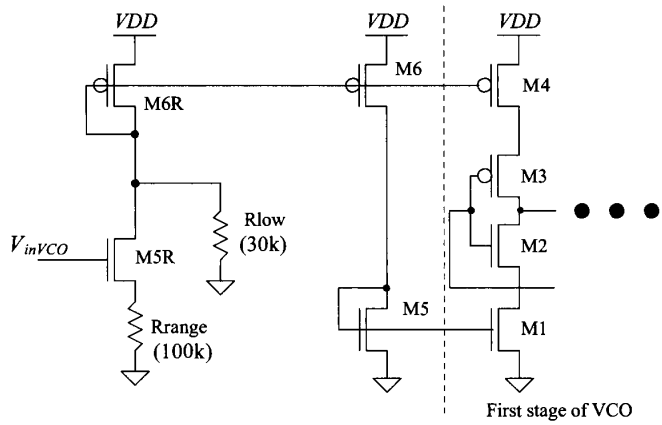


Figure 19.25 Limiting the current in a current-starved VCO.

to calculate the pull-in range since it is greater than the lock-in range. Again, the design of the oscillator is important in this DPLL. The lock time, T_L, of the DPLL is 40 ns. SPICE simulation results of this DPLL are seen in Fig. 19.27. The top waveform is the input control voltage of the VCO. The time that it takes the loop to pull-in the frequency and lock, from this figure, is roughly 250 ns. The lock time would be the time it takes the loop if the input frequency jumped after the loop was locked. Note how, in the bottom traces, the output clock rising edge is centered on the data. Notice that the VCO's control voltage varies by 200 mV. From Fig. 19.26, we can see a 5 MHz variation, Δf_{VCO}, in the VCO's output frequency. To estimate the jitter, Δt_{jitter}, in the output, we can use

$$\Delta t_{jitter} = \frac{1}{f_{center}} - \frac{1}{f_{center} + \Delta f_{VCO}} = \frac{1}{100\ MHz} - \frac{1}{105\ MHz} = 480\ ps \qquad (19.41)$$

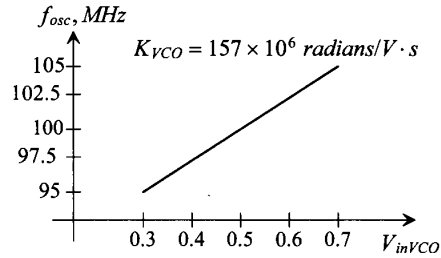


Figure 19.26 Gain of the VCO in Fig. 19.16 after linearizing and limiting with Fig. 19.25.

IPR2023-00698
Theta EX2019

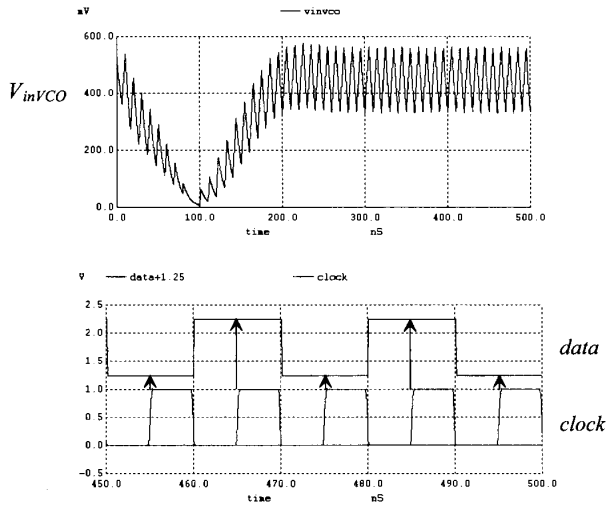


Figure 19.27 Simulating the DPLL in Ex. 19.2.

In this simulation (Fig. 19.27), we used the ideal input data, that is, a string of alternating ones and zeroes. Figure 19.28 shows what happens when the data is a one followed by seven zeroes. The rising edge of the clock is occurring a little offset from the center of the data. This error is sometimes called a *static phase error*. As mentioned earlier, when the input data is not changing, the VCO control voltage starts to wander back towards *VDD*/2.

Finally, *it's very important* to realize that if we were to increase the *RC* time constant of the filter, as seen above, the damping factor, ζ , would decrease and the loop would never lock. This is counter-intuitive and the reason we should *always use the design equations when designing DPLLs.* ■

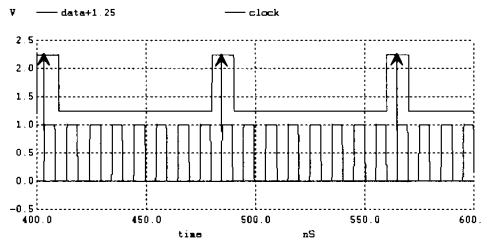
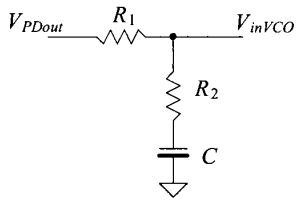


Figure 19.28 Showing how the DPLL doesn't lock up to the center of the data when the data isn't alternating ones and zeroes.

IPR2023-00698
Theta EX2019

Adding a zero to the simple passive RC loop filter, Fig. 19.29 (called a passive lag loop filter), the loop-filter pole can be made small (and thus the gain of the VCO can be made larger) while at the same time a reasonable damping factor can be achieved. The result is an increase in the lock range of a DPLL using the XOR PD and a shorter lock time see Eqs. (19.37) – (19.40). The passive lag loop filter is, in most situations preferred over the simple RC. Again, as Ex. 19.2 showed, if the center frequency of the VCO doesn't match the input frequency, the clock will not align at $\pi/2$ (in the center of the data).



$$K_F = \frac{V_{inVCO}}{V_{PDout}} = \frac{1 + j\omega R_2 C}{1 + j\omega(R_1 + R_2)C}$$

$$\omega_n = \sqrt{\frac{K_{PD}K_{VCO}}{N(R_1 + R_2)C}} \qquad \Delta\omega_L = \pi\zeta\omega_n$$

$$\zeta = \frac{\omega_n}{2} \cdot \left(R_2 C + \frac{N}{K_{PD}K_{VCO}}\right)$$

Figure 19.29 Passive lag loop filter used to increase DPLL lock range.

Active-PI Loop Filter

The clock misalignment encountered in a DPLL using an XOR PD and passive loop filter can be minimized by using the active proportional + integral (active-PI) loop filter shown in Fig. 19.30. The integrator allows the V_{inVCO} voltage to move, and stay, away from *VDD*/2. This then eliminates the static phase error (the clock will align to the middle of the data). The transfer function of this filter is given by

$$K_F = \frac{1 + sR_2C}{sR_1C} = \underbrace{\frac{R_2}{R_1}}_{\text{proportional}} + \underbrace{\frac{1}{sR_1C}}_{\text{integral}} \tag{19.42}$$

The natural frequency of the resulting second-order system is given by

$$\omega_n = \sqrt{\frac{K_{PD}K_{VCO}}{NR_1C}} \tag{19.43}$$

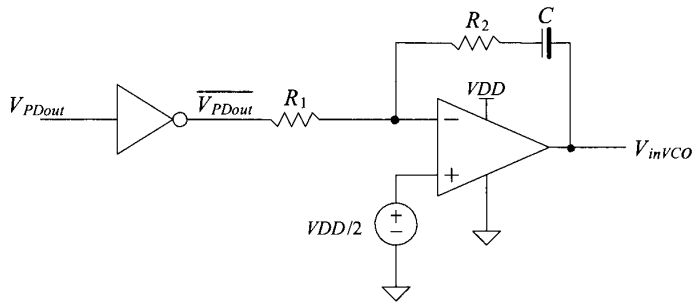


Figure 19.30 Active-PI loop filter.

IPR2023-00698
Theta EX2019

and the damping ratio is given by

$$\zeta = \frac{\omega_n R_2 C}{2} \tag{19.44}$$

The lock range is

$$\Delta\omega_L = 4\pi\zeta\omega_n \tag{19.45}$$

while the lock time remains $2\pi/\omega_n$. The pull-in range, using the active-PI loop filter, is limited by the VCO oscillator frequency. Consider the following example.

Example 19.3

Repeat Ex. 19.2 using the active-PI loop filter. The SPICE model for an op-amp (a voltage-controlled voltage source) seen in Fig. 20.19 can be used for the op-amp.

Using Eq. (19.43) in (19.44) gives

$$\zeta = \frac{R_2 C}{2}\sqrt{\frac{K_{PD}K_{VCO}}{NR_1 C}}$$

If we set the damping factor again to 1 we can write

$$4 = \frac{157 \times 10^6 R_2^2 C}{2\pi R_1}$$

Setting C to 10 pF and R_2 to 25k, then R_1 is 39k. The simulation results are seen in Fig. 19.31. The output clock is aligned to the center of the data. It's important to make sure that the step size used in the SPICE transient simulation isn't too big. For example, if a 100 MHz clock is used with a step size of 1 ns, it is likely that the loop either won't lock or if it does lock, both the VCO control voltage and thus the output data will oscillate. A maximum step size for 100 MHz signals may be 100 ps. Also note the sparse data pattern will increase the lock time (thus the reason for the long simulation time in Fig. 19.31).

Finally, it is important to remember that the VCO tuning scheme in Fig. 19.25 is not practical unless the components can be set manually (either off-chip or using fuses or switches on chip). Also, something very important that we are not discussing (yet) is power supply noise. The current-starved VCO's output frequency is not very stable with changes in *VDD*. ■

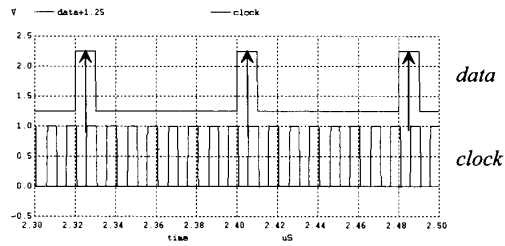


Figure 19.31 How the loop locks up on the center of the data.

IPR2023-00698
Theta EX2019

19.3.2 PFD DPLL

Tri-State Output

A block diagram of a DPLL using the PFD with tri-state output is shown in Fig. 19.32. The phase transfer function is the same as Eq. (19.30)

$$H(s) = \frac{\phi_{clock}}{\phi_{data}} = \frac{K_{PDtri}K_F K_{VCO}}{s + \beta \cdot K_{PDtri}K_F K_{VCO}} \tag{19.46}$$

where

$$K_F = \frac{1 + sR_2C}{1 + s(R_1 + R_2)C} \tag{19.47}$$

When this filter is driven with the tri-state output, no current flows in R_1 or R_2 with the output in the high impedance state. The voltage across the capacitor remains unchanged. We can think of the filter, tri-state output as an ideal integrator with a transfer function

$$K_F' = \frac{1 + sR_2C}{s(R_1 + R_2)C} \tag{19.48}$$

Substituting this equation into Eq. 19.46 and rearranging results in

$$H(s) = \frac{K_{PDtri}K_{VCO}\frac{1+sR_2C}{(R_1+R_2)C}}{s^2 + s\frac{K_{PDtri}K_{VCO}R_2C}{N(R_1+R_2)C} + \frac{K_{PDtri}K_{VCO}}{N(R_1+R_2)C}} = \frac{\phi_{clock}}{\phi_{data}} = \frac{f_{clock}}{f_{data}} \tag{19.49}$$

From this equation, the natural frequency is

$$\omega_n = \sqrt{\frac{K_{PDtri}K_{VCO}}{N(R_1 + R_2)C}} \tag{19.50}$$

and the damping factor is determined by solving

$$2\zeta\omega_n = \frac{K_{PDtri}K_{VCO}R_2C}{N(R_1 + R_2)C} \tag{19.51}$$

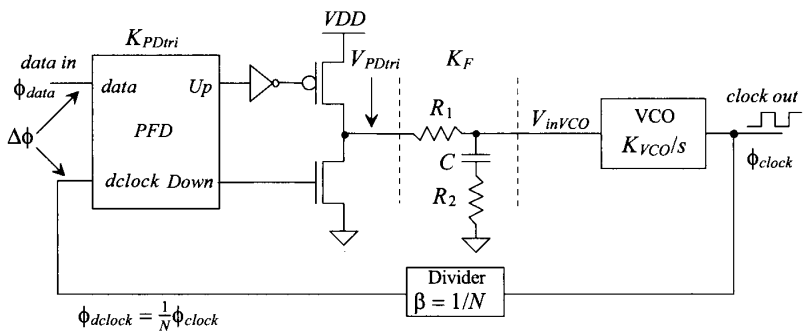


Figure 19.32 Block diagram of a DPLL using a sequential phase detector (PFD).

IPR2023-00698
Theta EX2019

which results in a damping factor given by

$$\zeta = \frac{\omega_n}{2} \cdot R_2 C \tag{19.52}$$

The lock range is given by

$$\Delta\omega_L = 4\pi\zeta\omega_n \tag{19.53}$$

while the lock time, T_L, remains $2\pi/\omega_n$. The pull-in range is limited by the VCO operating frequency. The pull-in time is given by

$$T_P = 2R_1 C \cdot \ln \frac{(K_{VCO}/N) \cdot (VDD/2)}{(K_{VCO}/N)(VDD/2) - \Delta\omega} \tag{19.54}$$

where $\Delta\omega$ is the magnitude of the *input* frequency step.

Implementing the PFD in CMOS

The PFD seen in Fig. 19.10 can be implemented using inverters and nand gates as seen in Fig. 19.33. Simulating this circuit gives the results seen in Fig. 19.34. When the *dclock* is lagging, the *data* the output of the PFD is the *up* pulse (indicating that the edge of *dclock* needs to speed up or occur earlier in time, that is, the VCO's control voltage should increase). When *data* is lagging *dclock* the *down* pulse goes high indicating *dclock* should slow down.

We have some fine points that need to be discussed here. For example, what happens to *up* and *down* as the rising edges of *dclock* and *data* move close together? What we may get is some small glitches in the *up* and *down* signals. Or we may see both *up* and *down* staying low as the two pulses move closer together. In the PFD of Fig. 19.33, the delay through the two inverters can be used to set how *up* and *down* behave as

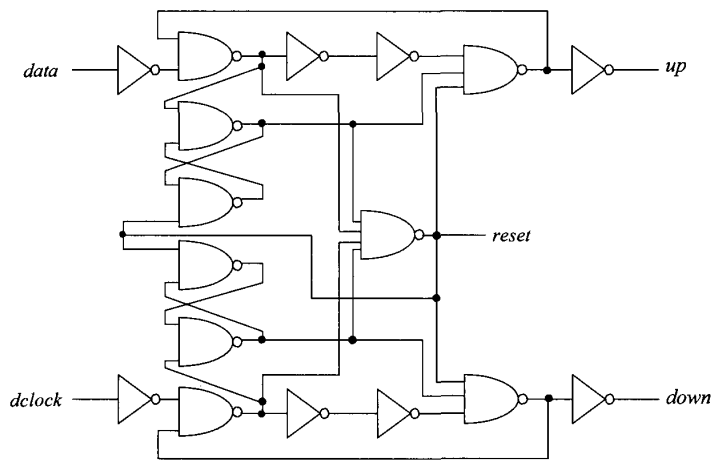


Figure 19.33 CMOS implementation of a PFD.

IPR2023-00698
Theta EX2019

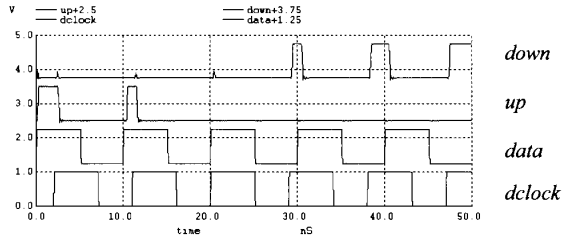


Figure 19.34 Simulation results for the PFD in Fig. 19.33.

the PFD's inputs move close together. If, for example, both stay low when the PFD's inputs get within 100 ps of each other, we get a *static phase error*. The loop won't lock any tighter than within 100 ps of its ideal position. Another way of looking at this is that as the phase difference, $\Delta\phi$, moves towards zero (see Fig. 19.11), the gain (the slope of the curve in Fig. 19.11) decreases. The phase detector is then said to have a *dead zone* where the gain of the phase detector (the slope of the curves) decreases.

Example 19.4

Design a DPLL using the tri-state topology seen in Fig. 19.32 that generates a clock signal at a frequency of 100 MHz from a 50 MHz square wave input. This application of the DPLL is called *frequency synthesis*.

The feedback path contains a divide by 2 circuit ($N = 2$). Let's use the VCO with the characteristics seen in Fig. 19.18

$$K_{VCO} = 1.57 \times 10^9 \; radians/V \cdot s$$

The gain of the phase detector (knowing *VDD* is 1 V) is

$$K_{PDtri} = \frac{1}{4\pi}$$

The lock range, Δf_L will be set to 20 MHz. Again let's set $\zeta = 1$. Using Eq. (19.53)

$$\Delta\omega_L = 4\pi\zeta\omega_n \rightarrow \omega_n = 10 \times 10^6 \; radians/V \cdot s$$

Using Eq. (19.52) gives

$$R_2C = 200 \; ns$$

Let's set the capacitor to 10 pF and R_2 to 20k. Solving Eq. (19.50) for R_1 gives 42.5k ($= R_1$).

The simulation results are seen in Fig. 19.35. We should first notice that the VCO's control voltage doesn't have the excessive ripple like we had in the DPLL using the XOR gate phase detector. The response of the loop shows a nice $\zeta = 1$ shape. Again note that a DPLL loop's pull-in range is limited by the VCO's operating frequency (which, in this example uses the VCO from Fig. 19.18, which ranges from 50 to 150 MHz). A good exercise to perform at this point is to change the divider in the feedback path (to divide by 1, 2, 4, etc.) and the input frequency (a signal we've called *data*) and look at the robustness of the loop. ■

IPR2023-00698
Theta EX2019

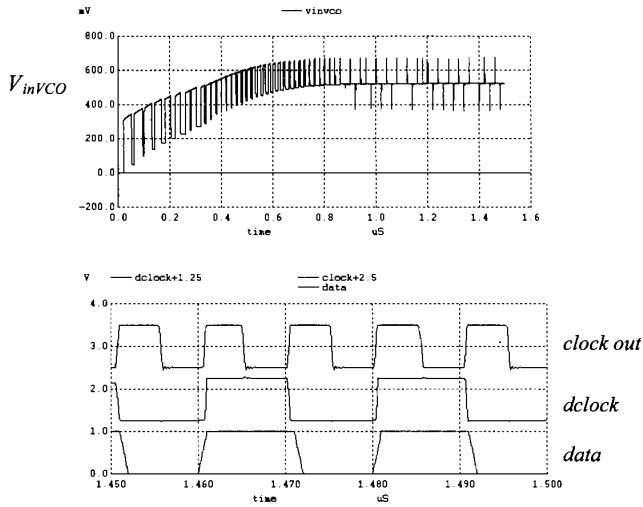


Figure 19.35 Simulation results for Ex. 19.4.

PFD with a Charge Pump Output

The PFD with a charge-pump output is seen in Fig. 19.36. A CMOS implementation of a DPLL using this configuration is, in general, preferred over the tri-state output because of the better immunity to power supply variations. In the tristate configuration seen in Fig. 19.32 note that when either the NMOS or PMOS switches are on, either *VDD* or ground are connected directly to the loop filter. Power supply or ground noise can thus feed directly into the loop filter and then into the VCO's control voltage.

The loop filter integrates the charge supplied by the charge pump. The capacitor C_2 prevents $I_{pump} \cdot R$ from causing voltage jumps on the input of the VCO and thus frequency jumps in the DPLL output. In general, C_2 is set at about one-tenth (or less) of C_1. The loop-filter transfer function neglecting C_2 is given by

$$K_F = \frac{1 + sRC_1}{sC_1} \tag{19.55}$$

The feedback loop transfer function is given by

$$H(s) = \frac{\phi_{clock}}{\phi_{data}} = \frac{K_{PDI}K_{VCO}(1 + sRC_1)}{s^2 + s(\frac{K_{PDI}K_{VCO} \cdot R}{N}) + \frac{K_{PDI}K_{VCO}}{NC_1}} \tag{19.56}$$

From the transfer function the natural frequency is given by

$$\omega_n = \sqrt{\frac{K_{PDI}K_{VCO}}{NC_1}} \tag{19.57}$$

and the damping factor is

$$\zeta = \frac{\omega_n}{2} \cdot RC_1 \tag{19.58}$$

IPR2023-00698
Theta EX2019

The lock range and lock time remain the same (using the different values for the natural frequency and damping ratio) as the PFD with the tri-state output. Again, the pull-in range is set by the VCO oscillator frequency range. The pull-in time is given by

$$T_P = 2RC_1 \ln\left[\frac{(K_{VCO}/N)\cdot(I_{pump})}{(K_{VCO}/N)\cdot(I_{pump}) - \Delta\omega}\right] \tag{19.59}$$

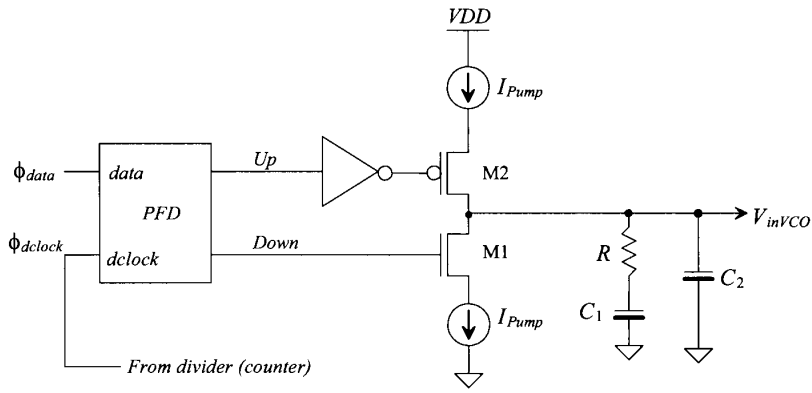


Figure 19.36 PFD using the charge pump.

Practical Implementation of the Charge Pump

The circuit seen in Fig. 19.36 is useful for illustrating the concept of the PFD with charge pump. However, in a practical circuit the fact that the sources of M1 and M2 charge to ground and *VDD* respectively when M1 and M2 are off creates some design issues. For example, suppose that the source of M1 is discharged to ground when the *down* signal goes high. When M1 turns on, it doesn't supply the charge set by I_{pump} to the loop filter but rather, until the voltage across the current sink increases, behaves like a switch simply connecting the filter's input to ground (meaning that we are not controlling the signal that is applied to the loop filter). To get around this problem, consider the circuit seen in Fig. 19.37. The bias voltages come from diode-connected MOSFETs (to form current mirrors, see Ch. 20). When *up* and *down* are low, M1L and M2L are on. The pump-up current source, MPup, is driving the pump-down current source, MNdwn. When either *up* or *down* goes high, the output is connected to one of the current sources. The MOSFETs M1L,R and M2L,R simply steer the current to the loop filter. The only other concern we have with this topology is the fact that when M1L and M2L are both on, the voltage on their drains won't precisely match the voltage across the loop filter (V_{inVCO}). In other words, the voltage across each of the current sources (MPup and MNdwn) won't be the same as it is when they are connected to the output node. The result is that charge sharing between the parasitic capacitance on the drains of MPup and MNdwn and the capacitance used in the loop filter cause a static phase error or jitter. To eliminate this problem, an amplifier (see dashed lines in the figure) can be inserted to set the drain voltage of M1L and M2L to the same value V_{inVCO}.

IPR2023-00698
Theta EX2019

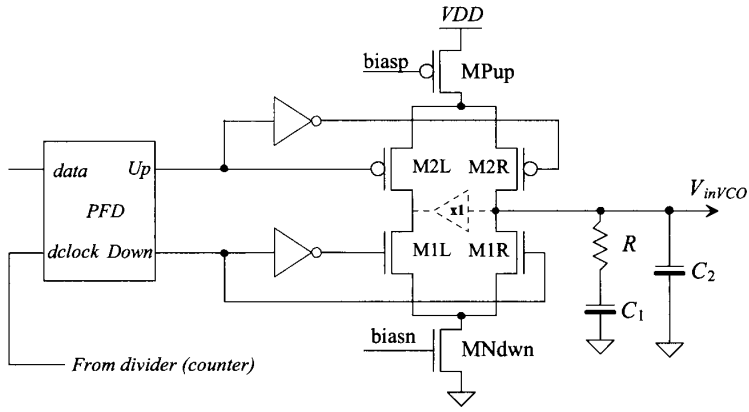


Figure 19.37 Practical implementation of the charge pump.

