are included, during slew-rate limiting, we have

$$I_{D12} + I_{D3} = I_{bias2}$$
$$I_{D3} = 10I_{D11}$$
$$I_{D11} = 20 \mu A + I_{D12}$$

$$I_{D11} = \frac{I_{bias2} + 20 \mu A}{11} = 30.9 \mu A$$

$$I_{D3} = I_{D4} = 10I_{D11} = 0.309 \text{ mA}$$

$$\text{SR} = \frac{I_{D4}}{C_L} = 124 \text{ V/\mu s}$$
Example 6.10 (page 274)

The opamp in Example 6.9 is simulated and found to have an equivalent second pole frequency at $\omega_{eq} = 2\pi \cdot 365$ MHz. Select a value for the lead compensation resistor, $R_C$, to provide a phase margin of $85^\circ$ when the opamp is used in a unity-gain configuration.

**Solution**

The phase margin without $R_C$ is given by

$$\text{Phase Margin} = 90^\circ - \tan^{-1}\left(\frac{\omega_l}{\omega_{eq}}\right) = 90^\circ - \tan^{-1}\left(\frac{2\pi \cdot 255 \text{ MHz}}{2\pi \cdot 365 \text{ MHz}}\right) = 55^\circ$$

In order to increase this by $30^\circ = \tan^{-1}(1/1.7)$, the lead compensation zero must be placed at $1/R_C C_L = 1.7 \omega_l$. Since $\beta = 1$, we have $\omega_l = \omega_{ta}$. Hence, a reasonable size for $R_C$ in series with $C_L$ is given by

$$R_C = \frac{1}{1.7 C_L \omega_l} = \frac{1}{1.7 g_{m1}} = 147 \Omega \quad (6.118)$$
6.5 Current mirror OpAmp

Another popular OpAmp when driving only on-chip capacitive loads is the current-mirror OpAmp. Note that at the Q2 side, more current mirrors need to be used to provide current $K_{I_2}=K_{I_1}$.

Also, it can be seen that all internal nodes have low impedance except the output node. By using proper current mirrors with high output impedance, good gain can be achieved.

The overall transfer function of this OpAmp closely approximate dominant-pole operation.

$$A_v = \frac{V_{out}(s)}{V_{in}(s)} = K_{g_{m1}}Z_L(s) = \frac{K_{g_{m1}}r_{out}}{1 + sr_{out}C_L} \approx \frac{K_{g_{m1}}}{sC_L}$$

unity-gain frequency $\omega_{1a} = \frac{K_{g_{m1}}}{C_L} = \frac{2K_{I_{D1}}}{C_L V_{eff,1}}$

![Figure 21](Image)

![Figure 22](Image)
If the power dissipation is specified, the total current, \( I_{\text{total}} = (3 + K)I_{D1} \),
\[
\omega_{ta} = \frac{K}{(3 + K) V_{\text{eff,i}} C_L} \frac{2I_{\text{total}}}{I_{\text{total}}}
\]

It can be seen that larger \( K \) increases the unity-gain frequency assuming the load capacitor dominates the time constants. Larger \( K \) also increases the gain. A typical upper limit for \( K \) is 5.

A detailed analysis reveal important nodes for determining the non-dominant poles, at the drain of Q1 first and drain of Q2 and Q9 secondly. Larger \( K \) increases the capacitances at these nodes while also increases the resistance (assuming a fixed \( I_{\text{total}} \)), which reduce the non-dominant poles. In this case, then \( C_L \) has to be increased to maintain a large phase margin. So, \( K \) should not be too large, i.e. \( K \leq 2 \) usually.

During slew rate, all of the bias current \( I_b \) of the first stage is diverted through Q1/2 and amplified by the current mirror gain to the output. The total current to charge/discharge the load is \( KI_b \). So the slew rate is
\[
SR = \frac{KI_b}{C_L}
\]

Due primarily to the larger unity-gain frequency and slew rate, the current-mirror OpAmp may be preferred over the folded-cascode OpAmp. However, one has to be careful that the current-mirror OpAmp has larger input noise as well, as its input stage is biased at a lower portion of the total bias current and therefore a relatively smaller \( g_m \) given the same power consumption.
Example 6.11 (page 277)

Assume the current-mirror opamp shown in Fig. 6.22 has all transistor lengths equal to 0.4 \( \mu \text{m} \) and transistor widths as given in Table 6.2. Notice that \( K = 2 \). Assume the process parameters for the 0.18-\( \mu \text{m} \) process in Table 1.5, a single 1.8-V power supply, and that the opamp’s total current dissipation is 0.4 mA. The load capacitance is \( C_L = 2.5 \, \text{pF} \).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Q₁</td>
<td>90/0.4</td>
</tr>
<tr>
<td>Q₂</td>
<td>90/0.4</td>
</tr>
<tr>
<td>Q₃</td>
<td>20/0.4</td>
</tr>
<tr>
<td>Q₄</td>
<td>20/0.4</td>
</tr>
<tr>
<td>Q₅</td>
<td>20/0.4</td>
</tr>
<tr>
<td>Q₆</td>
<td>20/0.4</td>
</tr>
<tr>
<td>Q₇</td>
<td>20/0.4</td>
</tr>
<tr>
<td>Q₈</td>
<td>40/0.4</td>
</tr>
<tr>
<td>Q₉</td>
<td>20/0.4</td>
</tr>
<tr>
<td>Q₁₀</td>
<td>40/0.4</td>
</tr>
<tr>
<td>Q₁₁</td>
<td>4/0.4</td>
</tr>
<tr>
<td>Q₁₂</td>
<td>8/0.4</td>
</tr>
<tr>
<td>Q₁₃</td>
<td>4/0.4</td>
</tr>
<tr>
<td>Q₁₄</td>
<td>8/0.4</td>
</tr>
</tbody>
</table>

Find the slew rate and the unity-gain frequency, assuming the equivalent second pole does not dominate. Estimate the equivalent second pole. Would it be necessary to increase \( C_L \) if a 75° phase margin were required with \( \beta = 1 \) and without using lead compensation? What if lead compensation were used?

\[
I_b = \frac{2I_{\text{total}}}{(3 + K)} = \frac{2(0.4 \, \text{mA})}{(3 + K)} = 160 \, \mu\text{A}
\]

\[
\omega_{la} = \frac{Kg_{m1}}{C_L} = 1.6 \times 10^9 \text{ rad/s} \Rightarrow f_{la} = 255 \, \text{MHz}
\]

\[
\text{SR} = \frac{KI_b}{C_L} = 128 \, \text{V/\mu s}
\]

\( V_{\text{eff1}} \) can be estimated to be about 51 mV so that \( g_m = 3.14 \, \text{mA/V} \) is sort of maximized.

Comparing to the previous example on FC-OpAmp with the same power and load, the current-mirror OpAmp can have better bandwidth and SR if \( K \) is made larger.

The dominant node almost certainly will occur at the drain of \( Q_1 \). The impedance at this node is given by

\[
R_1 = \frac{1}{g_{m5}} = 1.3 \, \text{k}\Omega
\]

the capacitance will be primarily due to the gate-source capacitances of \( Q_5 \) and \( Q_8 \).

\[
C_1 = C_{gs5} + C_{gs8} = (1 + K)C_{gs5} = (1 + K)(2/3)C_{ox}W_5L_5 = 0.14 \, \text{pF}
\]
If 75 degrees of phase margin is used, the unity-gain frequency must be 0.27 times of \( p_{2eq} \) or 126MHz, so the \( C_L \) must be increased from 2.5pF to 5pF to reduce from 255MHz to 126MHz.

If lead compensation is used, then unity-gain frequency can be designed to be 0.7 times of \( p_{2eq} \) so that 55 degrees of phase margin is achieved. Then, lead compensation can be used to achieve another 20 to 30 degrees of phase margin. Also, no additional load capacitance is necessary reducing the circuit area.
6.6 Linear settling time revisited

We saw in Chapter 5 that the time constant for linear settling time was equal to the inverse of $\omega_{-3\,\text{dB}}$ for the closed-loop circuit gain. We also saw that $\omega_{-3\,\text{dB}}$ is given by the relationship

$$\omega_{-3\,\text{dB}} = \omega_t$$

(6.137)

However, while for the classical two-stage CMOS opamp the unity-gain frequency remains relatively constant for varying load capacitances, the unity-gain frequencies of the folded-cascode and current-mirror amplifiers are strongly related to their load capacitance. As a result, their settling-time performance is affected by both the feedback factor as well as the effective load capacitance.