Example 19.5

Repeat Ex. 19.4 using the PFD and charge pump seen in Fig. 19.37.

One of the other benefits of using the charge pump is the fact that we can select I_{pump}, that is,

$$K_{PDI} = \frac{I_{pump}}{2\pi}$$

Again, let's set the lock range to 20 MHz. Using Eq. (19.53) with $\zeta = 1$ gives, again, $\omega_n = 10 \times 10^6$ *radians/V·s*. From Eq. (19.58) we get

$$RC_1 = 200\ ns \text{ so let's use } R = 20\text{k}, C_1 = 10\text{ pF}, \text{ and } C_2 = 1\text{ pF}$$

remembering that we generally set C_2 to one-tenth of C_1. Using Eq. (19.57), we can now find the value of I_{pump}

$$10 \times 10^6 = \sqrt{\frac{I_{pump} \cdot 1.57 \times 10^9}{2\pi \cdot 2 \cdot 10 \times 10^{-12}}} \rightarrow I_{pump} = 8\ \mu A$$

Because this value isn't that critical, we'll round it up to 10 μA. Figure 19.38 shows the simulation results. Notice the $\zeta = 1$ (or maybe a little less because the voltage does overshoot its final value by a little bit) shape of the VCO's control voltage. Let's modify the loop filter to show what V_{inVCO} would look like if $\zeta = 0.1$. All we have to do, from Eq. (19.58), is drop R by a factor of 10. The result is seen in Fig. 19.39. The control voltage is oscillating and the loop isn't locking. Of course, we can also increase the damping factor. The loop now behaves sluggishly and does not respond to fast changes in the input signal. For general design, where the process and temperature vary, it's better to center the design on a larger damping factor to avoid instability. ■

IPR2023-00698
Theta EX2019

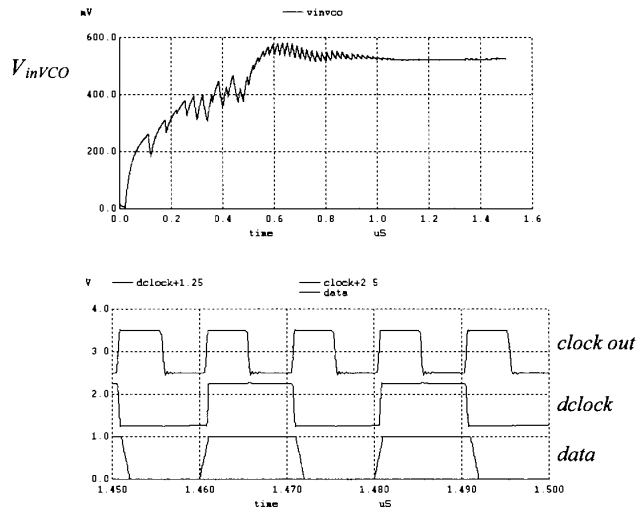


Figure 19.38 Simulation results for Ex. 19.5.

Discussion

When selecting values for the loop filter, we assumed that the output resistance of the phase detector was small (for the XOR and tri-state PD) compared to the impedances used in the loop filter. We also assumed that the input resistance of the VCO was infinite and the input capacitance of the VCO was small compared to the capacitance used in the loop filter. Considering the parasitics present in the DPLL is an important part of the design.

The center frequency of the VCO is critical for good DPLL performance when using the XOR gate with *RC* loop filter. If the center frequency, f_{center}, of the VCO (i.e., $V_{inVCO} = VDD/2$) does not match twice the input data rate, the DPLL will lock up at a phase different from $\pi/2$ (the input frequencies to any phase detector must be equal). The need for a precision center frequency is eliminated by using the XOR PD with an active-PI loop filter or by using the PFD. The other big benefit of using the PFD is that the VCO's gain can be larger.

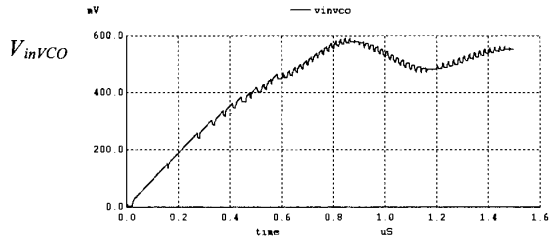


Figure 19.39 Showing what happens when the damping factor is reduced to 0.1.

IPR2023-00698
Theta EX2019

Finally, selecting of the loop's damping factor, ζ, is very important. If the value of ζ is too small, the loop will have trouble locking or the output will jitter excessively when the loop is locked. This last problem is sometimes (gratuitously) called *jitter peaking* since a step in the DPLL's input frequency causes excessive overshoot in V_{inVCO} with small ζ.

19.4 System Considerations

System concerns are often the driving force behind the design of a DPLL. Referring to Fig. 19.1, we observe that the data transmitted through the channel should ideally arrive at the receiver with the same shape with which it was transmitted. In reality, the data becomes distorted. Distortion arises from nonlinearities in the receiver input amplifier and the finite bandwidth of the channel. To understand the conditions for distortionless transmission, consider the block diagram shown in Fig. 19.40. The system has a transfer function in the frequency domain of $H(f)$ and in the time domain $h(t)$. For distortionless transmission, we can relate the input and output of the system by

$$y(t) = K \cdot x(t - t_o) \tag{19.60}$$

where t_o is the time delay through the system and K is a constant. This equation shows that for distortionless transmission through a system the output is simply a scaled, time-delayed version of the input. An interesting observation can be made by taking the Fourier Transform of both sides of this equation,

$$Y(f) = K \cdot X(f)e^{-j2\pi f t_o} \tag{19.61}$$

The transfer function of a distortionless system can then be written as

$$H(f) = \frac{Y(f)}{X(f)} = Ke^{-j2\pi f t_o} \tag{19.62}$$

Figure 19.41 shows the magnitude and phase responses of a distortionless system. A system is distortionless when its amplitude response, $|H(f)|$, is a constant, K, and its phase response, $\angle H(f)$, is linear with a slope of $-2\pi t_0$ over all frequencies of interest.

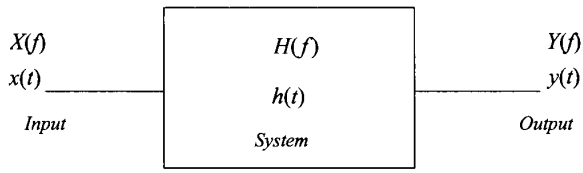


Figure 19.40 Representation of a system with input and output.

The responses shown in Fig. 19.41 are ideal. In practice, the magnitude response of a system may look similar to Fig. 19.42a. At higher frequencies, the magnitude rolls off. To compensate for this roll-off, or other imperfections, a circuit called an equalizer is added in series with the system (Fig. 19.42b). The equalizer has a transfer function in which its magnitude response increases with increasing frequency beyond a point (Fig. 19.42c). If the low-frequency gain of the equalizer is A/K and the low-frequency gain of the system is K, then the resulting gain of the system/equalizer combination is A.

IPR2023-00698
Theta EX2019

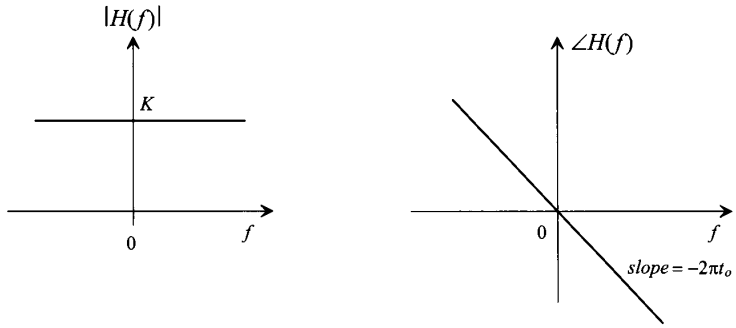


Figure 19.41 Magnitude and phase response of a distortionless system.

Another source of potential distortion occurs when the receiver input data is regenerated into digital levels. This was discussed back in Sec. 18.3. Timing errors occur when the input data is not precisely sliced through its middle (see Fig. 18.12). What makes slicing the data correctly even more difficult is the fact that the amplitude response of the channel can change with time and the data pattern can affect the average level of the data. There are two solutions to this problem. The first uses a circuit (see Fig. 18.29) that determines the peak positive and negative input analog amplitudes, averages the values, and feeds back the result to the comparator in the decision-making circuit. The second method encodes the digital data so that the duty cycle of the resulting encoded data is 50%. The encoding increases the channel bandwidth for a constant data rate.

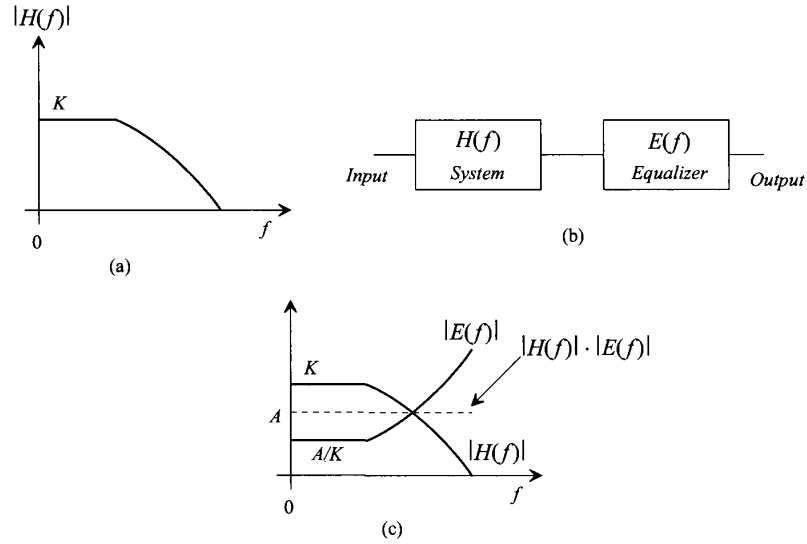


Figure 19.42 Using an equalizer to lower distortion in a system.

IPR2023-00698
Theta EX2019

Encoding the data can eliminate the need for a decision circuit. If the resulting encoded data has a 50% duty cycle, it can be passed through a capacitor in the receiver, resulting in an analog signal centered around ground. The noninverting input of the comparator will then be connected to ground. The comparator will then be able to slice the data at the correct moments in time (in the middle of the data bit). An example of an encoding scheme is shown in Fig. 19.43. Encoding occurs in the transmitter prior to transmission over the channel. This particular encoding scheme is referred to as the bi-phase format, or more precisely, the bi-phase-level (sometimes called bi-phase-L or Manchester NRZ) format. The cost of using this scheme over the NRZ data format is increased channel bandwidth. Other encoding schemes are shown in Fig. 19.44.

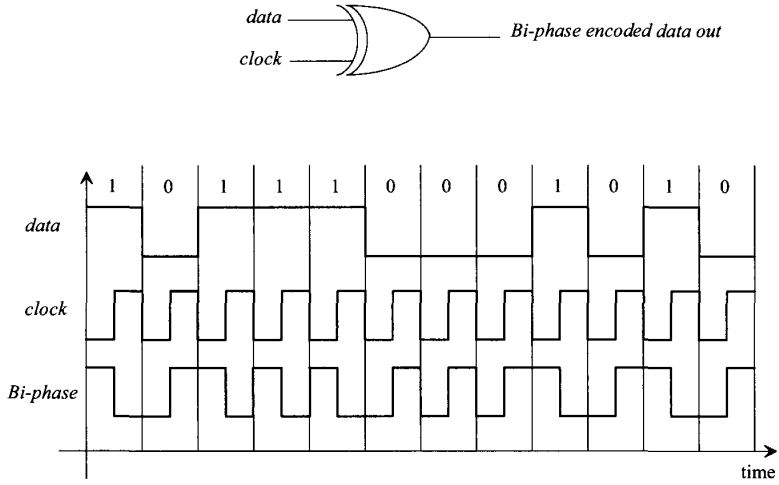


Figure 19.43 Bi-phase data encoding.

19.4.1 Clock Recovery from NRZ Data

One of the most important steps in the design of a communication system is the selection of the transmission format, that is, NRZ, bi-phase, or some other format, together with use of parity, cyclic redundancy code, or some other encoding format. In this section we discuss some of the considerations that go into the design of a clock-recovery DPLL in a system that uses NRZ.

Let's begin this discussion by considering the NRZ *data* and *clock* shown in Fig. 19.45. Let's further assume that these signals are the inputs to an XOR PD in a DPLL, which is not in lock since the clock is not aligned properly to the data. The resulting output of the XOR PD is shown in this figure as well. If we were to average this output using a loop filter, we would get *VDD*/2. In fact, it is easy to show that shifting the *clock* signal in time has no effect on the average output of the PD. Why? To answer this question, let's use some numbers. Assume that a bit width of *data* is 10 ns (which is also the period of the *clock*). The frequency of the square wave resulting from the alternating

IPR2023-00698
Theta EX2019

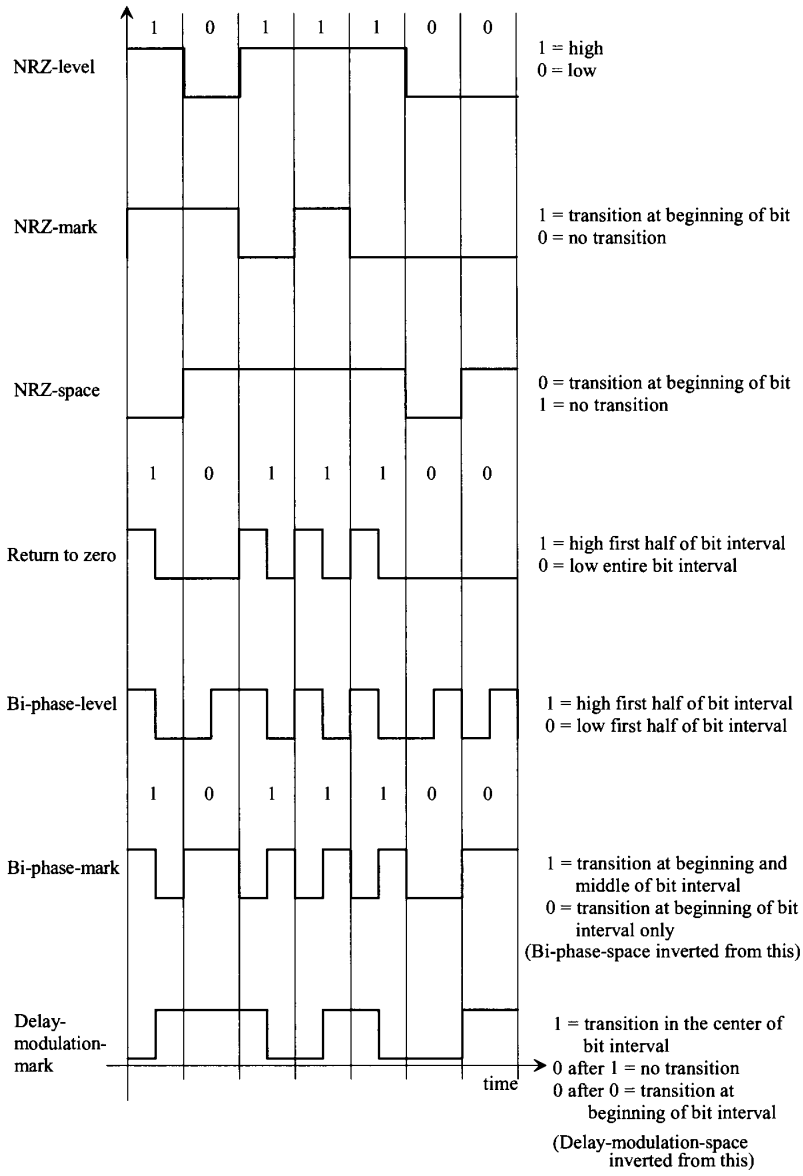


Figure 19.44 Data transmission formats.

IPR2023-00698
Theta EX2019

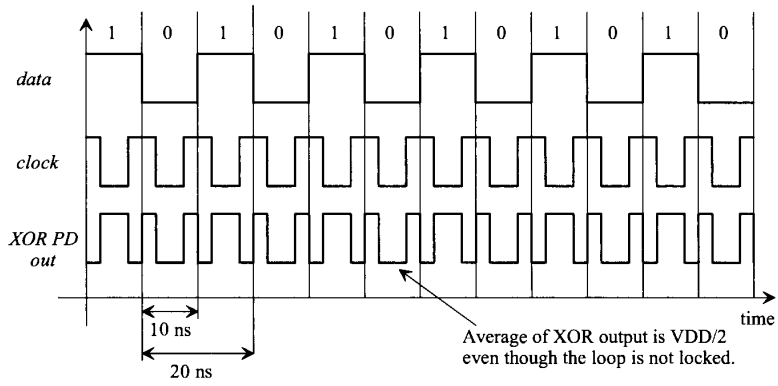


Figure 19.45 The problems of using clock without the divide by 2 to lock on data.

strings of ones and zeros is 50 MHz. We know that if we take the Fourier Transform of a square-wave, only the odd harmonics (i.e., 50, 150, and 250 MHz) are present. Since the *clock* signal is at 100 MHz, there is no energy or information common between the *clock* and *data* signals. To remedy this, we divide the clock down in frequency, *dclock*, so that it is at the same frequency as the alternating ones and zeros of the data (divide by two).

The next problem we encounter, if we use the divide by two in the feedback loop, occurs when we get *data* that is a repeating string of 2 ones followed by 2 zeroes, Fig. 19.46 (the *dclock* is not locked to the data). Again, there is no common information between the two inputs, and the resulting XOR PD output will always have an average of *VDD*/2. In this case, however, the *dclock* is running at 50 MHz, and the *data* is a square-wave of 25 MHz. Should we divide the *clock* down further to avoid this situation? The answer is no. However many times we divide the clock down, we can still come up

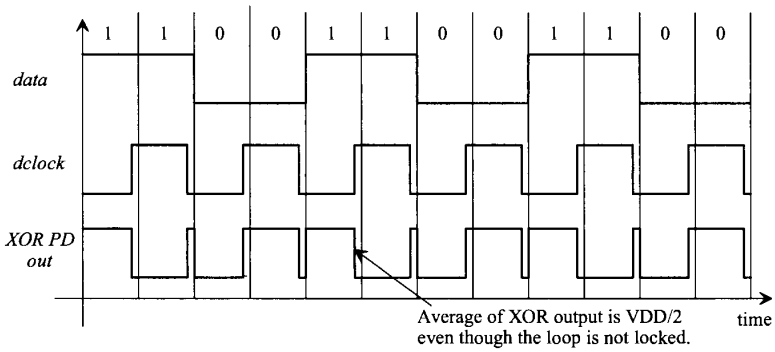


Figure 19.46 Problems trying to lock on a data stream that is one-half the dclock frequency.

IPR2023-00698
Theta EX2019

with a data string that will not allow the loop to lock. Also, it is the actual edge transitions (the frequency of the *data* and *dclock*) that is used when the inputs are not pure squarewaves. Increasing the width of *dclock* has the effect of removing information and making it more difficult to lock up to the data. One solution to this problem is to use odd-parity with an 8-bit word (9 bits total) and eliminate, at the transmitting end, the possibility of all 8 bits being high, that is, 11111111 or 255. It is then impossible to generate a square-wave.

The restrictions on the data pattern in a communications system using NRZ data can be reduced by detecting the edges of the input data with an edge detector circuit (Fig. 19.47). The delay through the inverters sets the width of the output pulses, labeled "*Edge out*" in this circuit. The frequency content of the output pulses will always contain energy at the *clock* frequency and thus the loop can lock up on the data.

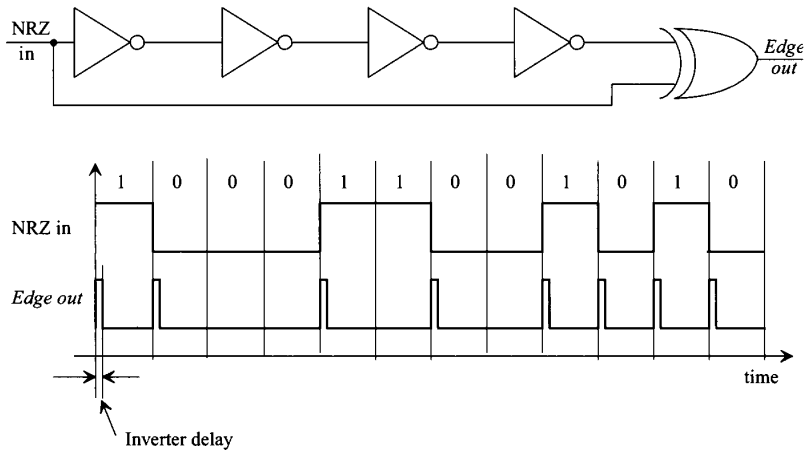


Figure 19.47 Detecting the edges in NRZ data.

As an example, consider the block diagram and data shown in Fig. 19.48. The *Edge output* is connected as the *input* of an XOR-based DPLL. The output of the VCO, *clock*, will lock up on the center of the *Edge output*, that is, the rising edge of the *clock* signal will become aligned with the center of *Edge out*. Averaging *PD out*, in this figure, results in *VDD*/2. If the clock is shifted to the left or right in time, the average value of *PD out* will shift down or up, causing the VCO frequency to change and keep the *clock* aligned to the center of *Edge out*. Several practical problems exist with this configuration. The delay through the inverters should be constant whether a high-to-low or a low-to-high transition is propagating through the inverters. Also, for best performance the delay of the inverters, or whatever element is used for the delay (one common element for high-speed applications is a microstrip line) should be close to one-half of the bit-interval time. This delay is important as it directly affects the gain of the phase detector and therefore the transient properties of the DPLL.

IPR2023-00698
Theta EX2019

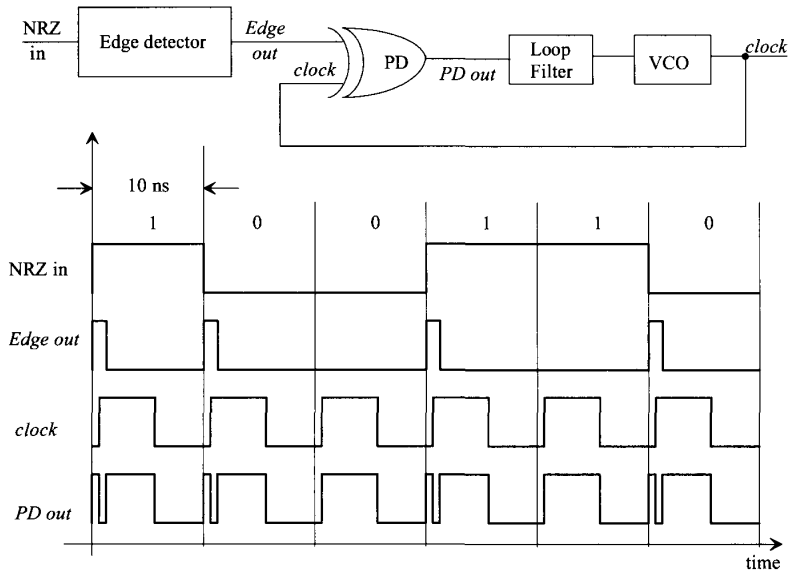


Figure 19.48 Clock-recovery circuit for NRZ using an edge detector. Note that the DPLL is in lock, when the rising edge of clock is centered on the edge output pulse.

Not having the *clock* aligned to the center of the data bit can cause problems in high-speed, clock-recovery circuits. Simply adding a delay in series with the *clock* signal does not solve this problem since the temperature dependence and process variations of the associated circuit do not guarantee proper alignment. What is needed is a circuit that is *self-correcting*, causing the clock signal to align to the center of the data bit independent of the data-rate, the temperature, or process variations.

The Hogge Phase Detector

The PD portion of a self-correcting, clock-recovery circuit is shown in Fig. 19.49 along with associated waveforms for a locked DPLL. Nodes A and B are simply the input NRZ data shifted in time by one-half bit-interval and one bit-interval, respectively. The outputs of the phase detector are labeled *Increase* and *Decrease*. If *Increase* is low more often than *Decrease*, the average voltage out of the loop filter and thus the frequency out of the VCO will decrease. A loop filter that can be used in a self-correcting DPLL is shown in Fig. 19.50a. This filter is the active-PI loop filter discussed in Sec. 19.3.1, with an added input to accommodate both outputs of the PD. Figure 19.50b shows the resulting waveforms in a DPLL where the clock is leading the center of the NRZ data, and thus *Increase* is high less often than *Decrease*. If the NRZ data was lagging the center of the data bit, *Decrease* would be high less often than *Increase,* resulting in an increase in the loop-filter output voltage. Note that in this discussion we have neglected the propagation delays present in the circuit. For a high-speed, self-correcting PD design, we would have to analyze each delay in the PD to determine their effect on the performance of the DPLL.

IPR2023-00698
Theta EX2019

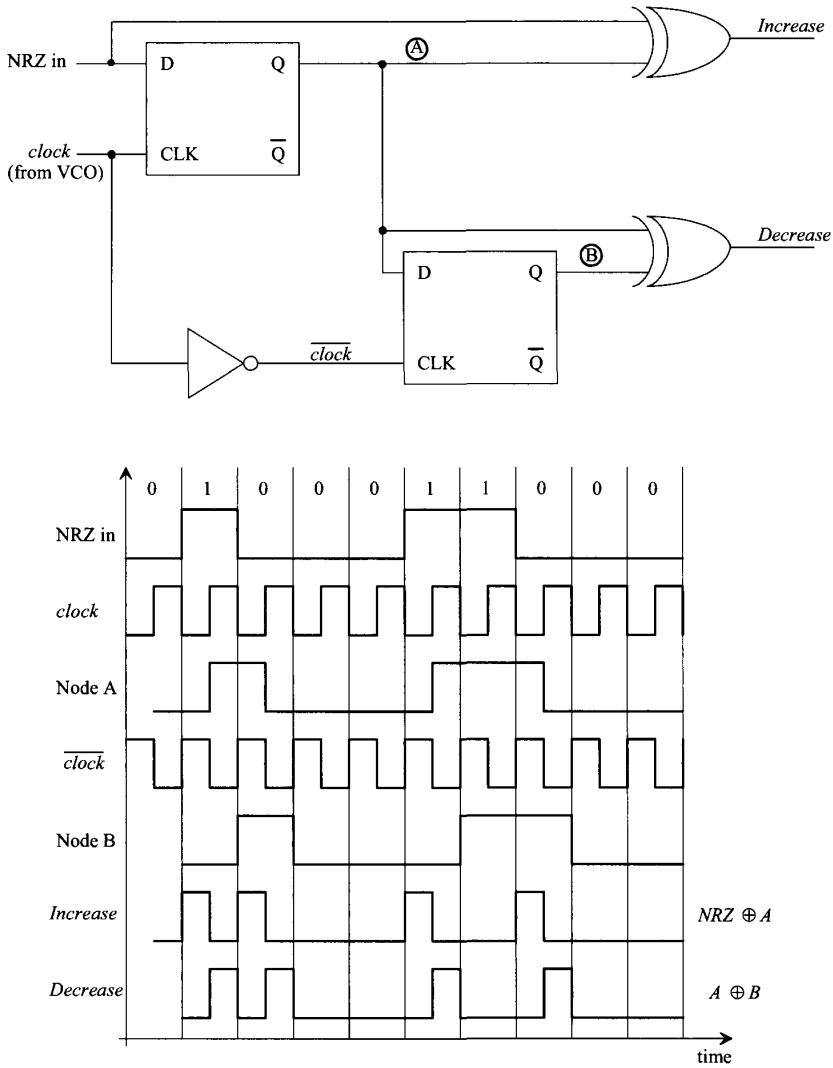


Figure 19.49 The PD (Hogge) portion of a self-correcting, clock-recovery circuit in lock.

IPR2023-00698
Theta EX2019

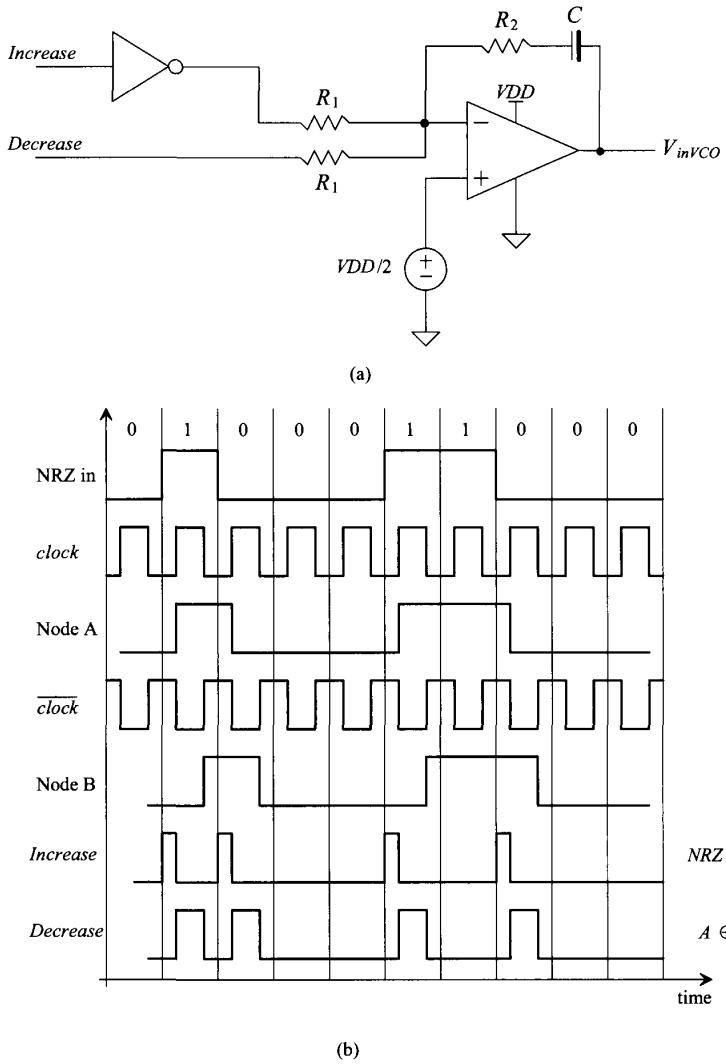


Figure 19.50 (a) Possible loop filter used in a self-correcting (Hogge) DPLL and (b) waveforms when the loop is not in lock and the clock leads the center of the data.

IPR2023-00698
Theta EX2019

Jitter

Jitter, in the most general sense, for clock-recovery and synchronization circuits, can be defined as the amount of time the regenerated clock varies once the loop is locked. Figure 19.51a shows the idealized case when the *clock* doesn't jitter, while Fig. 19.51b shows the actual situation where the clock-rising edge moves in time (jitters). In these figures, the oscilloscope is triggered by the rising edge of the *data*. In the following discussion, we neglect power supply and oscillator noise, that is, we assume that the oscillator frequency is an exact number that is directly related to the VCO input voltage. In the section following this one, we cover delay-locked loops and further discuss the limitations of the VCO.

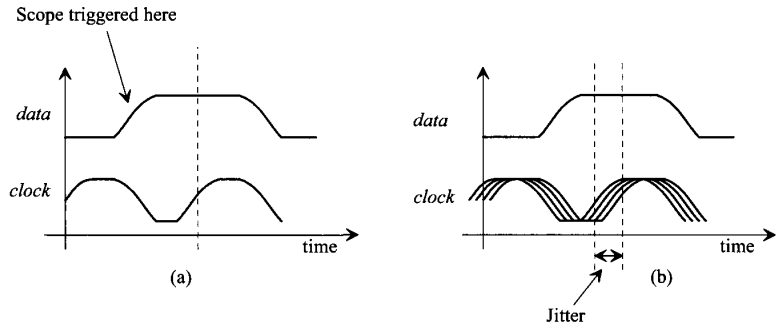


Figure 19.51 (a) Idealized view of clock and data without jitter and (b) with jitter.

Consider using the charge pump with the self-correcting PD shown in Fig. 19.52. When the loop is locked (from Fig. 19.49), both increase and decrease occur (with the same width) for every transition in the incoming data. Note that unlike a PFD/charge pump combination where the output of the charge pump remains unchanged when the loop is locked, the self-correcting PD/charge-pump combination generates a voltage ripple on the input of the VCO (similar to the XOR PD with RC or Active PI filter)[4]. Let's assume that this ripple is 10 mV and use the values for VCO gain and frequencies given in Ex. 19.5 to illustrate the resulting jitter introduced into the output clock. The change in output frequency resulting from this ripple is $10\text{mV}\cdot(1.57 \times 10^9$ radians/V·s)·$(1/2\pi)$ or 2.5 MHz. This means that the output of the DPLL will vary from, say, 100 MHz to 102.5 MHz. In terms of a jitter specification (see Eq. (19.41)), the clock jitter is (roughly) 250 ps (2.5% of the output clock's period).

From this example, *data dependent jitter*, can be reduced by

1. *Reducing the gain of the VCO.* Ripple on the input of the VCO has less of an effect on the output frequency. The main disadvantage, in the most general sense, of reducing the gain of the VCO is that the range of frequencies the DPLL can

[4] Of course, the self-correcting PD should not be used in most frequency synthesis applications. Similarly, the PFD should not be used in most clock-recovery applications.

IPR2023-00698
Theta EX2019

lock up on is reduced. Also, the VCO gain strongly affects the ability to fabricate the VCO without postproduction tuning.

2. *Reducing the bandwidth of the loop filter.* This has the effect of reducing the actual ripple on the input of the VCO. The main drawback here is the increase in chip real estate needed to realize the larger components used in the filter.

3. *Reducing the gain of the PD.* This also has the effect of reducing the ripple on the input of the VCO. This method is, in general, the easiest when using the charge pump since reducing the gain is a simple matter of reducing I_{pump}. Substrate noise can become more of a factor in the design.

Another way of stating the above methods of reducing jitter is simply to say that the forward loop gain of the DPLL, that is, $K_{PD}K_FK_{VCO}$, should be made small. The main problems with using small forward loop gain are the reductions in lock range and pull-in range coupled with the associated increases in pull-in and lock times.

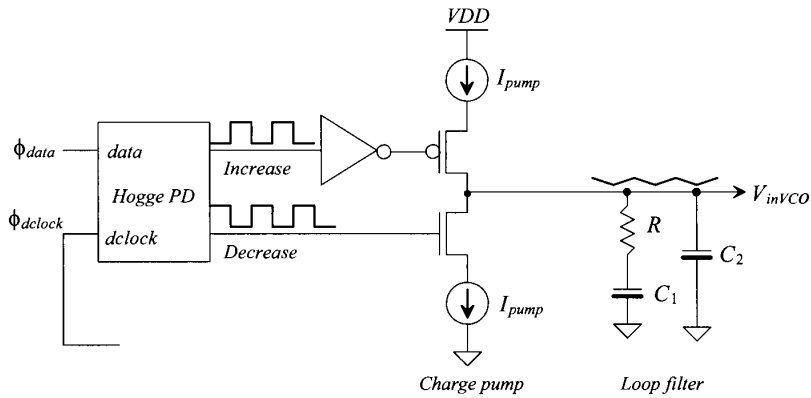


Figure 19.52 Self-correcting PD with charge pump output.

19.5 Delay-Locked Loops

Problems with PLL output jitter resulting from the VCO output frequency changing (often called oscillator or phase noise) with a constant input voltage (V_{inVCO} = constant) has led to the concept of a delay-locked loop (DLL). Figure 19.53 shows the basic block diagram of a DLL. *Assuming* that a reference clock is available at exactly the correct frequency, the input data is delayed through a voltage-controlled delay line (VCDL) a time t_o until it is synchronized with the reference clock. Jitter is reduced by using an element, the VCDL, that does not generate a signal (like the VCO did). The transfer function ϕ_{clock}/ϕ_{out} is zero (the phase of the reference clock is taken as the reference for the other signals in the DLL, i.e., $\phi_{clock} = 0$), so that oscillator noise and the resulting jitter are not factors in DLL design. The jitter considerations discussed in the last section, however, are still a concern since any ripple on the output of the loop filter will cause jitter.

IPR2023-00698
Theta EX2019

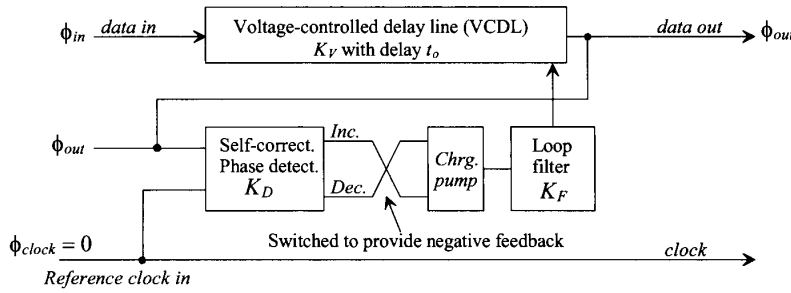


Figure 19.53 Block diagram of a delay-locked loop.

The phase (in radians) of the input data is related to the phase of the output data by

$$\phi_{out} = \phi_{in} + t_o \cdot \frac{2\pi}{T_{clock}} \tag{19.63}$$

where T_{clock} is the period of the reference clock (or half of the period of the *data in* for a string of alternating ones and zeros). The gain of the VCDL can be written in terms of the delay, t_o, by

$$t_o = K_V \cdot V_{indel} \tag{19.64}$$

where K_V has units of seconds/V and V_{indel} is the voltage input to the VCDL from the loop filter. The minimum and maximum delays of the VCDL should, in general, lie between $T_{clock}/2$ and $1.5T_{clock}$ for proper operation. The output of the loop filter (input to the VCDL) can be written as

$$V_{indel} = \phi_{out} \cdot K_D \cdot K_F \tag{19.65}$$

The overall transfer function may now be written as

$$\frac{\phi_{out}}{\phi_{in}} = \frac{1}{1 - K_D K_F K_V \cdot \omega_{clk}} \tag{19.66}$$

where $\omega_{clk} = 2\pi/T_{clock}$. The gain of the self-correcting PD with charge-pump output, with the help of Fig. 19.54 and noting that *Increase* and *Decrease* can occur at the same time, is

$$K_D = -\frac{I_{pump}}{\pi} \text{ (amps/radian)} \tag{19.67}$$

The negative sign is the result of switching the *Increase* and *Decrease* outputs of the self-correcting PD when connecting to the charge pump. This switch is required to provide negative feedback around the loop. Another benefit of the DLL is that the loop filter can be a simple capacitor, which results in a first-order feedback loop, that is,

$$K_F = \frac{1}{sC_1} \tag{19.68}$$

IPR2023-00698
Theta EX2019

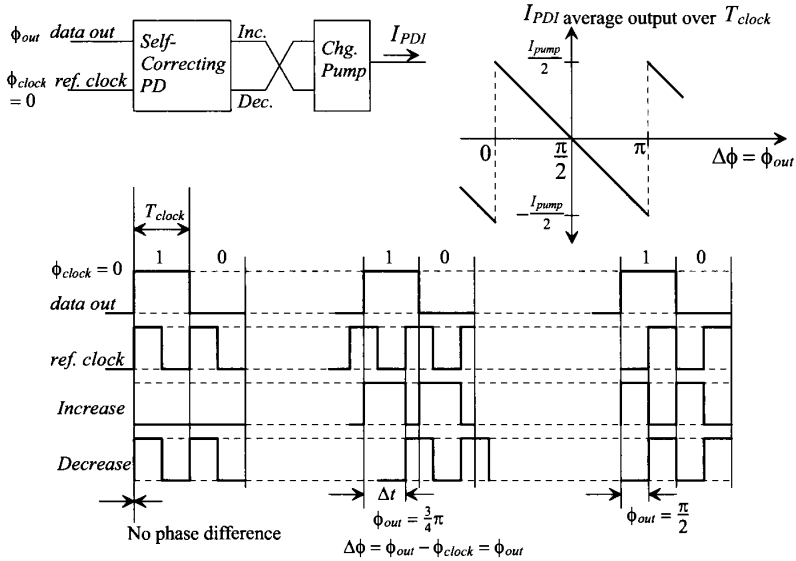


Figure 19.54 Self-correcting PD output for various inputs (assuming input data is a string of alternating ones and zeros).

The transfer function of the DLL relating the input data to the time-shifted output is

$$\frac{\phi_{out}}{\phi_{in}} = \frac{1}{1 + \frac{I_{pump}}{\pi} \cdot \frac{1}{sC_1} \cdot K_V \cdot \omega_{clk}} = \frac{s}{s + K_V \cdot \frac{2I_{pump}}{C_1 T_{clock}}} \tag{19.69}$$

We know that the frequency of the reference clock must be exactly related to the frequency of the input data. However, there will exist instantaneous changes in the phase of the input data which the output of the DLL should follow. Modeling instantaneous changes in ϕ_{in} by $\Delta\phi_{in}/s$ (a step function with an amplitude of $\Delta\phi_{in}$), we get a change in output phase given by

$$\Delta\phi_{out} = \frac{\Delta\phi_{in}}{s + K_V \cdot \frac{2I_{pump}}{C_1 T_{clock}}} \tag{19.70}$$

The time it takes the DLL to respond to an input step in phase is simply

$$T_r = 2.2 \cdot \frac{C_1 T_{clock}}{K_V \cdot 2I_{pump}} = \text{number of clock cycles} \cdot T_{clock} \tag{19.71}$$

This time can be decreased by making C_1/I_{pump} small, which from our discussion in the last section, has the result of increasing the output pulse jitter (i.e., jitter dependent on the input data pattern). Decreasing C_1/I_{pump}, increases the ripple on the control voltage of the VCDL. Similarly, increasing K_V (the time/volt delay of the VCDL) increases jitter since a given ripple on the control voltage of the VCDL will have a larger effect on the delay. Again, trade-offs must be made between responses to input variations and output jitter.

IPR2023-00698
Theta EX2019

Delay Elements

The VCDL is an important component of the DLL. Figure 19.55a shows the basic implementation of a VCDL using adjustable delay inverters. The last two inverters in the VCDL ensure that clean digital signals are output from the line. Figure 19.55b shows the circuit schematics for possible delay elements. The first delay element should be recognized as a current-starved inverter discussed earlier in the chapter. The second delay element is nothing more than an inverter with a variable load. In practice, these delay elements are rarely used because of the susceptibility to noise and power supply variations. Rather, fully-differential delay elements are used.

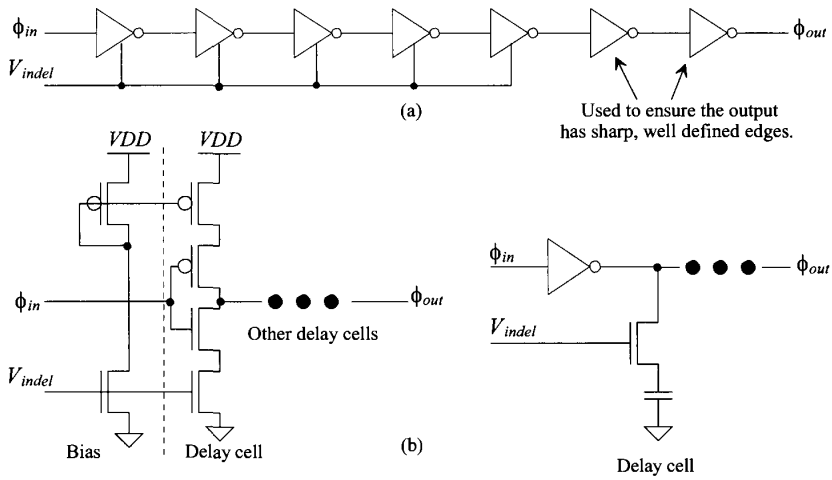


Figure 19.55 (a) VCDL made using inverter delay cells and (b) possible delay cells.

Figure 19.56 shows the connection of a fully differential VCDL. Using any number of stages, a fully differential VCO can be implemented with the delay elements (and the proper feedback), while using an even number with the inverting (noninverting) output fed back and connected to the noninverting (inverting) input in-phase (I) and quadrature (Q) signals can be generated.

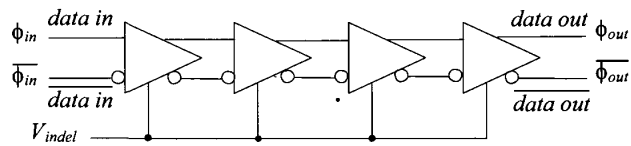


Figure 19.56 Implementation of a fully differential VCDL.

IPR2023-00698
Theta EX2019

Practical VCO and VCDL Design

An example of a practical delay stage is shown in Fig. 19.57. The control voltage (that controls the delay if used in a VCDL or the oscillation frequency if the element is used in a VCO) can be generated from some linear voltage to current converter (like the one seen in Fig. 19.17 using a resistor, M5R, and M6R). The reason that this circuit is labeled "practical" can be understood by first considering what happens when there is noise on *VDD*. The noise causes a voltage variation across the PMOS current source, MPC. Ideally, this doesn't change the current through MPC. The output voltages of the delay cell swing between V_{REF} and ground. If there is noise on ground, it feeds equally into each output. This common-mode noise is then, ideally, rejected by the differential amplification action of the next stage. The bias circuit is a half-replica of the delay stage and is used to bias the delay elements so that the swing is up to V_{REF} when the gate of one of the PMOS switches is at ground.

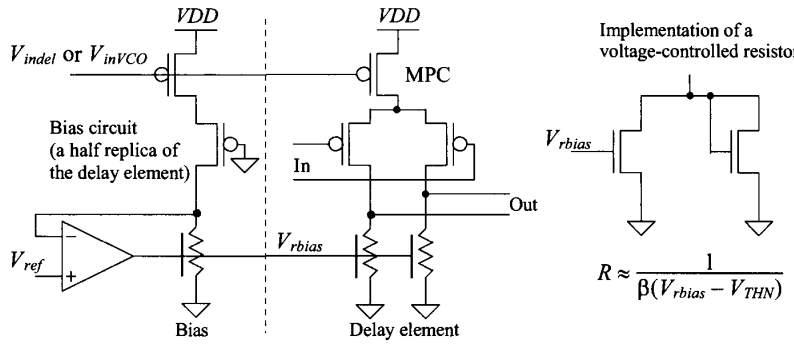


Figure 19.57 A differential delay element based on a voltage-controlled resistor. The bias circuit adjusts the value of the resistors used in the delay elements to sink the current sourced by the p-channel MOSFETs.

19.6 Some Examples

To finish up this chapter, let's present some examples with discussions and simulations.

19.6.1 A 2 GHz DLL

To begin, let's design a 500 ps delay line for a DLL that is used for *clock synchronization*. Let's design the delay line so that when it is used in a DLL, with the input and output of the delay line synchronized, we have eight phases of the input clock (we need eight stages in the delay line).

The first circuits we need to design are the bias circuit and the voltage-to-current converter used to generate $V_{outfilter}$ and V_{rbias} (see Fig. 19.57). We'll use the basic schemes of Fig. 19.17 and 19.57, as seen in Fig. 19.58. The important issues concerning this circuit are: 1) the linearity of the V_{indel} voltage against the generated I_{REF}, 2) the sensitivity of I_{REF} to changes in *VDD*, and 3) how well the amplifier regulates node n2 to V_{REF}.

IPR2023-00698
Theta EX2019

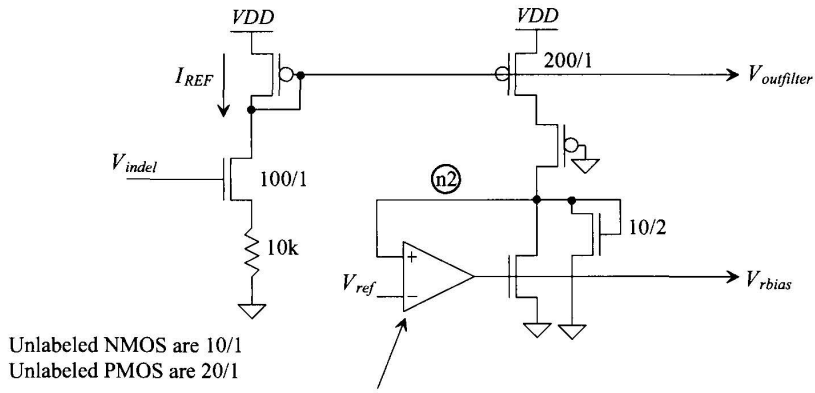


Figure 19.58 Bias circuit for a 500 ps delay line.