<table>
<thead>
<tr>
<th>Amplifier Type</th>
<th>$\omega_{ta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>classical two-stage CMOS opamp</td>
<td>$\frac{g_m}{C_C}$</td>
</tr>
<tr>
<td>folded-cascode opamp,</td>
<td>$\frac{g_m}{C_L}$</td>
</tr>
<tr>
<td>a current-mirror opamp,</td>
<td>$\frac{Kg_m}{C_L}$</td>
</tr>
</tbody>
</table>

Recall from Chapter 5 the 3db bandwidth of the closed loop amplifier is the unity-gain frequency of the loop gain, which is $\beta$ times the unity-gain frequency of the OpAmp, i.e.

$$\omega_t = \beta \omega_{ta}.$$
6.6 Linear settling time revisited

To determine the –3-dB frequency of a closed-loop opamp, consider the general case shown in Fig. 6.23. At the opamp output, $C_{\text{load}}$ represents the capacitance of the next stage that the opamp must drive, while $C_C$ is a compensation capacitance that might be added to maintain a sufficient phase margin. At the input side, $C_p$ represents parasitic capacitance due to large transistors at the opamp input as well as any switch capacitance.

Recall from Chapter 5 on negative feedback

$$\beta = \frac{1/[s(C_1 + C_p)]}{1/[s(C_1 + C_p)] + 1/(sC_2)} = \frac{C_2}{C_1 + C_p + C_2}$$

The load capacitance is more complicated. Treating the inverting terminal of OpAmp open, the effective $C_L$ is more than just $C_{\text{load}}$ and $C_C$, but

$$C_L = C_C + C_{\text{load}} + \frac{C_2(C_1 + C_p)}{C_1 + C_p + C_2}$$

This can be verified using the loop gain method introduced in Chapter 5: we can find out the loop gain first and directly find the unity-gain frequency of the loop gain:

$$V_r = -g_mV_t \frac{1}{S[\frac{C_1 + C_p}{C_1 + C_p + C_2} + C_L + C_C]} \frac{C_2}{C_1 + C_p + C_2}$$
Example 6.12 (page 280)

Consider the current-mirror opamp with no lead compensation in Example 6.11 being used in the circuit shown in Fig. 6.23 with $C_1 = C_2 = C_C = C_{\text{load}} = 5$ pF. What is the linear settling time required for 0.1 percent accuracy with?

\[
C_{\text{gs}1} = \frac{2}{3} \times 90 \times 0.4 \times 8.5 \, \text{fF/\mu m}^2 = 0.21 \, \text{pF}
\]

The capacitance seen looking into the inverting input of the opamp is one-half this value since the gate-source capacitances of the two input devices are in series. Thus, the parasitic capacitance, $C_P$, is 0.11 pF. Therefore, the effective load capacitance is given by

\[
C_L = 5 + 5 + \frac{5(5 + 0.11)}{5 + 5 + 0.11} = 12.53 \, \text{pF}
\]

\[
\omega_{ta} = \frac{K_{g_{\text{m}1}}}{C_L} = \frac{2 \times 2 \, \text{mA/V}}{12.53 \, \text{pF}} = 3.19 \times 10^8 \, \text{rad/s} \quad \beta = \frac{5}{5 + 0.11 + 5} = 0.49
\]

\[
\tau = \frac{1}{\beta \omega_{ta}} = 6.4 \, \text{ns}
\]

Finally, for 0.1 percent accuracy, we need a linear settling time of $7\tau$ or 45 ns.
6.7 Fully differential amplifiers

The main difference between single-ended amplifiers and fully-differential versions is that a current mirror load is replaced by two matched current sources in the later. Notice the power dissipation and slew rate is the same.

However, the voltage swing in fully-differential version is twice that of the single-ended version, because they use the differential voltage at two circuit nodes instead of one.

![Diagrams of single-ended and fully-differential amplifiers with equations for power, swing, and slew rate.](image)
The small signal performance of a fully differential amplifier is in many way equivalent to that of a single-ended amplifier with similar power consumption. Fig. 6.25 illustrates the point for single-stage amplifiers. Assuming that the circuit’s input stage is in both cases a differential pair under the same dc bias, and each half of the fully differential output has an output and load impedance similar to that of the single-ended output, the circuits have similar small-signal gain and bandwidth.

**dc gain:** $G_{ma}r_{out}$  

**$\omega_{-3dB}$** $= 1/(r_{out}C_L)$  

**$\omega_{ta}$** $= G_{ma}/C_L$

---

**dc gain:** $\frac{G_{ma}}{2}r_{out} = G_{ma}r_{out}$  

**$\omega_{-3dB}$** $= 1/(r_{out}C_L)$  

**$\omega_{ta}$** $= G_{ma}/C_L$

---

*Chapter 6 Figure 25*
Why Fully differential amplifiers?