Figure 19.59 shows the simulation results for the bias circuit in Fig. 19.58. In (a) the x-axis is the delay line's control voltage, V_{indel} , while the y-axis is the generated reference current. Also seen in the figure, although hard to discern, is the change in *VDD* from 900 mV to 1.1 V (*VDD* is changed from 900 mV to 1.1 V in 50 mV steps, while V_{indel} is swept from 300 mV to 800 mV). Figure 19.59b shows how well node n2 is regulated to V_{REF}. V_{REF} can be generated with the beta-multiplier discussed in the next chapter (see Eq. (20.38) and Fig. 20.22).

Note that we aren't setting a lower value of current in this circuit like we did in Fig. 19.25. If we were to do this (which may be useful to reduce jitter since the resulting delay line or VCO would have lower gain), we would use a topology like the one seen in Fig. 19.60. The reference voltage is used to set the minimum current (when V_{indel} is small) through Rlow. The problem with setting the lower current with the method seen in Fig. 19.25 is that any changes in *VDD* feed directly across Rlow and change the bias current. Using the method in Fig. 19.60 with wide NMOS devices, the lower current is (roughly) $(V_{REF} - V_{THN})$/Rlow.

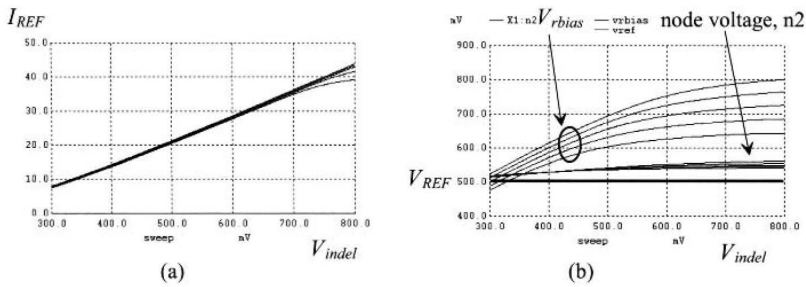


Figure 19.59 The performance of the bias circuit in Fig. 19.58.

IPR2023-00698
Theta EX2019

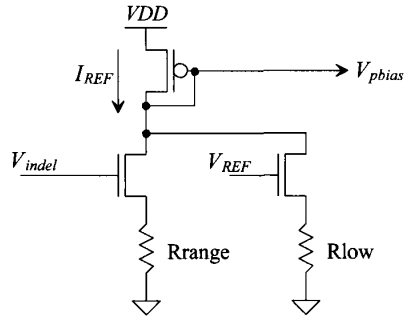


Figure 19.60 Lowering the current range in the bias circuit.

A schematic of the delay line is seen in Fig. 19.61. Notice that the current in the each delay stage was sized up by a factor of ten (the current sources in the delay cell are sized 200/1). This VCDL was simulated using the topology seen in Fig. 19.62. The control voltage was set to *VDD*/2 while the differential inputs were generated using an inverter and a transmission gate (to attempt to equalize the delays that the input signal sees to the VCDL). The outputs of the even stages are shown in this figure. Notice how they are evenly spaced. Notice, also, that the outputs only swing up to V_{REF} (= 500 mV here). As discussed earlier, this was done to minimize the effects of power supply and ground noise on the delay of the circuit.

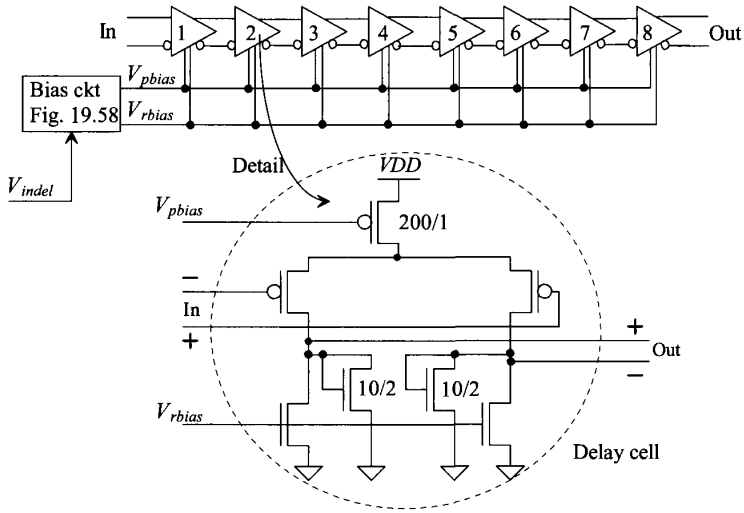


Figure 19.61 An eight-stage VCDL.

IPR2023-00698
Theta EX2019

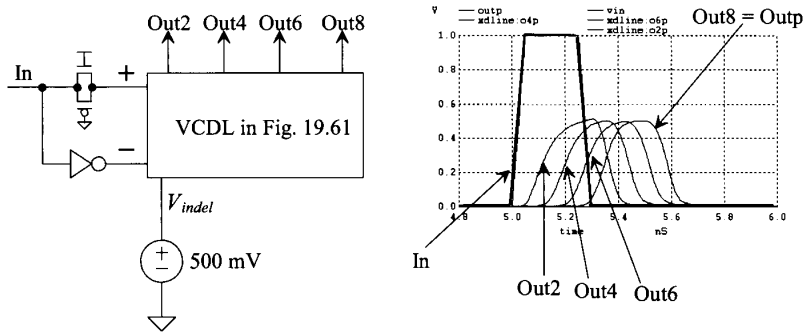


Figure 19.62 Simulating the VCDL in Fig. 19.61.

An important concern is how we regenerate full logic levels without introducing skew into our signals (see Sec. 18.3). Consider the modified VCDL seen in Fig. 19.63. Here we've changed the last stage to two diff-amps with swapped input signals (so we can generate the positive and minus outputs). Figure 19.64 shows the simulation results with this change. The shift in the eighth stage's output is small, in (a), and not much different than what is seen in Fig. 19.62. However, as needed, the output amplitude is larger. To get full logic levels, we pass these signals through inverters (that add more delay as seen Fig. 19.64b).

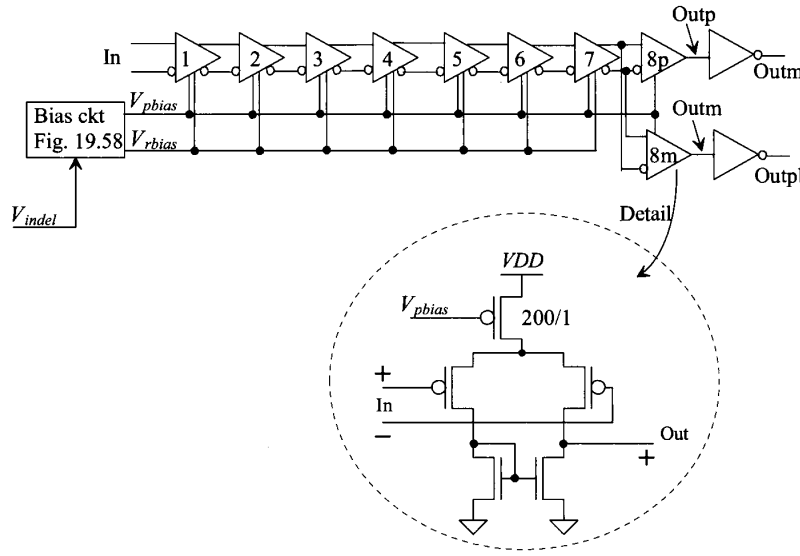


Figure 19.63 Modifying the VCDL to generate full output logic levels.

IPR2023-00698
Theta EX2019

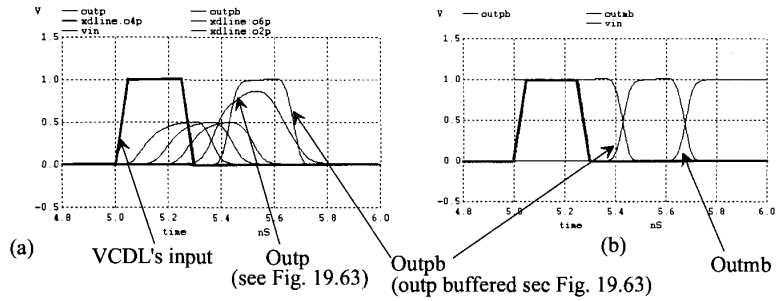


Figure 19.64 The operation of the VCDL in Fig. 19.63.

Using the simulations that generated Fig. 19.64, we can get a reasonable estimate for the VCDL's gain as

$$K_V = \frac{75\ ps}{100\ mV} = 750\frac{ps}{V}$$

The minimum delay through the VCDL is roughly 300 ps, while the maximum delay is infinite when V_{indel} moves towards V_{THN}. As seen in Fig. 19.60, we can set the minimum current that flows, I_{REF}, and thus the maximum delay. While this is an important concern we won't address it any further here.

For the PD in our DLL, let's use the PFD detector and charge pump from Ex. 19.5. The PFD's gain can be written as

$$K_D = \frac{I_{pump}}{2\pi} = \frac{10\ \mu A}{2\pi} = 1.59 \times 10^{-6}\ amps/radian$$

Using Eq. (19.71) with a lock time of 50 cycles gives

$$C_1 = \frac{(750 \times 10^{-12}) \cdot 2 \cdot (10 \times 10^{-6}) \cdot 50}{2.2} = 340\ fF$$

As indicated in the discussion following Eq. (19.71), however, this capacitor's selection is based, in some designs, not on lock time but rather the allowable jitter in the output signal once the loop is locked. A block diagram of our DLL is seen in Fig. 19.65. Simulations show that, indeed, the loop locks within 50 cycles with this value of loop filter capacitor (340 *fF*). However, the ripple on the control voltage is 20 mV. With the VCDL gain given above, we can estimate the output jitter as 15 ps. Since the period is 500 ps, this is 3% of the period. This much jitter, in a general application, may not be too bad. When we derived Eq. (19.71), we assumed that the gain of the VCDL was linear. If it is not linear the DLL can exhibit some second-order locking effects (meaning the response isn't truly first-order), resulting in a static phase error or a dead zone (the loop doesn't lock tightly to the input signal). To reduce the jitter and to linearize the loop's response, let's use a large capacitor for the loop filter (a 5 pF) in this example.

The simulated input and output signals of the DLL in Fig. 19.65 using a 5 pF capacitor are seen in Fig. 19.66. Notice the static phase error. The center of the rising edge of the input signal isn't exactly occurring at the same time as the center of the rising edge of the output signal. Further, notice the asymmetry in the output signal, which stays

IPR2023-00698
Theta EX2019

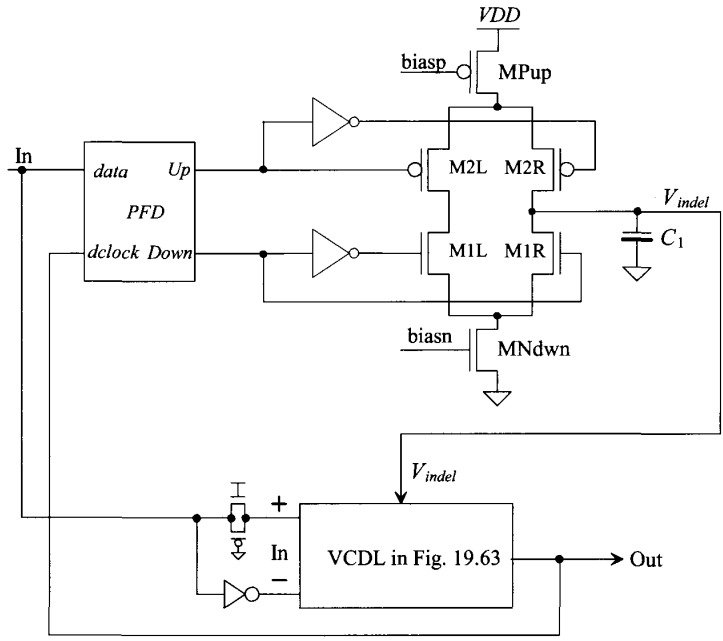


Figure 19.65 Block diagram of a DLL.

high longer than it stays low. This skew, as discussed in Sec. 18.3, is caused by different drive strengths of NMOS and PMOS devices or by differing slopes of input signals. We used the half-replica bias circuit in Fig. 19.58 to make the delay line more tolerant of power supply and ground noise. The drawback is that, at the end of the line, we do have to generate full logic levels. This causes skew between the final output of the VCDL (stage 8 in Fig. 19.63) and the outputs of the other stages in the line. The simple solution

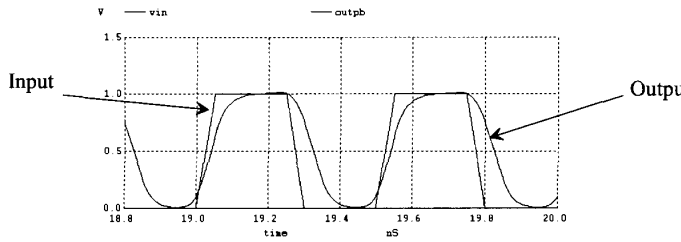


Figure 19.66 Input and output of the DLL in Fig. 19.65.

IPR2023-00698
Theta EX2019

to this problem is to use a delay stage where the outputs swing to full logic levels. To reduce the DLL's susceptability to power supply and ground noise, we can power it with its own on-chip regulated power supply, supply power and ground to the DLL from separate *VDD* and ground pads, or use some simple filtering to ensure the noise on *VDD* is slow enough that the DLL can respond and correct for variations in the VCDL's delay circuits. For example, we might, in Fig. 19.61, reduce the number of stages to four and add, on the output of each of the delay cells seen in this figure, two of the PMOS buffers seen in Fig. 18.21. We need two buffers because we swap the inputs of each buffer to generate both true and compelement output logic levels. An example of this delay cell is seen in Fig. 19.67. We only use four stages now (for the 500 ps delay line in this section) because of the extra capacitive loading on the output of each stage. The buffered delay cell outputs drive an external load while the basic delay cell outputs are used to drive the input of the next delay cell in the VCDL (only).

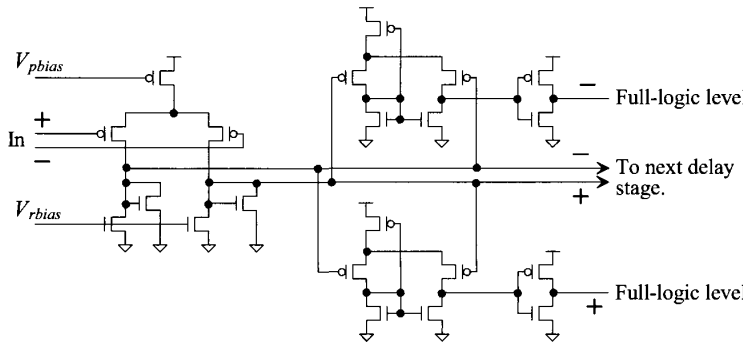


Figure 19.67 Generating full-logic levels in each delay stage.

19.6.2 A 1 Gbit/s Clock-Recovery Circuit

As another example, let's discuss and simulate the design of a 1 Gbit/s clock recovery circuit using the NRZ data format. Let's use the VCDL we've already developed in the VCO (see Fig. 19.68). The output of the VCDL is fed back to its input (with an inversion) to get the positive feedback needed for oscillations. Note that the VCO oscillates at 1 GHz when V_{inVCO} is (roughly) 350 mV (which is not *VDD*/2). While, in a production part, centering the VCO's operating frequency is important, we don't modify the VCDL in Fig. 19.63 for the PLL design here.

We will be using the Hogge phase detector in this design. Since the Hogge PD doesn't perform frequency detection (like the PFD), we must be concerned with locking on harmonics. For example, we might have an input NRZ data stream of 1000 0000. Our loop locks with the clock centered on the one and then six other edges centered on the seven zeroes. The loop is locked just not at the right clock frequency (it's locked at 7/8 of the correct frequency). While this is an **important practical** concern, we won't discuss it further until the end of the section. We will use initial conditions on the loop filter to set the initial clock frequency close to the correct value to avoid locking on a harmonic.

IPR2023-00698
Theta EX2019

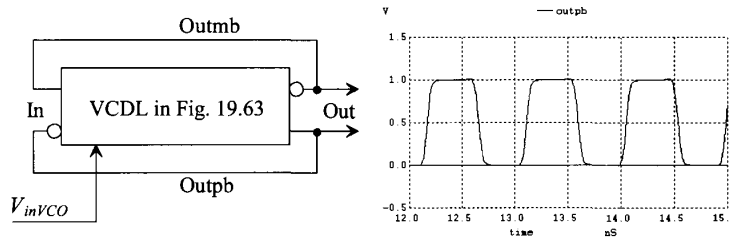


Figure 19.68 Making a VCO with the delay line in Fig. 19.63 and its output when Vinvco = 350 mV.

We know that the gain of the VCDL is 750 ps/V. Since, for one complete oscillation of the VCO, the signal must travel through the VCDL twice, we can estimate the gain of the VCO as

$$K_{VCO} = 2\pi \cdot \frac{1.175\,GHz - 1\,GHz}{100\,mV} = 11 \times 10^9\ radians/V \cdot s$$

where we assume at 350 mV the output frequency is 1 GHz (period is 1 ns). If we increase V_{inVCO} by 100 mV, we get a decrease in the period by 150 ps ($2 \times 100\,mV \cdot 750\,ps/V$) resulting in an output frequency of 1.175 GHz (period is 850 ps).

To implement the Hogge PD seen in Fig. 19.49, we'll use the same TSPC edge-triggered latch that we used for the divide-by-2 seen in Fig. 19.24. The result is seen in Fig. 19.69. We've been very careful to buffer the inputs and outputs of the PD to ensure high-speed output edges and to square up the input signals. To verify the operation, we've simulated the PD with the NRZ and clock signals seen in Fig. 19.49, Fig. 19.70. Looking at the *increase* and *decrease* signals, we see that the width of the *increase* signal is too large. On careful inspection, the cause of this error is the delay through the DFF the NRZ data sees to node A. To compensate for this delay, we'll add a delay in parallel with this path to the XOR gate, as seen in Fig. 19.71. Figure 19.72 shows the resulting simulation output. Again note that unlike the PFD, which only looks at the edges, the Hogge PD, like the XOR PD, uses the widths of its inputs to drive the loop filter.

When the Hogge PD is used with a charge pump the gain is, see Eq. (19.67),

$$K_D = \frac{I_{pump}}{\pi}$$

If we use the charge pump from the DLL in Fig. 19.65 with I_{pump} = 10 μA, the gain of the PD is 3.2 μA/radian. If we set the natural frequency, ω_n, to, 100×10^6 radians/s and ζ = 1, then using Eq. (19.58)

$$RC_1 = 20\ ns$$

and using Eq. (19.57) we can solve for C_1 (knowing that there isn't a feedback divider so N is 1) as

$$C_1 = \frac{K_D K_{VCO}}{\omega_n^2} = \frac{(3.2 \times 10^{-6})(11 \times 10^9)}{(100 \times 10^6)^2} = 3.5\ pF$$

IPR2023-00698
Theta EX2019

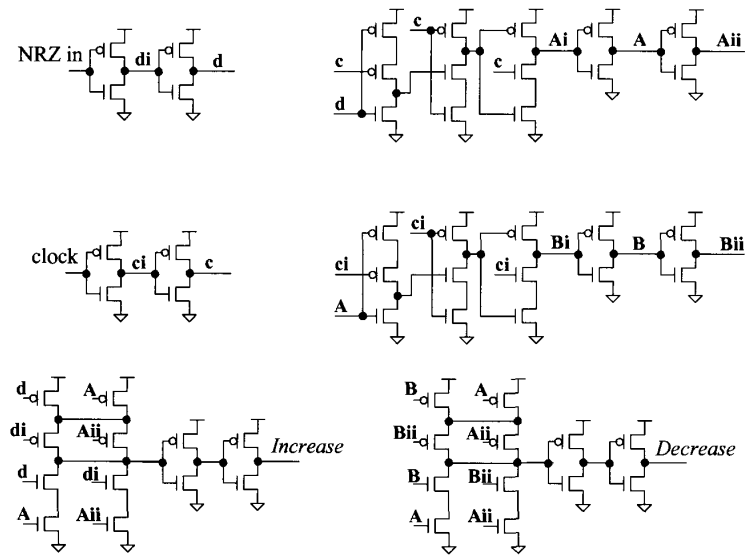


Figure 19.69 CMOS implementation of the Hogge PD.

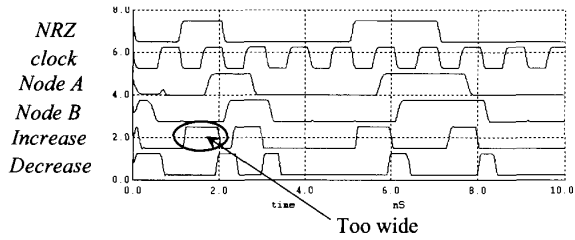


Figure 19.70 Simulating the Hogge PD in Fig. 19.69. These results should be compared to Fig. 19.49. Notice how the increased pulse widths are too wide. The result is that the Vinvco control voltage will increase above the desired value (and cause a static phase error or false locking).

Let's set C_1 to 3.5 pF, R to 5k, and C_2 to 0.35 pF. The schematic of the complete DPLL clock-recovery circuit is seen in Fig. 19.73.

In the first simulation, Fig. 19.74, we apply an alternating string of ones and zeroes to the PLL clock-recovery circuit. Since the data rate is 1 Gbit/s, the width of a one or a zero is 1 ns. In this simulation we didn't apply any initial conditions to the loop filter to start the simulation out. However, because the NRZ data is full of transitions, V_{inVCO} quickly moves to the correct voltage (around 350 mV). The important thing to note is that, again, the lock time depends on the input data.

IPR2023-00698
Theta EX2019

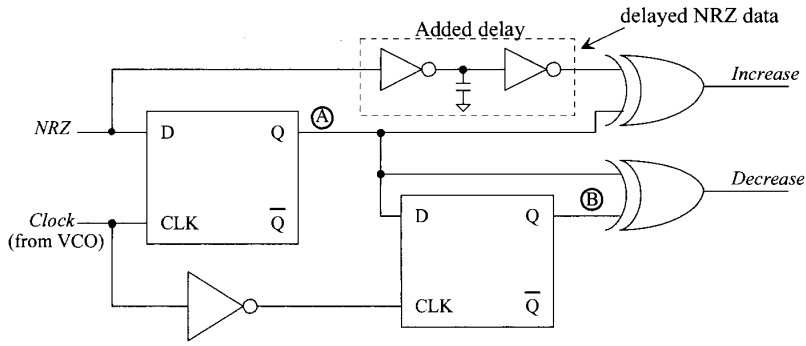


Figure 19.71 Adding a delay to the Hogge PD to compensate for the delay the NRZ data sees to node A.

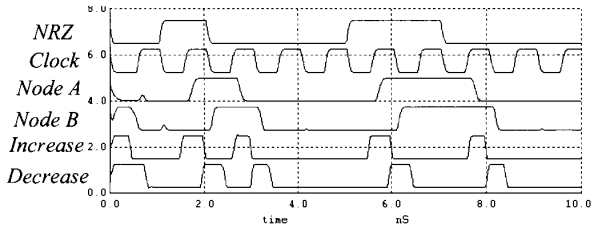


Figure 19.72 Resimulating the operation of the Hogge PD with the delay seen in Fig. 19.71 included.

For the second simulation, Fig. 19.75, we start the voltage across the loop filter, V_{inVCO} , out at 340 mV (slightly below the final value of around 350 mV). After approximately 600 ns, the loop locks. Had we not used this initial condition, the lock time (and the simulation time required to attain lock) would be quite long. Because of the large gain of the VCO, not using an initial voltage for V_{inVCO} would result in the VCO oscillating at the wrong frequency and the loop not locking correctly (locking on the wrong frequency) when the data is, as seen, a single one followed by seven zeroes (or some other sparse pattern). To avoid this problem, we can require the PLL sees a long pattern of alternating ones and zeroes to ensure lock prior to sending data (called sending a preamble) or we can reduce the gain of the VCO. The problem with sending a preamble is that it reduces the data rate through the channel. Further, large VCO gain makes it easier for the PLL to lose lock with steps in the phase shift through a communication channel. The practical solution is to reduce the gain of the VCO. In this example if we were to limit the VCO's oscillation range to, say, 950 MHz to 1050 MHz, then (with proper design of the loop filter) the PLL could acquire lock quickly and stay locked with reasonable variations in the phase shifts of the input NRZ data (these phase shifts are common in channels whose path length is not fixed, as in wireless communications). The practical problem with low VCO gain, again, is manufacturability. It's impossible to

IPR2023-00698
Theta EX2019

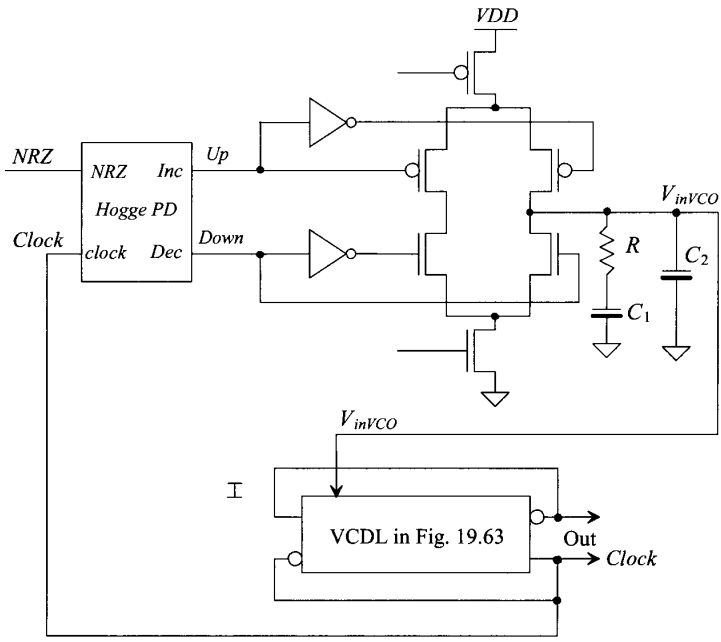


Figure 19.73 Block diagram of the DPLL used for clock-recovery discussed in this section.

design a purely monolithic CMOS oscillator (meaning no off-chip components) without some sort of hand tuning (meaning using fuses or some other sort of digital adjustment to center the oscillation frequency and limit the gain after the chips are made). Further, even if we hand-tailor the performance of each VCO in a production line, the temperature and power supply sensitivity will limit the minimum gain we can attain. Tuned circuits

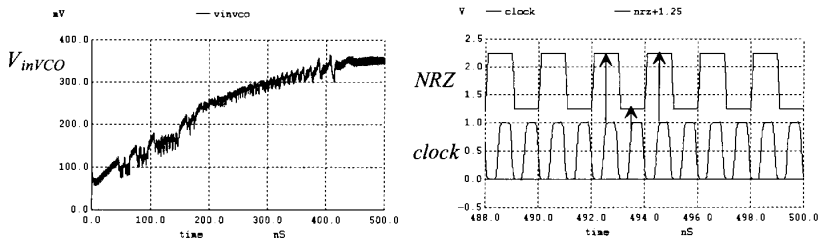


Figure 19.74 Simulating the PLL in Fig. 19.73 when the input NRZ data is an alternating string of ones and zeroes.

IPR2023-00698
Theta EX2019

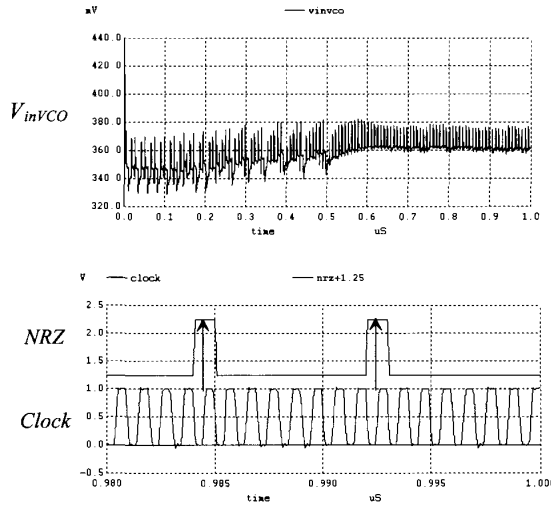


Figure 19.75 Simulating the PLL in Fig. 19.73 when the input data is a string of seven zeroes followed by a single one.

(parallel inductor-capacitor "tanks") are commonly used to help make purely monolithic clock-recovery circuits manufacturable. However, hand-tuning during Probe is still common to set the oscillation frequency and oscillation range.

ADDITIONAL READING

[1] R. E. Best, *Phase-Locked Loops: Design, Simulation, and Applications,* McGraw-Hill, 6th ed., 2007. ISBN 978-0071493758. Excellent book covering PLL design.

[2] W. F. Egan, *Phase-Lock Basics*, John Wiley and Sons, 1998. ISBN 0-4712 -4261-6 Good introduction to PLLs. See also the same author's *Frequency Synthesis by Phase Lock.*

[3] B. Razavi, *Monolithic Phase-Locked-Loops and Clock Recovery Circuits,* IEEE Press, 1996. ISBN 0-7803-1149-3. Good tutorial and collection of papers.

[4] J. Maneatis, "Low-Jitter Process-Independent DLL and PLL Based on Self-Biased Techniques", *IEEE Journal of Solid-State Circuits*, vol. 31, no. 11, pp. 1723-1732, 1996. Excellent paper discussing the design of PLLs and DLLs using self-biased techniques for very wideband operation.

[5] D. H. Wolaver, *Phase-Locked Loop Circuit Design*, Prentice-Hall, 1991. ISBN 0-1366-2743-9. Another excellent book covering PLL design.

[6] S. Haykin, *An Introduction to Analog and Digital Communications,* John Wiley and Sons, 1989. ISBN 0-471-85978-8. Covers distortionless transmission.

[7] M. G. Johnson and E. L. Hudson, "A Variable Delay Line PLL for CPU - Coprocessor Synchronization," *IEEE Journal of Solid-State Circuits,* vol. SC-23,

IPR2023-00698
Theta EX2019

pp. 1218–1223, October, 1988. Classic paper presenting the concept of a delay-locked loop.

[8] C. R. Hogge, Jr., "A Self Correcting Clock Recovery Circuit," *IEEE Journal of Lightwave Technology*, vol. LT-3, pp. 1312–1314, December 1985. Paper presenting the "Hogge" phase detector.

[9] F. M. Gardner, "Charge-Pump Phase-Lock Loops," *IEEE Transactions on Communications,* COM-28, no. 11, pp. 1849–1858, November 1980. The paper first discussing charge pump PLLs.

PROBLEMS

19.1 The XOR gate PD seen in Fig. 19.4 can exhibit input-dependent skew. In other words, the delay from one of the inputs changing to the output of the PD will not be precisely the same as the delay from the other input to the output. Design an XOR PD that doesn't exhibit input-dependent skew. Hint: see Fig. 18.16.

19.2 Verify, using simulations, that a locked PLL using an XOR PD will exhibit, after RC filtering, an average value of *VDD*/2. Show, using simulations and hand calculations, the filter's average output if the XOR PD sees a phase difference in its inputs of $-\pi/4$.

19.3 Why, in your own words, is it so important for the VCO's center frequency to match the input NRZ data rate when a passive loop filter is used. Use simulations to show the problems. Why does using the active loop filter eliminate this requirement?

19.4 Suppose, for a robust high-speed PLL using an XOR PD, that it desirable to adjust the PD's gain. Show a charge pump can be used towards this goal. Should the loop filter have two outputs? Discuss the PD's gain and how it would be adjusted. What topology would be used for a passive loop filter? for an active loop filter?

19.5 Demonstrate, using simulations, how the outputs of the XOR PD can be equivalent when the loop is true or false locking (locking on a harmonic).

19.6 Describe, in your own words and using simulations, what the dotted line in Fig. 19.8 indicates.

19.7 We know that a PFD isn't generally used in clock recovery applications because both edges (*data* and *dclock*) must present when doing a phase comparison. Suggest a scheme where the edge of the input *data* enables a phase comparison. When an edge of data is not present your circuit will "swallow" the pulse from *dclock* resulting in no edges being applied to the PFD (no phase- frequency comparison).

19.8 Demonstrate, using simulations, charge sharing between the charge pump and loop filter using the topology seen in Fig. 19.12b (and Fig. 19.13b). Show how the topology in Fig. 19.37 helps with charge sharing (simulate with and without the x1 amplifier). For the x1 amplifier, use the n-type diff-amp from Fig. 18.17 without the inverter on its output. Connect the loop filter to the gate of M1 (in the diff-amp). To make the gain "1," tie the output (which is connected to drains of M1L/M2L in Fig. 19.37) of the diff-amp (the drains of M2/M4) to the gate of M2.

IPR2023-00698
Theta EX2019

19.9 Using the VCO that generated the simulation data in Fig. 19.18, plot the VCO's center frequency against changes in *VDD*. Is this VCO insensitive to changes in *VDD*? Where does the sensitivity come from?

19.10 Discuss, and demonstrate with simulations, how to reduce the gain of the VCO used to generate the data in Fig. 19.18 (see Fig. 19.25 and the associated discussion). Your discussion should include some insight into the manufacturability of low-gain VCOs.

19.11 Design and simulate the operation of a 100 MHz VCO using the topology seen in Fig. 19.19a. The output of your VCO should be full logic levels. Your design should show a linearly frequency against V_{inVCO} curve. What is the gain of your design?

19.12 Using the step response of an RLC circuit, Fig. 19.76, demonstrate how selection of the resistor, inductor, and capacitor affect the output voltage's damping factor and natural frequency. From this plot show how natural frequency and lock time are related.

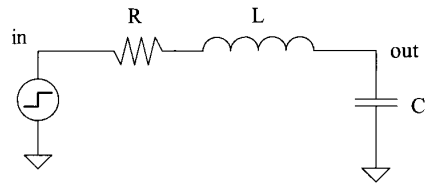


Figure 19.76 Using a second-order circuit to demonstrate how lock time and natural frequency are related.

19.13 Replace, in the DPLL that generated the waveforms in Figs. 19.27 and 19.28, the simple RC loop filter with the passive lag loop filter in Fig. 19.29. Show how the filter's component values are selected. Comment on, using simulations to support your conclusions, how the performance of the DPLL is affected.

19.14 Using the PFD in Fig. 19.33 and the charge pump in Fig. 19.36, demonstrate, with simulations, how the gain of PFD/charge-pump configuration can have a dead zone (the gain, K_{PDI}, decreases) as seen in Fig. 19.77.

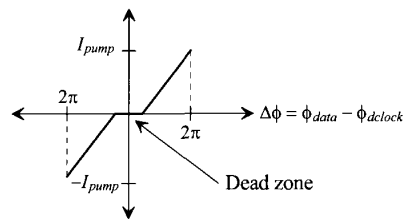


Figure 19.77 Showing the dead zone in a PFD/charge-pump.

IPR2023-00698
Theta EX2019

19.15 Suppose that the pump currents used in the DPLL in Ex. 19.5 are mismatched by 50%. Will this cause a static phase error? Why or why not? Use simulations to support your answer.

19.16 Redesign the DPLL in Ex. 19.5 so that the output remains a 100 MHz square wave signal when the input signal is changed to 25 MHz.

19.17 We used pF values for the capacitors in the loop filters in this chapter. Why not reduce the filter's layout area by using fF size capacitors? Demonstrate the problems with using such small loop filter capacitors.

19.18 What are we sacrificing by using an equalizer? What does the minus sign indicate in the slope of the phase in Fig. 19.41?

19.19 Using the DPLL from Ex. 19.2, demonstrate the false locking as seen in Fig. 19.46.

19.20 Design an edge detector, like the one seen in Fig. 19.47, for use in the DPLL in Ex. 19.2. Regenerate the simulation data seen in Figs. 19.27 and 19.28. Remember to eliminate the divide by two in the feedback path.

19.21 Derive Eq. (19.71).

19.22 Design a nominal delay line of 1 ns (when V_{indel} = 500 mV) using the current starved delay element seen in Fig. 19.55. Determine the delay's sensitivity to variations in *VDD*.

19.23 Repeat Problem 19.22 for the inverter delay cell in Fig. 19.55.

19.24 Suppose, as seen in Fig. 19.78, that instead of using a half-replica of the delay cell in Fig. 19.58, the full delay cell is used to generate V_{rbias}. Electrically is there any difference in V_{rbias} when comparing the full- and half-replica circuits? What may be the benefit of using a full-replica of the delay cell?

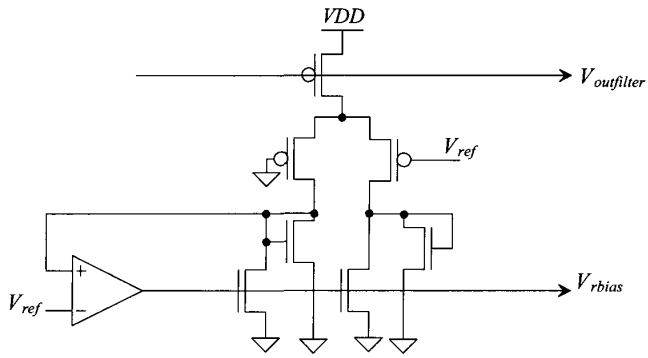


Figure 19.78 Using a full-replica of the delay element for generating Vrbias.

IPR2023-00698
Theta EX2019

19.25 For the delay line that generated the simulation results in Fig. 19.62 determine, using simulations, the delay's sensitivity to changes *VDD*. Plot the VCDL's delay as a function of *VDD* with V_{indel} held at 500 mV.

19.26 The delay element seen in Fig. 19.79 doesn't use a reference voltage. Show how this element can be used in the VCDL of Fig. 19.61. Note that the input capacitance of this delay stage is twice as large as the input capacitance of the element in Fig. 19.61 (and so you may have to use fewer stages to attain the same overall delay). How does the NMOS gate potential change with the inputs/outputs switching? Why? Is the output amplitude a function of *VDD*?

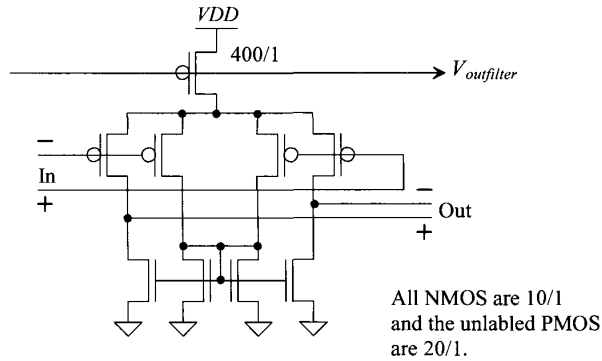


Figure 19.79 A delay element that doesn't use a reference voltage.

19.27 Use the delay element in Fig. 19.67 to implement a VCDL. Use the VCDL in the DLL seen Fig. 19.65 to generate the waveforms in Fig. 19.66. Show the 90, 180, 270, and 360 degree outputs of the VCDL swinging to full-logic levels.

19.28 The VCO output seen in Fig. 19.68 doesn't have an exactly 50% duty cycle. Will this affect a clock-recovery circuit's operation? Why or why not? Use simulations to support you answer.

19.29 Suggest another method, other than the one seen in Fig. 19.72, to equalize the widths of the *Increase* and *Decrease* outputs of the Hogge PD. Verify your design with simulations.

19.30 Must the currents in the charge pump in Fig. 19.73 be equal for proper DPLL clock-recovery operation? Why or why not? What happens if the currents aren't equal? What happens if the NMOS switches turn on at different speeds than the PMOS switches? Use SPICE to support your answers.

19.31 Modify the simulation inputs in Fig. 19.74 to show false locking. Comment on what can be done in a practical clock-recovery circuit to eliminate false locking.

IPR2023-00698
Theta EX2019

19.32 Replace the charge pump used in the DPLL in Fig. 19.73 with the active-PI loop filter seen in Fig. 19.50. Calculate the loop filter component values. Using the new loop filter, regenerate Figs. 19.74 and 19.75.

IPR2023-00698
Theta EX2019

Chapter

20

Current Mirrors

In this chapter we turn our attention towards the design, layout, and simulation of current mirrors (a circuit that sources [or sinks] a constant current). As we observed back in Fig. 9.1, and the associated discussion, the ideal output resistance, r_o, of a current source is infinite. Achieving high output resistance (meaning that the output current doesn't vary much with the voltage across the current source) will be the main focus of this chapter.

It's very important that the reader first understand the material in Ch. 9 concerning the selection of biasing currents and device sizes and how they affect the gain/speed of the analog circuits. We'll use the parameters found in Tables 9.1 and 9.2 in many of the examples in this chapter.

20.1 The Basic Current Mirror

The basic NMOS current mirror, made using M1 and M2, is seen in Fig. 20.1. Let's assume that M1 and M2 have the same width and length and note that $V_{GS1} = V_{DS1} = V_{GS2}$. Because the MOSFETs have the same gate-source voltages, we expect (neglecting channel-length modulation) them to have the same drain current. If the two resistors in the drains of M1/M2 are equal, the drain of M2 will be at the same potential as the drain of M1 (this is important). By matching the size, V_{GS} , and I_D of two transistors, we are assured that the two MOSFETs have the same drain-source voltage, ($V_{GS1} = V_{DS1} = V_{GS2} = V_{DS2}$).

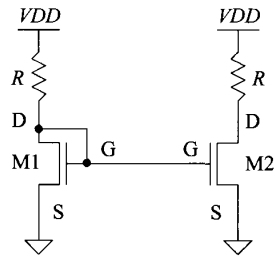


Figure 20.1 A basic current mirror.

IPR2023-00698
Theta EX2019

20.1.1 Long-Channel Design

Examine Fig. 20.2. In this figure we show a current mirror and the equivalent circuit representation of a current source. Looking at M1, we can write

$$I_{REF} = I_{D1} = \frac{KP_n}{2}\frac{W_1}{L_1}(V_{GS1} - V_{THN})^2(1 + \lambda(V_{DS1} - V_{DS1,sat})) \qquad (20.1)$$

knowing $V_{DS1} = V_{GS1}$ and $V_{DS1,sat} = V_{GS1} - V_{THN}$. For M2, we write

$$I_O = I_{D2} = \frac{KP_n}{2}\frac{W_2}{L_2}(V_{GS1} - V_{THN})^2(1 + \lambda(V_O - V_{DS1,sat})) \qquad (20.2)$$

noting $V_{GS1} = V_{GS2}$, $V_{DS1,sat} = V_{DS2,sat}$, and V_O is the voltage across the current source. Looking at the ratio of the drain currents, we get

$$\frac{I_O}{I_{REF}} = \frac{W_2/L_2}{W_1/L_1} \cdot \frac{1 + \lambda(V_O - V_{DS1,sat})}{1 + \lambda(V_{DS1} - V_{DS1,sat})} \qquad (20.3)$$

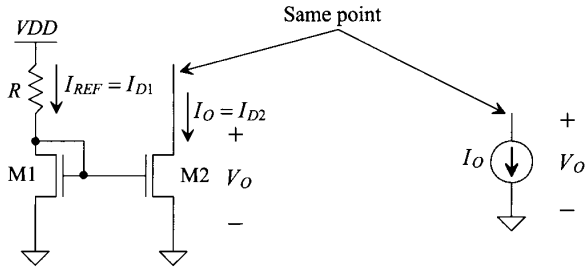


Figure 20.2 The current mirror and how we think about it.

Generally, the lengths of the devices in the current mirror are equal (let's assume they are for the moment). If, also at this time, we don't concern ourselves with channel-length modulation ($\lambda = 0$), we get a very useful result, that is,

$$\frac{I_O}{I_{REF}} = \frac{W_2}{W_1} \qquad (20.4)$$

By simply scaling the width of M2, we can adjust the size of our output current. Figure 20.3 shows an example of this scaling (using PMOS devices).

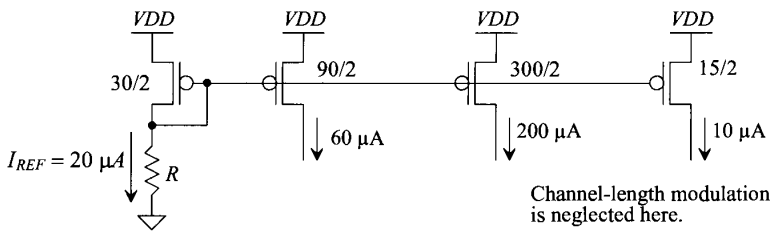


Figure 20.3 How current mirrors are ratioed.

IPR2023-00698
Theta EX2019

Example 20.1

Determine the value of the resistor, R, needed in Figs. 20.2 and 20.3 so that the reference drain currents are 20 μA. Use the long-channel parameters seen in Table 9.1. Simulate the operation of the NMOS mirror with the calculated resistor value.

In Fig. 20.2 (the NMOS current mirror), we can write

$$I_{REF} = 20\ \mu A = \frac{VDD - V_{GS1}}{R} \approx \frac{KP_n}{2} \cdot \frac{10}{2} \cdot \left(\overbrace{V_{GS1}}^{1.05\ V} - \overbrace{V_{THN}}^{0.8} \right)^2 = \frac{KP_n}{2} \cdot \frac{10}{2} \cdot (0.25)^2$$

(notice how we indicate "approximately" because channel-length modulation is not included in the equations) or

$$R = \frac{5 - 1.05}{20\ \mu A} \approx 200\ k\Omega$$

For Fig. 20.3 (the PMOS current mirror), we can write

$$I_{REF} = 20\ \mu A = \frac{VDD - V_{SG}}{R} \approx \frac{KP_p}{2} \cdot \frac{30}{2} \left(\overbrace{V_{SG}}^{1.15\ V} - \overbrace{V_{THP}}^{0.9} \right)^2 = \frac{KP_p}{2} \cdot \frac{30}{2} \cdot (0.25)^2$$

or

$$R = \frac{5 - 1.15}{20\ \mu A} \approx 200\ k\Omega$$

Simulation of the operation of the NMOS current mirror is seen in Fig. 20.4. The reference current isn't exactly 20 μA (and we shouldn't expect it to be). The x-axis is a sweep of the voltage across the current source, V_O. Note that below $V_{DS,sat}$ (= 250 mV here) M2 triodes and the output current, I_O , goes to zero. The output *compliance* range for this current source (the range of output voltages where the current source behaves like a current source, that is, not an open or a resistor) is between VDD and $V_{DS,sat}$. The point where $V_O = V_{DS1} = V_{GS1}$ is where $I_O = I_{REF}$ (again this is important for matching two currents). Finally, note that I_{REF} and V_{GS1} are not dependent on V_O. ■

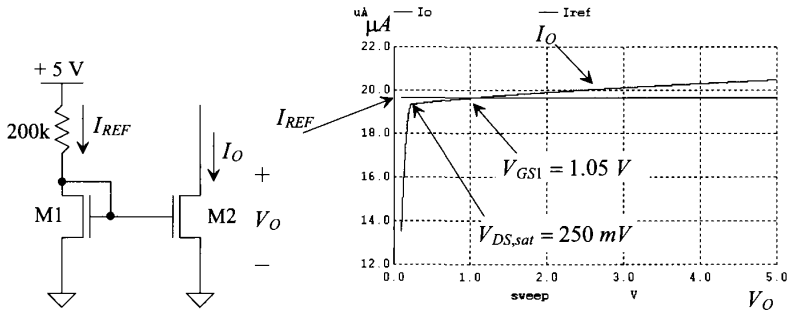


Figure 20.4 The operation of an NMOS current mirror.

IPR2023-00698
Theta EX2019

20.1.2 Matching Currents in the Mirror

Many analog applications are susceptible to errors due to layout. In circuits in which devices need to be matched, layout becomes a critical factor. For example, in the basic current mirror shown in Fig. 20.2, first-order process errors can cause the output current, I_O, to be significantly different from the reference current. Process parameters such as gate-oxide thickness, lateral diffusion, oxide encroachment, and oxide charge density can drastically affect the performance of a device. Layout methods can be used to minimize the first-order effects of these parameter variations.