One of the main driving forces behind the use of fully differential amplifiers is to help reject common-mode noise. The common-mode noise, $n_{cm}$, appears identically on both half signals and is therefore cancelled when the difference between them is taken.

Many noise sources, such as power supply noise, bias voltage noise and switches noise act as common mode noise and can therefore be well rejected in fully-differential amplifiers.

$\text{Ni}_1$ and $\text{Ni}_2$ in the figure represent random noise sources added to the two outputs, and the overall signal-to-noise ratio is still better than the single-ended version.
Why Fully differential amplifiers?

Fully-differential amplifiers have another benefit that if each output is distorted symmetrically around the common-mode voltage, the differential signal will have only odd-order distortion, which are often much smaller.

With the above mentioned advantages, most modern analog circuits are realized using fully differential structures.

One major drawback of using fully-differential OpAmp is that common-mode feedback circuit (CMFB to be discussed later) must be added to establish the common-mode output voltage. Another minor overhead is that in practice fully-differential OpAmp may need some additional power consumption due to CMFB and to produce the two outputs.

\[ v_{p1} = k_1 v_1 + k_2 v_1^2 + k_3 v_1^3 + k_4 v_1^4 + \cdots \]

\[ v_{n1} = -k_1 v_1 + k_2 v_1^2 - k_3 v_1^3 + k_4 v_1^4 + \cdots \]

\[ v_{\text{diff}} = 2k_1 v_1 + 2k_3 v_1^3 + 2k_5 v_1^5 + \cdots \]
6.7.1 Fully differential folded-cascode OpAmp

Compared to the singled-ended version, the n-channel current mirror has been replaced by two cascode current sources of Q7/8 and Q9/10. Also, a CMFB circuit is introduced. The gate voltage $V_{\text{cntrl}}$ is the output of the CMFB. Note that when OpAmp is slewing, the maximum current for negative slew rate is limited by the bias current of Q7 or Q9 (as there is no current mirror like the singed-ended one). So, fully-differential is usually designed with bias current in the output stage equal to the bias currents in the input transistors.

Note that each signal path now consists of only one node in addition to the output nodes, which is the drain nodes of Q1/2. These nodes are responsible for the second pole.

When load capacitance is relatively small so it is important to push the second pole away, then one can consider using pMOS for Q1/2 and nMOS for Q5/6, as the impedance at the drain of Q1/2 would be larger that way, resulting in smaller time constants. However, the tradeoff is DC gain may be smaller.
6.7.2 Alternative fully differential OpAmps

The previous singled-ended current mirror OpAmp can be converted to a fully-differential one as below.

Similarly, the complementary design using pMOS at input stage is possible. Which one to use depends on whether the load capacitance or second pole are limiting the bandwidth and whether DC gain or bandwidth is more important. (in the former case, then nMOS input is preferred.)

For a general-purpose amplifier, this design with large pMOS transistors, a current gain of \( K=2 \) and wide-wing enhanced output-impedance cascode mirrors and current sources may be a good choice compared to other designs.

![Circuit Diagram](image)
One limitation for fully differential OpAmp seen so far is that the maximum current at the output for singled-ended slewing is limited by fixed current sources. It is possible to modify the design to get bi-directional drive capability at the output.

In the revised circuit, the current mirrors at the top have been replaced by current mirrors having two outputs. The first output has a gain of K and goes to the output of the OpAmp as before. The second output has a gain of one and goes to a new current mirror that has a current gain of K, where it is mirrored the second time and then goes to the opposite output.

In this OpAmp, when slewing (suppose a very large input voltage), then the current going to Vout+ is Kibias, whereas the current sunked from Vout- is also Kibias.

This OpAmp has an improved slew rate at the expense of slower small-signal response due to addition of extra current mirrors. But it may be worthwhile in some applications.
Another alternative design to have bi-direction driving capability is to use two singled-ended output OpAmps with their inputs connected in parallel and each of their output being one output side of the fully-differential version.

The disadvantage is the additional current mirrors and complexity. (note in the figure, the CMFB loop is not shown).
6.7.3 Low supply voltage OpAmps

Low supply voltage complicates the OpAmp design. For the folded-cascode OpAmp, the input common-mode voltage must be larger than $V_{gs1} + V_{eff}$ in order to keep the tail current source device in active mode (a typical value is 0.95V which is difficult for 1.2 power supply).