Threshold Voltage Mismatch

In a given current mirror application, the values for the threshold voltages are critical in determining the overall accuracy of the mirror. Again, examine the basic current mirror shown in Fig. 20.2. Since both devices have the same value for V_{GS} and assuming the sizes and transconductance parameters for both devices are equal, let's examine the effect of a mismatch in threshold voltages between the two devices. If it is assumed that the threshold mismatch is distributed across both devices such that

$$V_{THN1} = V_{THN} - \frac{\Delta V_{THN}}{2} \tag{20.5}$$

$$V_{THN2} = V_{THN} + \frac{\Delta V_{THN}}{2} \tag{20.6}$$

where V_{THN} is the average value of V_{THN1} and V_{THN2} and ΔV_{THN} is the mismatch, then

$$\frac{I_O}{I_{REF}} = \frac{\frac{KP_n}{2}\frac{W}{L}(V_{GS} - V_{THN} - \frac{\Delta V_{THN}}{2})^2}{\frac{KP_n}{2}\frac{W}{L}(V_{GS} - V_{THN} + \frac{\Delta V_{THN}}{2})^2} = \frac{\left[1 - \frac{\Delta V_{THN}}{2(V_{GS}-V_{THN})}\right]^2}{\left[1 + \frac{\Delta V_{THN}}{2(V_{GS}-V_{THN})}\right]^2} \tag{20.7}$$

If both expressions are squared and the higher order terms are ignored, then the first-order expression for the ratio of currents becomes

$$\frac{I_O}{I_{REF}} \approx 1 - \frac{2\Delta V_{THN}}{V_{GS} - V_{THN}} = 1 - \frac{2\Delta V_{THN}}{V_{DS,sat}} \tag{20.8}$$

Equation (20.8) is quite revealing because it shows that as V_{GS} decreases, the difference in the mirrored currents increases due to threshold voltage mismatch. This is particularly critical for devices that are separated by relatively long distances because the threshold voltage is susceptible to process gradients. *To attain high speed and to reduce the effects of threshold voltage mismatch, a large gate overdrive voltage should be used* (remembering for a long-channel process that $V_{ovn} = V_{DS,sat} = V_{GS} - V_{THN}$). Of course the drawback, for a current mirror, is a reduced range of compliance (the MOSFET enters the triode region earlier).

Transconductance Parameter Mismatch

The same analysis can be performed on the transconductance parameter, KP_n. If $KP_{n1} = KP_n - \Delta KP_n/2$ and $KP_{n2} = KP_n + \Delta KP_n/2$, where KP_n is the average of KP_{n1} and KP_{n2}, then assuming perfect matching on all other parameters, the difference in the currents becomes

$$\frac{I_O}{I_{REF}} = \frac{KP_n + 0.5\Delta KP_n}{KP_n - 0.5\Delta KP_n} \approx 1 + \frac{\Delta KP_n}{KP_n} \tag{20.9}$$

IPR2023-00698
Theta EX2019

Since KP_n is a process parameter, we might think that as CMOS scales downwards (C'_{ox} goes up) the mismatch due to oxide variations and mobility differences might get better. However, there is less averaging of these variations with the smaller layout sizes. Practically, as we're about to see, differences in V_{DS} dominate the matching behavior.

Drain-to-Source Voltage and Lambda

One aspect of current mirror design that is *critical* for generating accurate currents is the drain-to-source voltage. As seen in Fig. 20.4, the only point where the currents of the devices are actually the same is when their V_{DS} values are equal. In Eq. (20.3) the ratio of the output current to the reference current is affected by both the matching in the drain-to-source voltages (V_O and V_{DS1}) and the device λs. If, for example in the short-channel process (see Table 9.2), $V_{DS1} = .35$ V, $V_{DS2} = V_O = .75$ V, and $\lambda_1 = \lambda_2 = 0.6$ V^{-1}, then

$$\frac{I_O}{I_{REF}} = \frac{1+\lambda_2 \cdot V_O}{1+\lambda_1 \cdot V_{DS1}} = \frac{1+0.6 \cdot 0.75}{1+0.6 \cdot 0.35} = 1.20 \qquad (20.10)$$

resulting in 20% error! It is extremely important, for good matching, that the V_{DS} values of the MOSFETs in the current mirror are equal.

Layout Techniques to Improve Matching

Most general analog applications require the length of the gate to be longer than minimum since the channel-length modulation, λ , has less effect on longer devices than on shorter ones (as discussed in Ch. 9). As a result, the minimum-sized devices found in digital circuits are not used as often in general analog design (however, see Ch. 26). However, the larger devices can result in larger parasitics if some layout issues are not considered. Figure 20.5a illustrates a basic MOSFET device with a large *W/L*. The implant resistance of the source and drain can be modeled as shown in Fig. 20.5b. The implant resistance can easily be reduced by simply adding as many contacts as possible along the width of both the source and the drain as seen in Fig. 20.5c. The increase in the number of contacts results in lower resistance, more current capability, and a more distributed current load throughout the device. However, as the device width increases, another technique is used which distributes the parasitics (both resistive and capacitive) into smaller contributions.

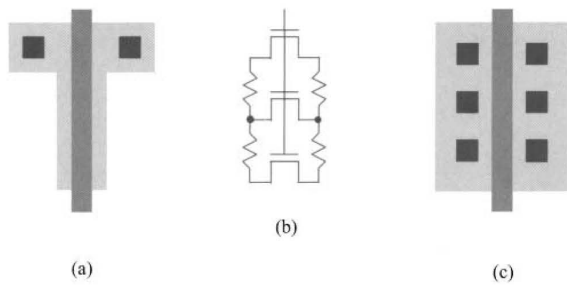


Figure 20.5 (a) Large device with a single contact and (b) its equivalent circuit. (c) Adding more contacts to reduce parasitic resistance.

IPR2023-00698
Theta EX2019

Examine Fig. 20.6. In this figure, a single device with a large *W/L* is split into several parallel devices, each with a width one-fourth of the original *W*. One result of splitting the device into several parts is smaller overall parasitic capacitance associated with the reverse-biased implant substrate diode (the drain or source depletion capacitance to substrate). Since values of C_{db} and C_{sb} are proportional to *W*, the split device reduces these parasitics by a factor of $(n + 1)/2n$ where n is the number of parallel devices and is odd. If n is even, then the C_{sb} is reduced by one-half and C_{db} is reduced by $(n + 2)/2n$. As seen in Fig. 20.6b, putting the devices in parallel also reduces the parasitic resistance in series with both the source and the drain (it gets cut roughly in half too).

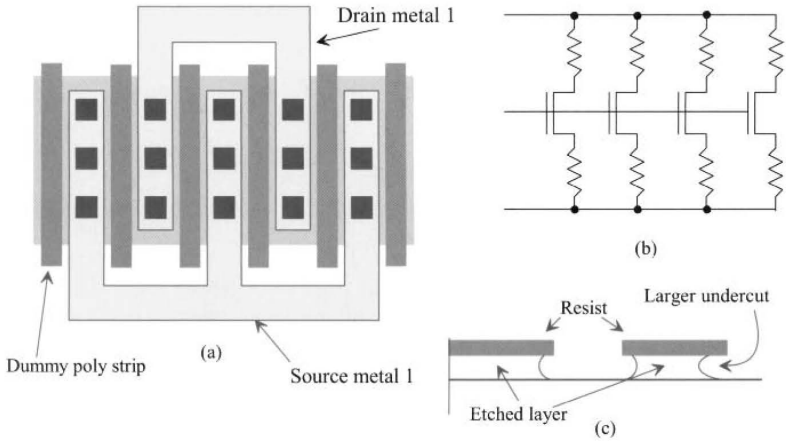


Figure 20.6 (a) A parallel device with dummy strips, (b) the equivalent circuit, and (c) undercutting.

Notice also, that Fig. 20.6a has dummy poly strips on both sides of the device. These strips are used to help minimize the effects of undercutting the poly on the outer edges after patterning, Fig. 20.6c. If the dummy strips had not been used, the poly would have been etched out more under the outermost gates, resulting in a mismatch between the four parallel devices.

When matching two devices, it is imperative that the two devices be as symmetrical as possible. Always orient the two devices in the same direction, unlike those illustrated in Fig. 20.7.

Splitting the devices into parallel devices and interdigitizing them can distribute process gradients across both devices and thus improve matching. An example of this is seen in Fig. 20.8. In (a) the current mirror seen in (b) is laid out. Each MOSFET in (b) is split up into four MOSFETs. If the W/L of each MOSFET in (b) is 80/2, then the size of each MOSFET (finger) in (a) is 20/2. Note the use of dummy poly strips on this layout. A good exercise at this point is to lay out the mirror of Fig. 20.8 in a common-centroid arrangement (see Ch. 5). Using a large layout area (long lengths and wide widths) and common-centroid layouts can result in significant improvements in matching.

IPR2023-00698
Theta EX2019

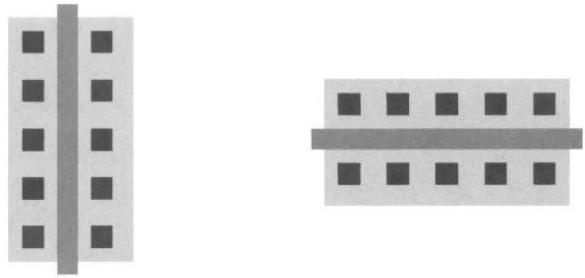


Figure 20.7 Devices with differing orientation (bad).

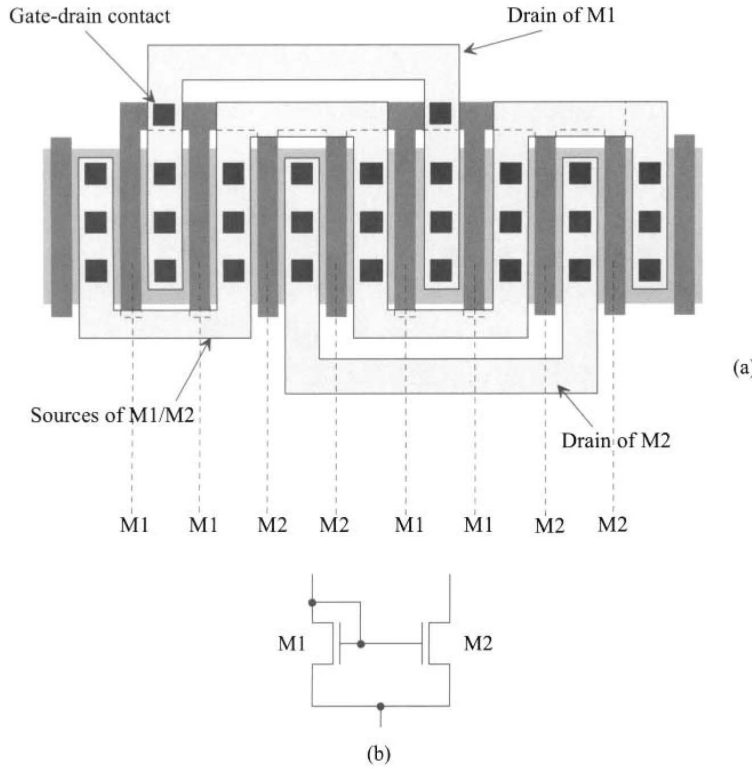


Figure 20.8 (a) Layout of a simple current mirror using interdigitation and (b) equivalent circuit.

IPR2023-00698
Theta EX2019

Layout of the Mirror with Different Widths

When laying out the current mirror, the lateral diffusion, L_{diff}, under the gate oxide (Fig. 5.13) and the oxide encroachment, W_{enc}, can affect the actual length and width, respectively, of the MOSFET. If not careful with the layout, this will affect the mirror's ratio. Equation 20.3 can be rewritten, without including the differences due to drain-source voltage and lambda, as

$$\frac{I_O}{I_{REF}} = \frac{(W_{2drawn} - 2W_{enc}) \cdot (L_{1drawn} - 2L_{diff})}{(W_{1drawn} - 2W_{enc}) \cdot (L_{2drawn} - 2L_{diff})} \qquad (20.11)$$

If the requirement that $L_{1drawn} = L_{2drawn}$ is imposed, then the widths of the devices determine the relative currents in the mirror. Figure 20.9a shows a current mirror layout without width compensation. If W_{enc} is 0.1, then for this layout,

$$\frac{I_O}{I_{REF}} = \frac{40 - .2}{20 - .2} = 2.01 \text{ (a 1\% error due to poor layout)}$$

Figure 20.9b shows how to lay out a current mirror to avoid these problems. The layout of M2 is two MOSFETs in parallel. This can be specified in SPICE by adding M = X after the MOSFET statement in the netlist, where X is the number of MOSFETs,

```
M1    Vd1   Vd1   0    0    NMOS    L=2   W=20
M2    Vd2   Vd1   0    0    NMOS    L=2   W=20   M=2
```

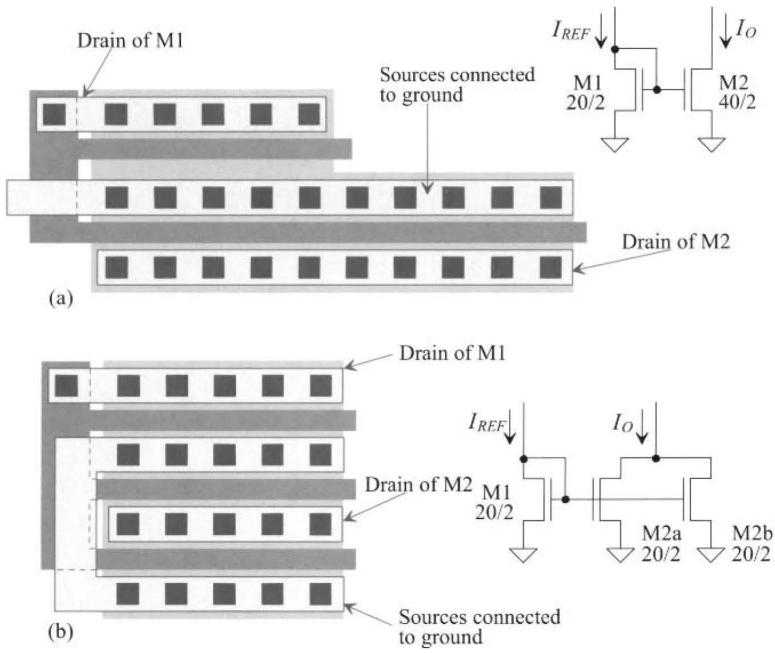


Figure 20.9 Layout of a current mirror (a) without width correction and (b) with width correction.

IPR2023-00698
Theta EX2019

20.1.3 Biasing the Current Mirror

Using a resistor to set the bias current, as seen in Figs. 20.2–20.4, can result in currents that are too dependent on the power supply value and temperature. Consider the current mirror seen in Fig. 20.10a. In this design we've used the sizes and bias current given in Table 9.2 (the short-channel CMOS process) to select the resistor. In particular, $V_{SG} = 0.35\ V$ so the gate potential is 0.65 *V*. Figure 20.10b shows how the reference and output currents vary if *VDD* is swept from 900 mV to 1 V. The reference current is linearly dependent on *VDD* (as seen in Ex. 20.1). The output current is dependent on both the reference current and the V_{DS} (λ) of M2. In general, we want a reference current that isn't dependent on power supply or ground variations (noise). This is an important point. Consider the following example.

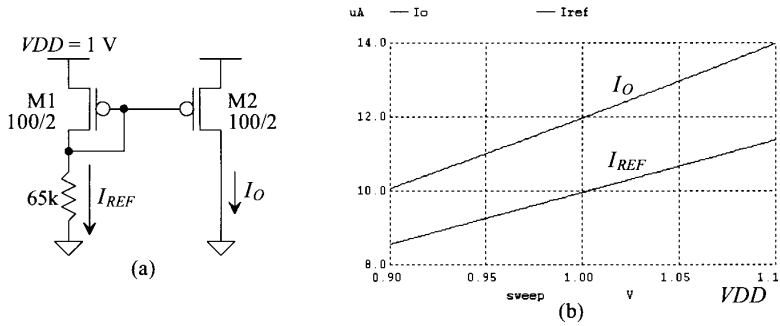


Figure 20.10 How reference and output current vary with VDD.

Example 20.2

Regenerate Fig. 20.10 if the 65k resistor is replaced with an ideal 10 μA current source, Fig. 20.11a. Explain the results.

The simulation results are seen in Fig. 20.11b. Note how the output current variation with *VDD* is much better than what is seen in Fig. 20.10. The reference current through M1 is, of course, a constant. The output current is larger than the

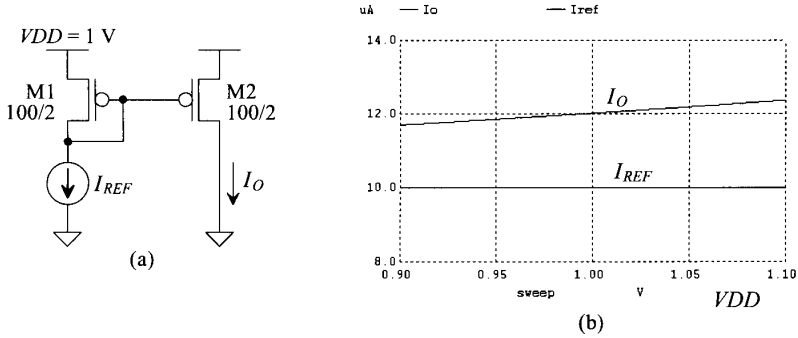


Figure 20.11 Showing that VDD variations don't affect the reference current.

IPR2023-00698
Theta EX2019

reference current because $V_{SD1} < V_{SD2}$ (notice how often this point is coming up in the design of current mirrors). The reason I_O varies is due to the finite output resistance ($\lambda \neq 0$) of M2, Eq. (20.2).

The point of this example is that it is desirable to have a bias current that doesn't vary with changes in *VDD* or ground. ■

Using a MOSFET-Only Reference Circuit

Let's try replacing the resistor bias with a MOSFET, Fig. 20.12. If we assume long-channel behavior, then

$$VDD = V_{SG3} + V_{GS1} \tag{20.12}$$

or

$$VDD = \sqrt{\frac{2I_{REF}}{KP_p \frac{W_3}{L_3}}} + V_{THP} + \sqrt{\frac{2I_{REF}}{KP_n \frac{W_1}{L_1}}} + V_{THN} \tag{20.13}$$

If we are going to think of M3 as a resistor, as in Fig. 20.2, so that we mirror the current in M1, then we solve for the size of M3. Using the data from Table 9.1, we get

$$5 = \sqrt{\frac{2 \cdot 20}{40 \cdot \frac{W_3}{L_3}}} + 0.9 + \sqrt{\frac{2 \cdot 20}{120 \cdot \frac{10}{2}}} + 0.8 \rightarrow \frac{W_3}{L_3} = 0.11 \approx \frac{10}{90} \tag{20.14}$$

If we are going to think of M1 as a resistor, as in Fig. 20.10a, so that we mirror the current in M3, then we solve for the size of M1 as we did in Eq. (20.14) and get

$$\frac{W_1}{L_1} \approx \frac{10}{270} \tag{20.15}$$

Figure 20.13 shows the operation of these MOSFET bias circuits with *VDD* swept from 4.5 to 5.5 V. Note that, at 5 V, we should have a bias current of 20 μA. What we get is 14 μA. The reason that we have such a large difference is due to neglecting the output resistance of the MOSFETs (and the effects of mobility degradation). Note that the sensitivities, how the output current changes with *VDD*, are roughly 8 μA/V. (Let's compare this to hand-calculated values of sensitivity next.)

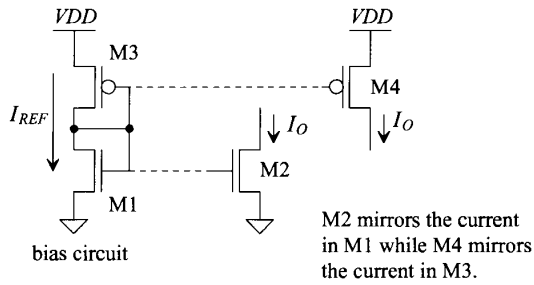


Figure 20.12 A MOSFET-only bias circuit.

IPR2023-00698
Theta EX2019

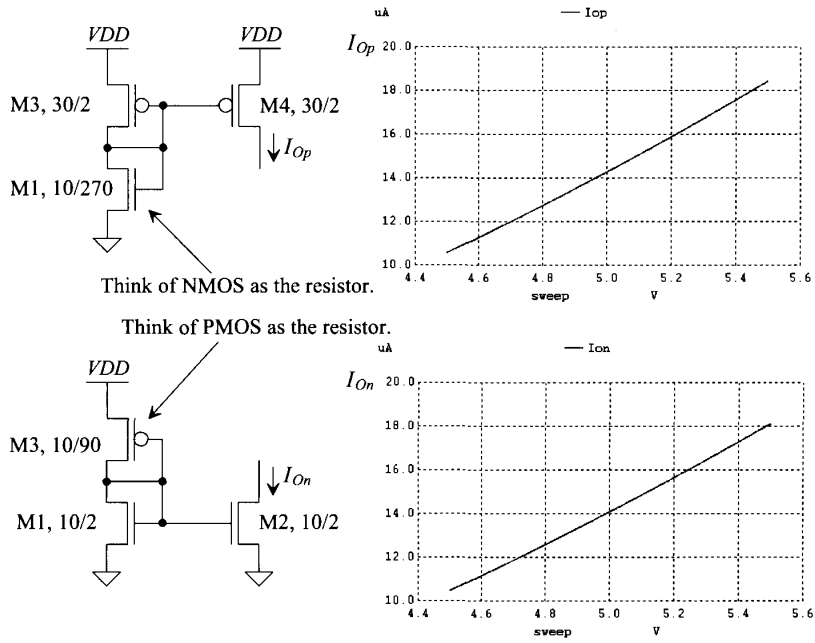


Figure 20.13 Behavior of MOSFET-only bias circuits with changes in VDD.

To determine I_{REF}'s sensitivity to *VDD*, let's take the derivative of the I_{REF} in Eq. (20.13) with respect to *VDD*. First let's write

$$(VDD - V_{THN} - V_{THP})^2 = I_{REF} \cdot \overbrace{\left[\sqrt{\frac{2L_3}{KP_p \cdot W_3}} + \sqrt{\frac{2L_1}{KP_n \cdot W_1}}\right]^2}^{K} \tag{20.16}$$

and next

$$\frac{\partial I_{REF}}{\partial VDD} = \frac{2 \cdot VDD}{K} - \frac{2 \cdot (V_{THN} + V_{THP})}{K} \tag{20.17}$$

Using the values in Table 9.1 and Eqs. (20.14) or (20.15), K is approximately 547×10^3 V^2/A so

$$\frac{\partial I_{REF}}{\partial VDD} = 12\ \mu A/V \tag{20.18}$$

For every millivolt change in *VDD* (around *VDD* = 5 V), we get 12 nA change in I_{REF}. As a comparison from Ex. 20.1, we can write (assuming the change in V_{GS} with *VDD* small)

$$\frac{\partial I_{REF}}{\partial VDD} \approx \frac{1}{R} = \frac{1}{200k} = 5\ \mu A/V \tag{20.19}$$

For every millivolt change in *VDD* we get 5 nA change in bias current.

IPR2023-00698
Theta EX2019

Supply Independent Biasing

Instead of putting the resistor in the drain side of the current mirror, let's consider placing it in the source side, Fig. 20.14a. One of the problems with this approach is that if we try to mirror the current in M2 using M5 we won't know the value of the mirrored current. It's not obvious how V_{GS2} and V_{GS5} are related. In (b) we add a diode connected M1 so that its current can be mirrored (by M5). The next question is how do we force the same current through both M1 and M2? Figure 20.14c shows the addition of a PMOS current mirror to do this. From (c) we can write

$$V_{GS1} = V_{GS2} + I_{REF} \cdot R \tag{20.20}$$

which can only be valid if $V_{GS1} > V_{GS2}$. To ensure that this is the case, we use a larger value of β in M2, that is, we *multiply-up* M1's β in M2 so that less gate-source voltage is needed to conduct I_{REF}. Generally, this is done by simply using a larger width in M2. The resulting circuit is called a *Beta-multiplier* reference circuit. Knowing

$$V_{GS} = \sqrt{\frac{2I_D}{\beta}} + V_{THN} \tag{20.21}$$

$\left(\beta = KP_n \cdot \frac{W}{L}\right)$ and

$$\beta_2 = K \cdot \beta_1 \text{ (which is satisfied by } W_2 = K \cdot W_1) \tag{20.22}$$

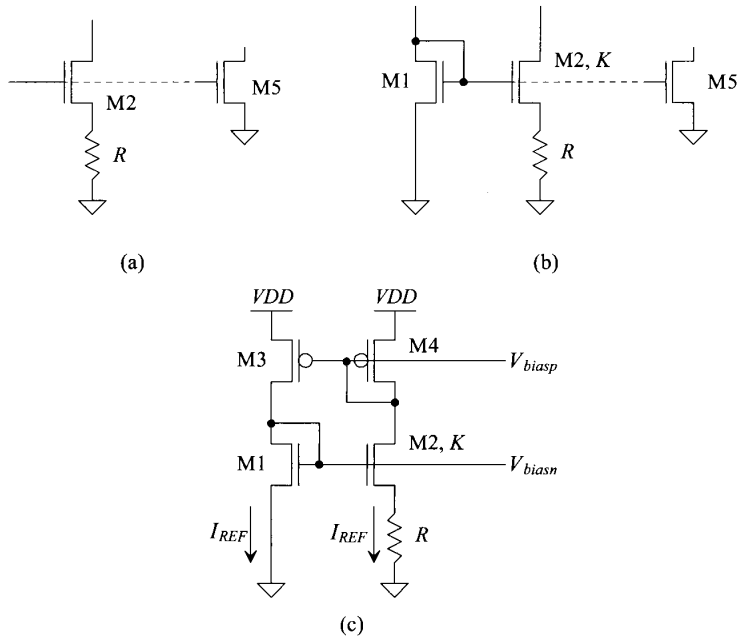


Figure 20.14 Developing the Beta-multiplier reference.

IPR2023-00698
Theta EX2019

we can write

$$I_{REF} = \frac{2}{R^2 KP_n \cdot \frac{W_1}{L_1}}\left(1 - \frac{1}{\sqrt{K}}\right)^2 \text{ or } V_{DS,sat} = V_{GS} - V_{THN} = \frac{2}{R \cdot KP_n \cdot \frac{W_1}{L_1}}\left(1 - \frac{1}{\sqrt{K}}\right) \quad (20.23)$$

It's important to note I_{REF} and V_{ovn} ($= V_{GS} - V_{THN}$), neglecting the finite output resistance of the MOSFETs which wasn't included in the derivation of Eq. (20.23), are independent of *VDD*. Solving for R with the values given in Table 9.1 and a K of 4 gives $R = 6.5\ k\Omega$. Note that this circuit, when $K = 4$, is sometimes called a *constant-g_m* bias circuit because

$$g_m = \sqrt{2KP_n \frac{W}{L} \cdot I_{REF}} = \frac{1}{R} \quad (20.24)$$

a constant independent of MOSFET process shifts. Figure 20.15 shows the schematic of a bias circuit based on the sizes, bias currents, and $V_{DS,sat}$ given in Table 9.1.

Note that a "start-up circuit" was included in Fig. 20.15. In any self-biased circuit there are two possible operating points: the one we just described and the unwanted one where zero current flows in the circuit. This unwanted state occurs when the gates of M1/M2 are at ground while the gates of M3/M4 are at *VDD*. When in this state, the gate of MSU1 is at ground and so it is off. The gate of MSU2 is somewhere between *VDD* and $VDD - V_{THP}$. MSU3, which behaves like an NMOS switch, turns on and leaks current into the gates of M1/M2 from the gates of M3/M4. This causes the current to snap to the desired state and MSU3 to turn off. *Note that during normal operation the start-up circuit should not affect the Beta-multiplier's operation.* The current through MSU3 should be zero (or very small).

The Beta-multiplier is an example of a circuit that uses positive feedback. The addition of the resistor kills the closed loop gain (a positive feedback system can be stable if its closed loop gain is less than one). However, if we decrease the size of the resistor, we increase the gain of the loop and push the feedback system closer to instability. An example of when this could occur is if the parasitic capacitance on the source of M2 to ground is large (effectively shorting M2's source to ground). If the resistor, for example, is bonded out off-chip to set the current, it is likely that this bias circuit will oscillate.

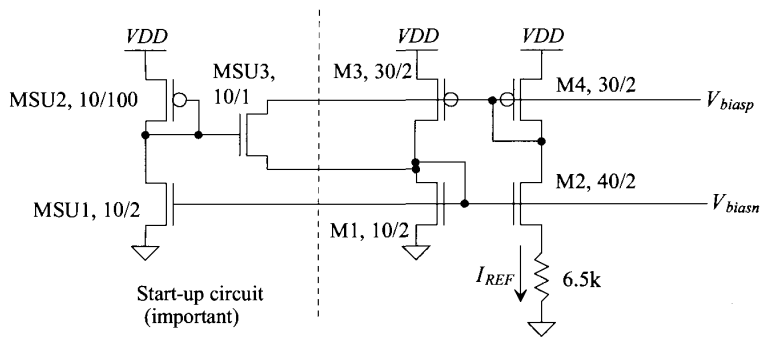


Figure 20.15 Beta-multiplier reference for biasing in the long-channel process described in Table 9.1.

IPR2023-00698
Theta EX2019

Figure 20.16 shows how the reference currents through M1 and M2 vary with *VDD*. The minimum value of *VDD* can be estimated by looking at the minimum value of voltage across the drain-source of M3 and M1. For M3 this is $V_{SD3,sat}$ or 250 mV (because we are using the parameters from Table 9.1 in the design of this current reference). For M1 this is $V_{DS1} = V_{GS1} = 1.05$ V. We can then write

$$VDD_{\min} = V_{SD3,sat} + V_{GS1} = 1.3\ V \tag{20.25}$$

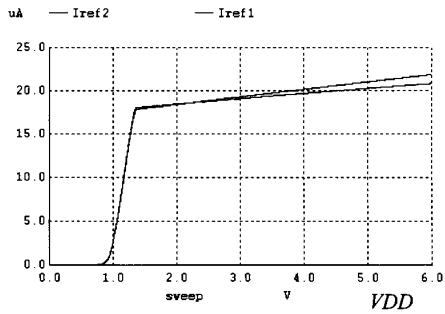


Figure 20.16 The reference currents through M1 and M2 in the Beta-multiplier.

Finally, the sensitivity of I_{REF} is directly dependent on the output resistance of the MOSFETs. From the simulations

$$\frac{\partial I_{REF}}{\partial VDD} \approx 800\ nA/V$$

or almost an order of magnitude better than the previous reference circuits

Example 20.3

Estimate the voltage on the gate of M6 and its drain current in Fig. 20.17.

M3 and M4 are biased to source 20 μA of current. Since $V_{GS1} = V_{GS2} = V_{DS1}$ (and M1/M2 have the same drain current, see Fig. 20.1 and the associated discussion), it follows that $V_{GS6} = V_{DS2} = V_{GS1} = 1.05$ V. *We treat M6, for biasing purposes, as if its gate were tied to the gates of M1 or M2.* It follows then that $I_{D6} = 20$ μA. ■

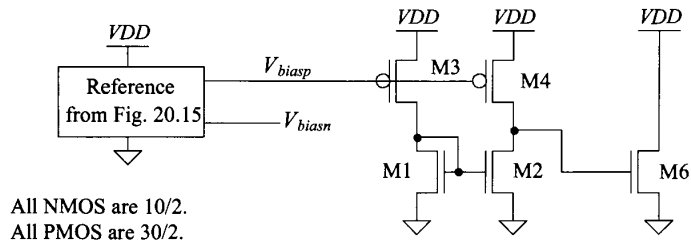


Figure 20.17 Circuit used in Ex. 20.3

IPR2023-00698
Theta EX2019

20.1.4 Short-Channel Design

Figure 20.18 shows a Beta-multiplier biasing circuit based on the values given in Table 9.2 as well as simulation results showing how the reference currents vary with *VDD*. From Eq. (20.25) and Table 9.2 we would expect VDD_{min} to be 400 mV. At *VDD* = 1 V the reference currents are indeed 10 μA. However, what we see is a horrible sensitivity to changes in *VDD*. Reviewing Figs. 9.31 and 20.10 we see the low output resistance of the short-channel devices causes the drain current to change significantly with changes in drain-to-source voltage. In some analog applications this variation isn't that harmful. However, if the Beta-multiplier circuit is to behave like a true current reference, the reference currents shouldn't vary with changes in *VDD*.

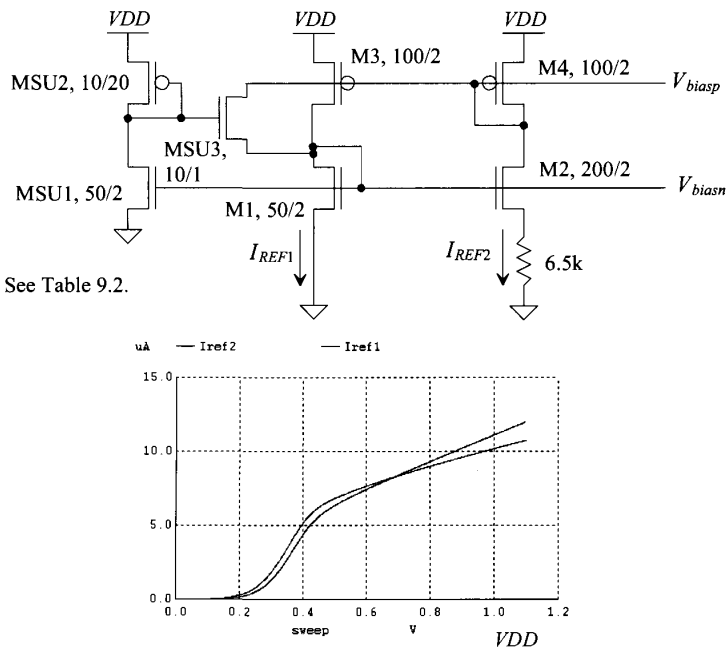


Figure 20.18 Beta-multiplier reference for short-channel design (see Table 9.2).

To reduce the sensitivity, we need to reduce the variations in the drain-to-source voltages of the NMOS devices with changes in *VDD*. Consider adding a differential amplifier (diff-amp) to the basic Beta-multiplier seen in Fig. 20.19. Note that M4 is no longer diode-connected so its drain can move to the same potential as M2's drain, that is, V_{biasn}. The start-up circuit (required) is not shown. The idea is to use the amplifier to compare the drain voltage of M1 (V_{biasn}) with the drain voltage of M2 (V_{reg}) and regulate them to be equal. The result is an effective increase in M2's output resistance. For example, if V_{reg} is above V_{biasn} , the amplifier's output voltage increases. This drives the

IPR2023-00698
Theta EX2019

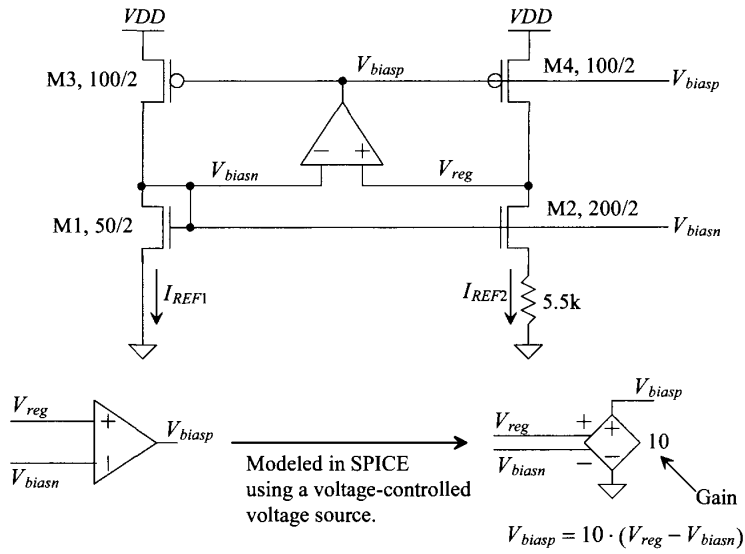


Figure 20.19 Increasing the output resistance of short-channel MOSFETs using feedback. The result, for the Beta-multiplier circuit, is better power supply sensitivity.

gate of M4 upwards, lowering the current it supplies and causing V_{reg} to drop back down. At the same time the gate of M3 is also increased, causing it to source less current. This causes a drop in V_{biasn} (the same as V_{reg} because of the symmetry as discussed in Fig. 20.1 or Ex. 20.3). Figure 20.20 shows how the reference current, in Fig. 20.19, changes with *VDD* when the added amplifier has a gain of 10. We've lowered *R* to 5.5k to more accurately set the current and used, in the simulation, a voltage-controlled voltage source for the amplifier.

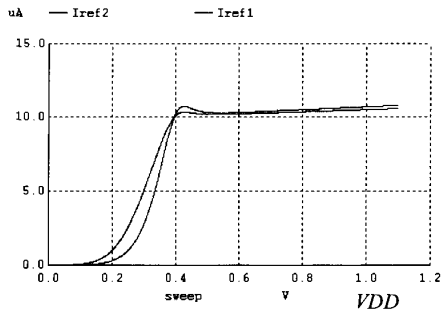


Figure 20.20 Improvement with the added amplifier.

IPR2023-00698
Theta EX2019

Figure 20.21 shows a possible differential amplifier configuration. The PMOS devices form a current mirror while the NMOS devices will, when $V_{biasn} = V_{reg}$, simply mirror the current in M1. If $V_{biasn} \neq V_{reg}$ then the imbalance causes the amplifier output to swing up or down providing the desired action.

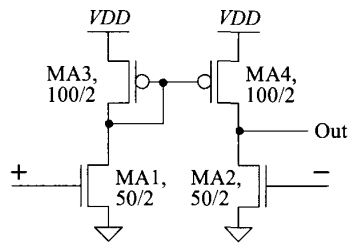
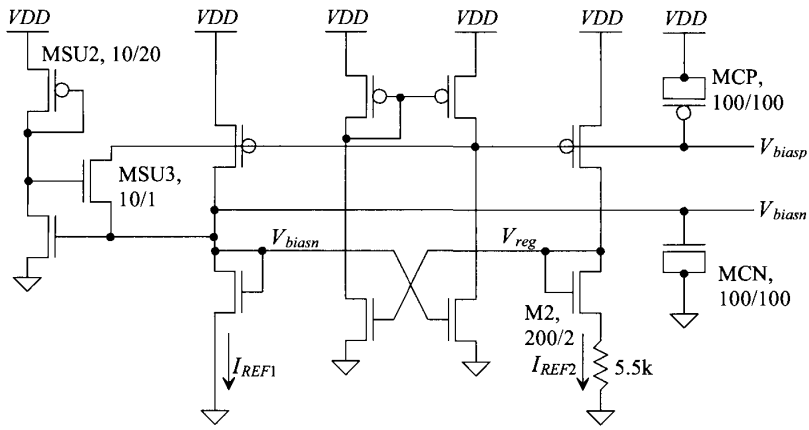


Figure 20.21 A possible implementation of the diff-amp in Fig. 20.19.

Figure 20.22 shows the improved current reference for use in a short-channel CMOS process (where devices have very low output resistances). The amplifier in Fig. 20.21 is placed in the reference, as indicated in Fig. 20.19. As with any feedback scheme, stability is of critical importance. As we'll see in the coming chapters, a feedback amplifier with only a single high-impedance node (meaning no diode-connected MOSFETs or MOSFET-source terminals are connected to the node) is straightforward to compensate (make stable). So that the reference has a single high-impedance node, that is V_{biasp}, we connected M2's gate to its drain (instead of to V_{biasn}). We get the same effect, as



All unlabeled NMOS are 50/2.
All unlabeled PMOS are 100/2.

Figure 20.22 Improved current reference for short-channel devices. See also Sec. 23.1.3 for more information.

IPR2023-00698
Theta EX2019

seen in Fig. 20.19 but the gain around the loop is reduced. To make the reference stable, we add capacitors (MCP and MCN) to the circuit. The high-impedance node is the critical point for adding the compensation capacitor (that is MCP is the critical capacitance). If the reference drives a significant number of MOSFETs (so these driven MOSFETs provide the load capacitance), then MCP and MCN may be excluded from the design. Figure 20.23 shows how the reference currents change with *VDD*.

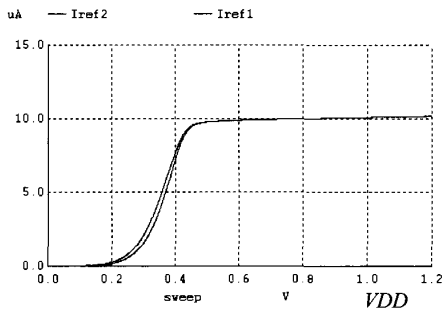
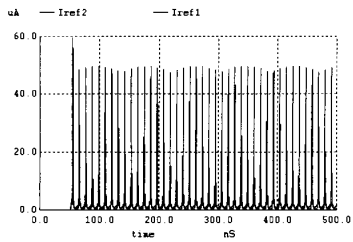


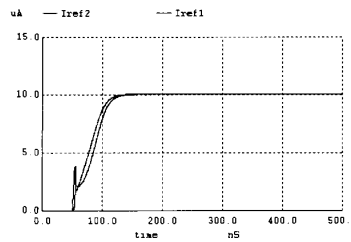
Figure 20.23 Variation of reference currents with VDD for the circuit in Fig. 20.22.

An Important Note

It's extremely important to understand that together with the benefits of better power supply sensitivity are the undesired traits of feedback, that is, the potential for an unstable circuit. For example, Fig. 20.24a shows what happens to the reference currents if we remove MCP and MCN in Fig. 20.22 and then apply a step voltage to *VDD* (that is, *VDD* steps from 0 to 1 V at 50 ns in the simulation). Clearly, the reference is not stable and the currents oscillate. In (b) we do the same thing but with MCP and MCN present. The response shows first-order behavior, and oscillations are not present. We might think we are done with the reference and move on to the other analog circuit designs. However, further characterization of the current reference using simulations is warranted. For example, what happens if *VDD* has a 50 mV squarewave signal coupled to it at 100 MHz? How does the reference behave with this signal (a squarewave *VDD* that oscillates between 1 and 1.05 V at 100 MHz)? In general, no analog circuit has good power supply noise rejection at high frequencies. To remedy this, the power supply is decoupled (a large capacitor is placed from *VDD* to ground to remove high-frequency noise).



(a) MCP and MCN not present



(b) MCP and MCN present

Figure 20.24 What happens when VDD is pulsed from 0 to 1 at 50 ns.

IPR2023-00698
Theta EX2019

20.1.5 Temperature Behavior

While we've been focusing on how the reference current changes with *VDD,* it's also important to know how the current changes with temperature. Looking at the basic current mirror in Fig. 20.11a, note that if I_{REF} is a constant, independent of temperature, then the output current will also be independent of temperature. That's not to say that the MOSFET characteristics aren't changing because they are; they change at the same rate. Because they change together, the current mirror relationship is still valid and M2 mirrors the current in M1. Of course, if M1 is located at a different physical location than M2 and each is heated differently, a mismatch in the currents will result. *The point is that the temperature behavior of the reference current determines the temperature behavior of all mirrored currents.*

Using the short-channel CMOS parameters we know, from Ch. 9 (Fig. 9.30b), that the change in V_{THN} with temperature is –0.6 mV/C° (for an NMOS device). For the PMOS device we can vary temperature and look at the change in V_{THP}. Using the mirror in Fig. 20.11 and looking at the V_{SG} at different temperatures results in the plots seen in Fig. 20.25. We see that it also varies at a rate of –0.6 mV/C°.

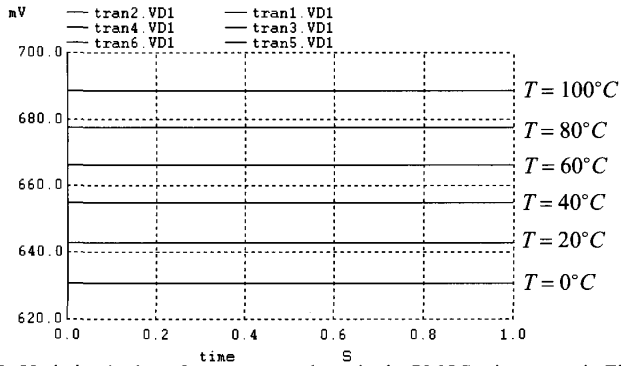


Figure 20.25 Variation in the reference gate voltage in the PMOS mirror seen in Fig. 20.11.

Resistor-MOSFET Reference Circuit

For a resistor-MOSFET reference circuit (Fig. 20.2), we can write (neglecting the MOSFET's finite output resistance)

$$I_{REF} = \frac{VDD - V_{GS}}{R} \tag{20.26}$$

The change in I_{REF} with temperature is

$$\frac{\partial I_{REF}}{\partial T} = -\frac{VDD - V_{GS}}{R^2} \cdot \frac{\partial R}{\partial T} - \frac{1}{R} \cdot \frac{\partial V_{GS}}{\partial T} \tag{20.27}$$

or

$$\frac{\partial I_{REF}}{\partial T} = -I_{REF} \cdot \frac{1}{R}\frac{\partial R}{\partial T} - \frac{1}{R} \cdot \frac{\partial V_{GS}}{\partial T} \tag{20.28}$$

IPR2023-00698
Theta EX2019

We know that to write the reference current as a function of temperature we use

$$I_{REF}(T) = I_{REF}(T_0) \cdot (1 + TCI_{REF} \cdot (T - T_0)) \tag{20.29}$$

where the temperature coefficient of the reference current is

$$TCI_{REF} = \frac{1}{I_{REF}} \cdot \frac{\partial I_{REF}}{\partial T} \tag{20.30}$$

Rewriting Eq. (20.28) using Eq. (20.26) yields

$$TCI_{REF} = \frac{1}{I_{REF}} \cdot \frac{\partial I_{REF}}{\partial T} = -\frac{1}{R}\frac{\partial R}{\partial T} - \frac{1}{VDD - V_{GS}}\frac{\partial V_{GS}}{\partial T} \tag{20.31}$$

Noting, again, that the temperature coefficient is not a constant but rather changes with temperature (even though this isn't indicated in Eq. (20.29)) just like a small-signal parameter, e.g., g_m, changes with the DC operating point.

Example 20.4

Determine the temperature behavior of the reference current in the mirror seen in Fig. 20.10. Verify the answer with SPICE.

If the temperature coefficient of the resistor, $TCR = \frac{1}{R}\frac{\partial R}{\partial T} = 2000\ ppm/C^\circ$ (= 0.002) and we assume after looking at Fig. 9.30 that $\frac{\partial V_{GS}}{\partial T} \approx \frac{\partial V_{THN}}{\partial T} \approx -0.6\ mV/C^\circ$ then, knowing the parameters used in Table 9.2 were the basis for the design of the current mirror in Fig. 20.10,

$$TCI_{REF} = -0.002/C^\circ - \frac{1}{1 - 0.35} \cdot (-0.6\ mV/C^\circ) \approx -1,000\ ppm/C^\circ$$

and so

$$I_{REF}(T) = 10 \cdot (1 - 0.001 \cdot (T - 27))\ \mu A$$

At 100 C° the reference current is 9.27 μA and at 0 C° it's 10.27 μA (of course, the reference current is approximately 10 μA at 27 C°). Figure 20.26 shows the SPICE simulation results. ■

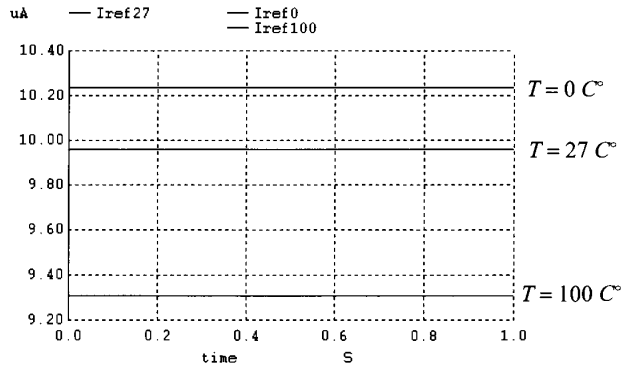


Figure 20.26 Example 20.4 showing the temperature behavior of the reference current in Fig. 20.10.