The low-voltage design shown below (CMFB circuit is not shown) makes use of both nMOS and pMOS in the two differential input pairs. When the input common-mode voltage range is close to one of the power supply voltages, one of input differential pairs turns off while the other one remains active. To keep the OpAmp gain relatively constant, the bias currents of the still-active pair is dynamically increased.

For example, when input common-mode voltage is close to $V_{dd}$, $Q_3/4$ turns off and $Q_6$ conduct all of $I_2$ so that the bias current of $I_1$ is increased.
Another challenge of low supply voltage designs is that the signal output swing is very small, especially for the single-stage folded-cascode OpAmp (referring to the OpAmp in Fig 6.28, it can be shown that the signal swing is as small as 0.3V if $V_{\text{eff}}=0.2V$).

One possible design to alleviate that problem is an enhanced two-stage OpAmp with a folded-cascode first stage and a common-source second stage. The first stage can provide high gain (small voltage swing for first stage output is not an issue) and the second stage can provide relatively large signal swing. (note CMFB is not shown and also the lead compensation is used.)
6.8 Common-mode feedback circuits

To understand the need for CMFB, let us begin with a simple realization of a differential amplifier [Fig. 9.30(a)]. In some applications, we short the inputs and outputs for part of the operation [Fig. 9.30(b)], providing differential negative feedback. The input and output common-mode levels in this case are quite well-defined, equal to \( V_{DD} - I_{SS} R_D/2 \).

![Diagram of differential amplifier with CMFB](image)

**Figure 9.30** (a) Simple differential pair. (b) Circuit with inputs shorted to outputs.

6.8 Common-mode feedback circuits

To understand the need for CMFB, let us begin with a simple realization of a differential amplifier [Fig. 9.30(a)]. In some applications, we short the inputs and outputs for part of the operation [Fig. 9.30(b)], providing differential negative feedback. The input and output common-mode levels in this case are quite well-defined, equal to $V_{DD} - I_{SS} R_D / 2$.

![Diagram of CMFB circuits](image)

**Figure 9.30** (a) Simple differential pair. (b) circuit with inputs shorted to outputs.

6.8 Common-mode feedback circuits

Now suppose the load resistors are replaced by PMOS current sources so as to increase the differential voltage gain [Fig. 9.31(a)]. What is the common-mode level at nodes X and Y? Since each of the input transistors carries a current of \( I_{SS}/2 \), the CM level depends on how close \( I_{D3} \) and \( I_{D4} \) are to this value. In practice, as exemplified by Fig. 9.31(b), mismatches in the PMOS and NMOS current mirrors defining \( I_{SS} \) and \( I_{D3,4} \) create a finite error between \( I_{D3,4} \) and \( I_{SS}/2 \). Suppose, for example, that the drain currents of \( M_3 \) and \( M_4 \) in the saturation region are slightly greater than \( I_{SS}/2 \). As a result, to satisfy Kirchhoff current law at nodes X and Y, both \( M_3 \) and \( M_4 \) must enter the triode region so that their drain currents fall to \( I_{SS}/2 \). Conversely, if \( I_{D3,4} < I_{SS}/2 \), then both \( V_X \) and \( V_Y \) must drop so that \( M_5 \) enters the triode region, thereby producing only \( 2I_{D3,4} \).

Figure 9.31 (a) High-gain differential pair with inputs shorted to outputs, (b) effect of current mismatches.
6.8 Common-mode feedback circuits

The above difficulties fundamentally arise because in high-gain amplifiers, we wish a p-type current source to balance an n-type current source. As illustrated in Fig. 9.32, the difference between $I_p$ and $I_N$ must flow through the intrinsic output impedance of the amplifier, creating an output voltage change of $(I_p - I_N)(R_p \parallel R_N)$. Since the current error depends on mismatches and $R_p \parallel R_N$ is quite high, the voltage error may be large, thus driving the p-type or n-type current source into the triode region. As a general rule, if the output CM level cannot be determined by “visual inspection” and requires calculations based on device properties, then it is poorly defined. This is the case in Fig. 9.31 but not in Fig. 9.30. We emphasize that differential feedback cannot define the CM level.

The foregoing study implies that in high-gain amplifiers, the output CM level is quite sensitive to device properties and mismatches and it cannot be stabilized by means of differential feedback. Thus, a common-mode feedback network must be added to sense the CM level of the two outputs and accordingly adjust one of the bias currents in the amplifier.

![Figure 9.32](image-url) Simplified model of high-gain amplifier.
6.8 Common-mode feedback circuits

Typically, a CMFB circuit should have three operations:
1. Sense the common-mode voltage level of the differential output;
2. Compare the common-mode voltage to a reference voltage (the desired voltage);
3. Return the error to the amplifier’s bias network to adjust the current and eventually the output voltage.

The following circuit illustrates the idea.

CMFB circuit design may well be the most difficult part of the OpAmp design.
6.8 Common-mode feedback circuits

There are two typical approaches to designing CMFB circuits—a continuous-time approach and a switched-capacitor approach. The former approach is often the limiting factor on maximizing the signal swings, and, if nonlinear, may actually introduce common-mode signals. The latter approach is typically only used in switched-capacitor circuits, since in continuous-time applications it introduces clock-feedthrough glitches.

The one shown below is a continuous one. To illustrate, assume CM output voltage, $V_{out,CM}$, equal to reference voltage $V_{ref,CM}$, and that $V_{out+}$ is equal in magnitude but opposite in sign to $V_{out-}$. Also, assume the two differential pairs Q1/2 and Q3/4 have infinite CMRR (i.e. output of them depend only on their differential voltage).

Since two pairs have the same differential voltages, current in Q1 is equal to current in Q3 and that in Q2 equal to Q4. Denoting the current in Q2 as $I_{D2} = I_{B}/2 + \Delta I$, and current in Q3 is $I_{D3} = (I_{B}/2) - \Delta I$, and the current in Q5 is $I_{D5} = I_{D2} + I_{D3} = (I_{B}/2 + \Delta I) + ((I_{B}/2) - \Delta I) = I_{B}$.

In the nominal case, when $V_{out,CM} = V_{ref,CM}$, then there will no voltage change for $V_{cntrl}$ stays constant.

If $V_{out,CM} > V_{ref,CM}$, then the differential voltage across Q1/2 increases while that for Q3/4 decreases, so the current in Q2 and Q3 will be larger than before, which increases the voltage $V_{cntrl}$. 
6.8 Common-mode feedback circuits

Now this voltage $V_{\text{cntrl}}$ can be the bias voltage that sets the current levels in the nMOS current sources at the output of the OpAmp (see below), which will bring down the common-mode output voltage, $V_{\text{out,CM}}$ to decrease toward the nominal $V_{\text{ref,CM}}$.

So, as long as the common-mode loop gain is large enough, and the differential signals are not so large as to cause either differential pair Q1/2 or Q3/4 to turn off, $V_{\text{out,CM}}$ can be kept very close to $V_{\text{ref,CM}}$. The later requires that we maximize the $V_{\text{eff}}$ for these transistors.

Finally, the IB should be high output impedance cascode current sources to ensure good CMRR.
Another CMFB

This circuit generates the senses the common-mode voltage of the output signals (minus a DC level shift) at node VA. This voltage is then compared to a reference voltage, Vref, using a separate amplifier.

One limitation is that the voltage drop across Q1/2 may severely limits the differential signals that can be processed, which is important in lower supply voltage applications.

\[ V_A = V_{CM} - (V_{off1} + V_{t1}) \]

\[ V_{ref} = -(V_{off1} + V_{t1}) \]
Design considerations of CMFB loop

One important design consideration is that CMFB is part of the negative feedback loop, and therefore must be well compensated if needed, otherwise the injection of common-mode signal can cause output ringing and even unstable. Thus, phase margin (break at Vcntrl to find loop gain from Chapter 5) and step response (giving Vref a step input) of the common-mode loop should be checked.

Often, the common-mode loop is stabilized using the same capacitors used to compensate the differential loop (for example by connecting two comp. or load capacitors from outputs to ground).

Also, it is important to maximize the speed of CMFB loop by having as few nodes in the design as possible (to prevent high frequency common-mode noise). For this reason, the CFMB output is usually used to control current sources in the output stage of the OpAmp. For the same reason, the CM output of each stage in a multi-stage amplifier is individually compensated (for example the on in Fig 6.33. ).

This is an active research area.
Switch-capacitor CMFB circuit

In this approach, Capacitors $C_c$ generates the $V_{out,CM}$, which is then used to create control voltages $V_{cntrl}$. The bias voltage $V_{bias}$ is designed to be equal to the difference between the desired $V_{ref,CM}$ and the desired $V_{cntrl}$ used for OpAmp current sources.

This CMFB circuit is mostly used in switched-capacitor circuits since they allow a larger output signal swing.