IPR2023-00698
Theta EX2019

MOSFET-Only Reference Circuit

For the MOSFET-only bias circuit, Fig. 20.12, we can write the reference current as (see Eq. [20.16])

$$I_{REF} = \frac{VDD^2 - 2VDD \cdot (V_{THN} + V_{THP}) + (V_{THN} + V_{THP})^2}{\left(\sqrt{\frac{2L_3}{KP_p \cdot W_3}} + \sqrt{\frac{2L_1}{KP_n \cdot W_1}}\right)^2} \tag{20.32}$$

Knowing $KP_n = M \cdot KP_p$, where M is 3 for our long-channel process and 2 for our short-channel process (see Tables 9.1 and 9.2), we can write

$$I_{REF} = KP_n \cdot \frac{VDD^2 - 2VDD \cdot (V_{THN} + V_{THP}) + (V_{THN} + V_{THP})^2}{\left(\sqrt{\frac{M \cdot 2L_3}{W_3}} + \sqrt{\frac{2L_1}{W_1}}\right)^2} \tag{20.33}$$

or

$$\frac{\partial I_{REF}}{\partial T} = \frac{\partial KP_n}{\partial T} \cdot \frac{VDD^2 - 2VDD \cdot (V_{THN} + V_{THP}) + (V_{THN} + V_{THP})^2}{\left(\sqrt{\frac{M \cdot 2L_3}{W_3}} + \sqrt{\frac{2L_1}{W_1}}\right)^2} +$$

$$\frac{KP_n}{\left(\sqrt{\frac{M \cdot 2L_3}{W_3}} + \sqrt{\frac{2L_1}{W_1}}\right)^2} \cdot \left(-2VDD \cdot \left(\frac{\partial V_{THN}}{\partial T} + \frac{\partial V_{THP}}{\partial T}\right) + 2(V_{THN} + V_{THP})\left(\frac{\partial V_{THN}}{\partial T} + \frac{\partial V_{THP}}{\partial T}\right)\right) \tag{20.34}$$

Dividing both sides by Eq. (20.33) gives

$$\frac{1}{I_{REF}} \frac{\partial I_{REF}}{\partial T} = \frac{1}{KP_n} \frac{\partial KP_n}{\partial T} + \frac{2\left(\frac{\partial V_{THN}}{\partial T} + \frac{\partial V_{THP}}{\partial T}\right)(V_{THN} + V_{THP} - VDD)}{VDD^2 - 2VDD \cdot (V_{THN} + V_{THP}) + (V_{THN} + V_{THP})^2} \tag{20.35}$$

noting, from Eq. (9.52), that the first term is simply −1.5/T.

Example 20.5

Determine the temperature behavior of the reference current in the MOSFET-only bias circuit seen in Fig. 20.13.

We note that both references in Fig. 20.13 have the same temperature behavior since Eq. (20.35) doesn't show a width or length dependence (there is a dependence in the reference current though). Figure 20.13 used the long-channel MOSFET parameters from Table 9.1 where

$$V_{THN} = 0.8,\ V_{THP} = 0.9,\ \frac{\partial V_{THN}}{\partial T} = -1\ mV/C^\circ,\ \frac{\partial V_{THP}}{\partial T} = -1.4\ mV/C^\circ$$

and VDD = 5 V with I_{REF} = 13.5 μA (measured in the simulation that generated Fig. 20.13) at T = 300 K°. Plugging the numbers into Eq. (20.35) gives

$$TCI_{REF} = \frac{1}{I_{REF}} \frac{\partial I_{REF}}{\partial T} = -\frac{1.5}{300} + \frac{2(-0.0024)(-3.3)}{25 - 17 + 2.89} = -0.005 + 0.00146 \approx -3,500\ ppm/C^\circ$$

and so $I_{REF}(T) = 13.5(1 - 0.0035(T - T_0))\ \mu A$. ■

IPR2023-00698
Theta EX2019

Temperature Behavior of the Beta-Multiplier

To determine the temperature behavior of the Beta-multiplier, we take the derivative of Eq. (20.23) with respect to temperature

$$\frac{\partial I_{REF}}{\partial T} = \frac{-4}{R^3 KP_n \cdot \frac{W_1}{L_1}}\left(1-\frac{1}{\sqrt{K}}\right)^2 \cdot \frac{\partial R}{\partial T} - \frac{2}{R^2 KP_n^2 \cdot \frac{W_1}{L_1}}\left(1-\frac{1}{\sqrt{K}}\right)^2 \frac{\partial KP_n}{\partial T} \quad (20.36)$$

Dividing both sides by I_{REF} (Eq. [20.23]) gives

$$TCI_{REF} = \frac{1}{I_{REF}} \cdot \frac{\partial I_{REF}}{\partial T} = -2\left(\frac{1}{R}\frac{\partial R}{\partial T}\right) - \frac{1}{KP_n}\frac{\partial KP_n}{\partial T} \quad (20.37)$$

Example 20.6

Determine the temperature behavior of the Beta-multiplier seen in Fig. 20.22.

If the temperature coefficient of the resistor, $TCR = \frac{1}{R}\frac{\partial R}{\partial T} = 2000\ ppm/C^\circ$ then

$$TCI_{REF} = -0.004 + \frac{1.5}{300} = 1000\ ppm/C^\circ$$

and so

$$I_{REF}(T) = 10 \cdot (1 + 0.001 \cdot (T - 27))\ \mu A$$

The positive TC of the resistor subtracts from the negative TC of the mobility to stabilize the current reference. If the mobility change with temperature is complicated by velocity saturation effects, then simulations become invaluable to determine the actual temperature performance of the circuit. ■

Voltage Reference Using the Beta-Multiplier

Let's try designing the Beta-multiplier where the gate voltage of M1 (see Figs. 20.15 or 20.22) is a constant independent of temperature. In this type of design we aren't using this voltage directly to bias anything so we'll change the label from V_{biasn} to V_{REF}. Using Eqs. (20.22) and (20.23) gives

$$V_{REF} = V_{GS1} = \frac{2}{R \cdot KP_n \cdot \frac{W}{L}}\left(1-\frac{1}{\sqrt{K}}\right) + V_{THN} \quad (20.38)$$

Taking the derivative with respect to temperature gives

$$\frac{\partial V_{REF}}{\partial T} = \frac{\partial V_{THN}}{\partial T} - \frac{2}{R \cdot KP_n \cdot \frac{W}{L}}\left(1-\frac{1}{\sqrt{K}}\right) \cdot \left(\frac{1}{R}\frac{\partial R}{\partial T} + \frac{1}{KP_n} \cdot \frac{\partial KP_n}{\partial T}\right) \quad (20.39)$$

If we want the change with temperature to be zero, $\partial V_{REF}/\partial T = 0$, then we can select the resistor based on

$$R = \frac{2}{\frac{\partial V_{THN}}{\partial T} \cdot KP_n \cdot \frac{W}{L}}\left(1-\frac{1}{\sqrt{K}}\right) \cdot \left(\frac{1}{R}\frac{\partial R}{\partial T} + \frac{1}{KP_n} \cdot \frac{\partial KP_n}{\partial T}\right) \quad (20.40)$$

Using the parameters from Table 9.1 with $K = 4$ and a resistor temperature coefficient of 0.002 gives a resistor value of 5 kΩ. Figure 20.27 shows the SPICE simulation results. The variation is approximately 60 mV/100 C° or 600 μV/C°.

IPR2023-00698
Theta EX2019

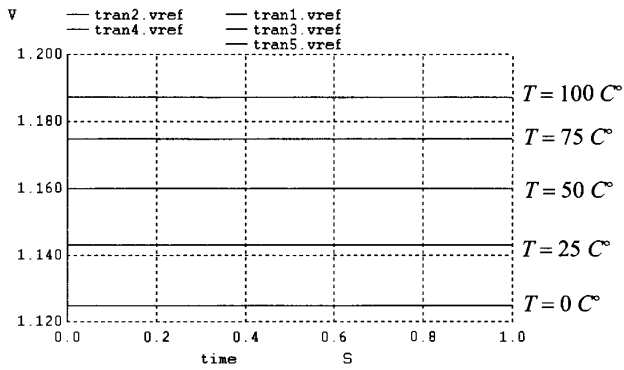


Figure 20.27 Temperature performance of a voltage reference using the Beta-multiplier.

20.1.6 Biasing in the Subthreshold Region

In some CMOS designs (for example, low-power designs) the MOSFETs may be operated in the weak or subthreshold regions. If a resistor is used to bias a current mirror, as seen in Fig. 20.10, its value can be very large. For example, if we use the short-channel process with VDD = 1 V and a V_{SG} of approximately V_{THP} (= 280 mV) with an I_D of 10 nA, we need a resistor with a value of

$$I_{REF} = \frac{VDD - V_{SG}}{R} \rightarrow R = 72\ M\Omega\ ! \tag{20.41}$$

Clearly, this isn't practical. To bias circuits for subthreshold operation, let's use the Beta-multiplier. From Eq. (9.17) we can write

$$V_{GS1} = nV_T \cdot \ln\left(\frac{I_{REF} \cdot L_1}{I_{D0} \cdot W_1}\right) + V_{THN} \tag{20.42}$$

and

$$V_{GS2} = nV_T \cdot \ln\left(\frac{I_{REF} \cdot L_1}{I_{D0} \cdot K \cdot W_1}\right) + V_{THN} \tag{20.43}$$

Knowing

$$I_{REF} = \frac{V_{GS1} - V_{GS2}}{R} \tag{20.44}$$

we get

$$I_{REF} = \frac{n \cdot V_T}{R} \cdot \ln K \tag{20.45}$$

or

$$R = \frac{n \cdot V_T}{I_{REF}} \cdot \ln K \tag{20.46}$$

Using the values above, that is, $K = 4$ and an I_{REF} of 10 nA, gives an R (assuming $n = 1$) of approximately 3.5 MΩ. Still large but much better than 72 MΩ.

IPR2023-00698
Theta EX2019

20.2 Cascoding the Current Mirror

In this section we discuss increasing the output resistance of a current mirror so that it behaves more ideally. Figure 20.28 shows the problem with using a single MOSFET for a current source. Instead of being a constant, the output current, I_O, increases as the voltage across the current source, V_O, increases. We model this increase using the MOSFET's output resistance, r_o (see Table 9.2). If we can hold the drain-source voltage of the MOSFET constant, then the current doesn't vary. However, this requires a fixed V_O. To avoid this, we'll add circuitry in between the current-mirrored MOSFET and V_O.

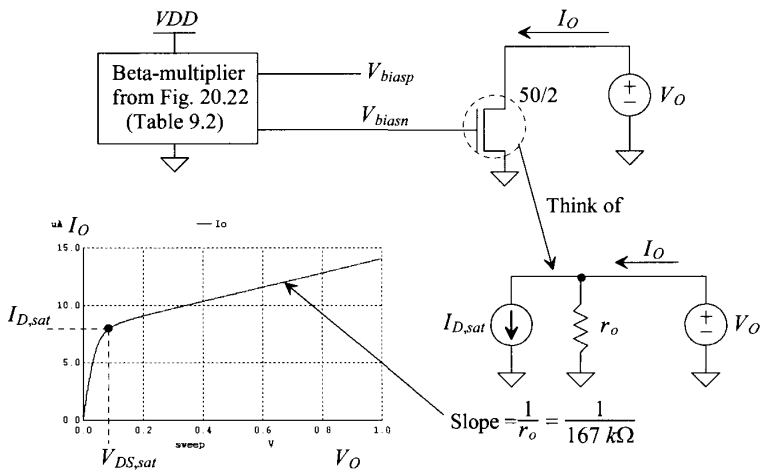


Figure 20.28 How the finite output resistance of the MOSFET affects the output current.

20.2.1 The Simple Cascode

The simple cascode current mirror made with M1–M4 is seen in Fig. 20.29. The name "cascode" is a vestige from the days of vacuum tubes when a common-cathode amplifier was cascaded (in series with) a common-grid amplifier. Drawing an analogy with MOSFETs, the cathode of a tube is equivalent to the source of a MOSFET. The grid of a tube corresponds to the gate of a MOSFET and the anode to the drain of the MOSFET.

It's important to realize that I_O is still determined by the gate-source voltages of M1 and M2. Changing the sizes of M3 and M4 simply changes the drain-source voltages of M1/M2, affecting the matching of their drain currents, as indicated by Eq. (20.10). Again, this (the gate-source voltages of M1/M2 determine the currents) is important to understand since, in the next several pages, we present different modifications to this basic cascode circuit. These modifications attempt to hold the drain source voltages of M1/M2 more constant to increase the current mirror's output resistance (make I_O change less with changes in V_O).

IPR2023-00698
Theta EX2019

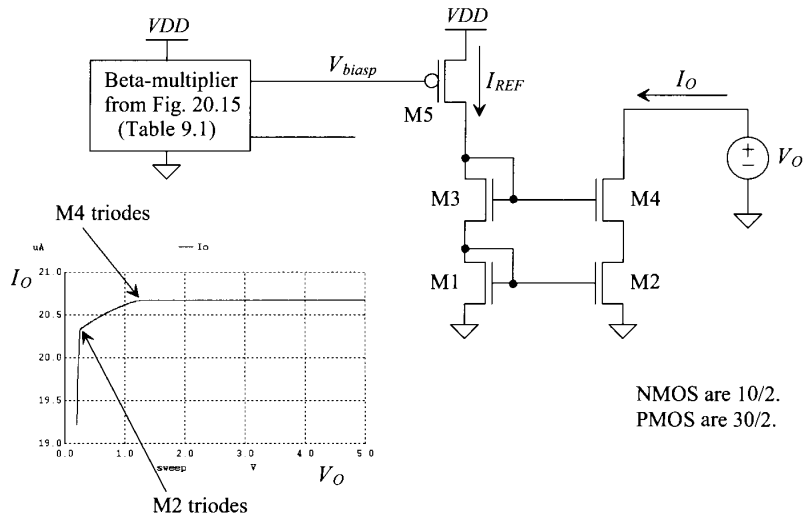


Figure 20.29 Biasing of the cascode current source and its operation.

DC Operation

The voltage on the gate of M4 in Fig. 20.29 is (remembering the actual values for these specific biasing conditions is given in Table 9.1)

$$2V_{GS} = 2(V_{DS,sat} + V_{THN}) \tag{20.47}$$

The voltage on the drain of M2, assuming that it is operating in the saturation region, is V_{GS}. The minimum voltage across the current source is then

$$V_{O,\min} = V_{DS,sat4} + V_{GS} = 2V_{DS,sat} + V_{THN} \tag{20.48}$$

Plugging in the numbers from Table 9.1, we get $V_{O,\min} = 1.3\ V$. Reviewing Fig. 20.29, we see that M4 starts to move into the triode region at this voltage.

To keep both M2 and M4 operating in the saturation region, we need only a $V_{DS,sat}$ across each one. In the next section we'll use this fact to design a *low-voltage cascode* current mirror.

Cascode Output Resistance

To estimate the output resistance of the cascode current mirror, let's apply a test voltage to the simplified cascode schematic seen in Fig. 20.30. We treat M1 and M3 as DC bias sources (AC grounds) and assume that the DC voltage on the drain of M4 is large enough to ensure that M2 and M4 are operating in the saturation. Further we draw the output resistances external to the devices. Note that M2's AC gate source voltage is zero and its drain voltage is $-v_{gs4}$. The resistance seen looking into the drain of M4 is

$$.R_o = \frac{v_T}{i_T} \tag{20.49}$$

IPR2023-00698
Theta EX2019

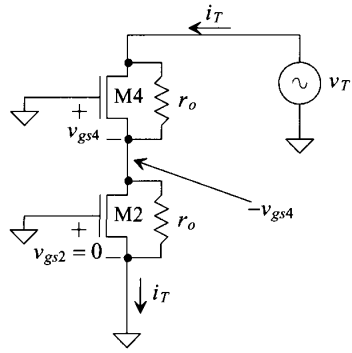


Figure 20.30 AC circuit used to determine the output resistance of a cascode current source.

Looking at Fig. 20.30, we can write

$$-v_{gs4} = i_T \cdot r_o \tag{20.50}$$

The current through M4 is then

$$i_T = g_m v_{gs4} + \frac{v_T - (-v_{gs4})}{r_o} \tag{20.51}$$

Substituting (20.50) into (20.51) gives

$$i_T = g_m(-i_T \cdot r_o) + \frac{v_T}{r_o} - i_T \tag{20.52}$$

Solving for the output resistance

$$R_o = (2 + g_m r_o) r_o \approx g_m r_o^2 \tag{20.53}$$

This result should be remembered because it will be referred to often. Using the numbers from Tables 9.1 and 9.2, we can tabulate the output resistance of cascode structures as seen in Table 20.1.

Table 20.1 Output resistances for cascode structures.

Cascode type	Long-channel process, Table 9.1, R_o	Short-channel process, Table 9.2, R_o
PMOS	2.4 GΩ	16.6 MΩ
NMOS	3.75 GΩ	4.2 MΩ

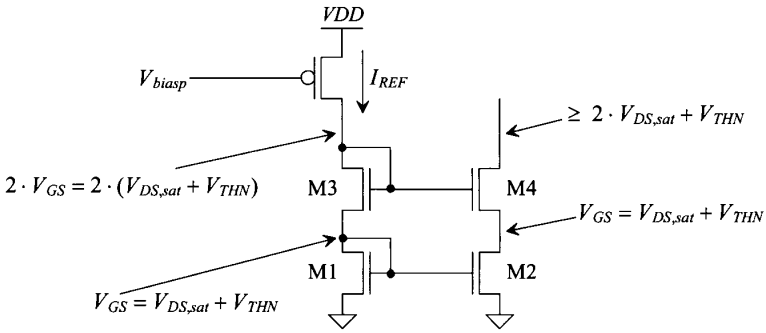
Notice the large difference (almost three orders of magnitude) between the output resistance of the short- and the long-channel devices. Again, this is the reason we need to use special circuit techniques when designing in short-channel processes.

Finally, note that the output resistance can be determined from a plot like the one seen in Fig. 20.29 (or Fig. 20.28) by plotting the reciprocal of the derivative of the output current.

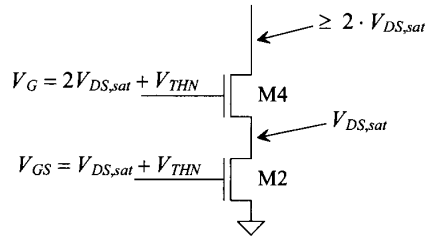
IPR2023-00698
Theta EX2019

20.2.2 Low-Voltage (Wide-Swing) Cascode

Figure 20.31a shows the regularly biased cascode structure. The key point of this figure is that the drain of M2 is not at the minimum voltage required to keep it in saturation, $V_{DS,sat}$, but rather it is held at a threshold voltage above this minimum. In (b) the cascode is biased for lowest voltage operation. What this means is that the drain of M4, V_o, can go to the minimum possible voltage that keeps M2/M4 in saturation, that is, $2V_{DS,sat}$. Again, it's important to notice that the gate-source voltages of M1/M2 are what set the currents in the mirror.



(a) Regular cascode structure



(b) Low-voltage (aka wide-swing) structure

Figure 20.31 DC voltages in (a) a cascode current mirror and in (b) a low-voltage cascode.

To generate the gate voltage of M2, we simply use a diode-connected MOSFET as seen in almost every schematic in this chapter. From Eq. (20.1), for example, we can write

$$I_{REF} = \frac{KP_n}{2} \cdot \frac{W}{L}(V_{GS} - V_{THN})^2 \qquad (20.54)$$

neglecting channel-length modulation. To generate the bias voltage for M4, we can adjust the width and length of a MOSFET, MWS, as seen in Fig. 20.32. We can write, knowing $V_{DS,sat} = V_{GS} - V_{THN}$

IPR2023-00698
Theta EX2019

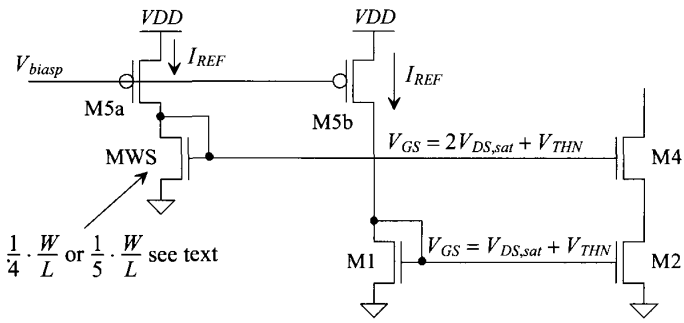


Figure 20.32 Generating a bias voltage for M4.

$$I_{REF} = \frac{KP_n}{2} \cdot \frac{W_{MWS}}{L_{MWS}} (2(V_{GS} - V_{THN}) + V_{THN} - V_{THN})^2 \qquad (20.55)$$

or

$$I_{REF} = \frac{KP_n}{2} \cdot \frac{W_{MWS}}{L_{MWS}} \cdot 4(V_{GS} - V_{THN})^2 \qquad (20.56)$$

Using our device sizes from either Tables 9.1 or 9.2, we can make the length of MWS four times the length of the other MOSFETs or

$$\frac{W}{L} = \frac{W_{MWS}}{L_{MWS}} \cdot 4 \qquad (20.57)$$

If we use the same widths, then

$$L_{MWS} = 4 \cdot L \qquad (20.58)$$

If Eq. (20.58) is valid, then Eq. (20.56) is equal to Eq. (20.54).

Note that, as seen in Fig. 20.31b, we are biasing M2 right on the edge of the triode region when we use Eq. (20.58) to size MWS. In many situations we may want to move M2 further into the saturation region. Intuitively, we would think that by increasing the length of MWS we increase its effective resistance and so the voltage drop across it goes up. If we increase MWS's length further, the gate voltage of M4 goes up and thus so does the drain voltage of M2. M2 is moved further away from the edge of the triode region into the saturation region. The minimum voltage across the current source increases. It's important to understand that it's OK for M2 to be biased away from the triode region by a small amount.

Example 20.7

Regenerate Fig. 20.29 using a wide-swing cascode structure.

The simulation results are seen in Fig. 20.33. It appears that the output resistance is really large until $V_O < V_{DS,sat}$. However, if we zoom in around $2V_{DS,sat}$ (= 500 mV here), we would see that the slope increases and M4 triodes. ■

IPR2023-00698
Theta EX2019

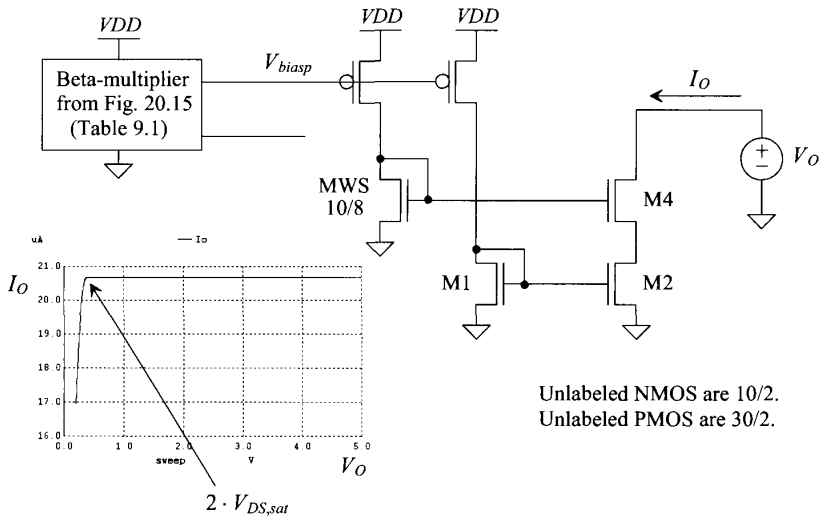


Figure 20.33 Wide-swing cascode current source in the long-channel process.

An Important Practical Note

In general we won't design a wide-swing current sink by setting the length to four times the length used in the rest of the design. As just mentioned, this biases M2 right on the edge of the triode region (point A) in Fig. 20.34. Stealing a current from the drain of M2 will move M2 into the triode region, point B. If we bias M2 further into the saturation region, point C, then more current must be removed before M2 triodes and the output resistance of the current source decreases. When we design folded-cascode amplifiers later, we will steal or add current in this way. For general *long-channel design,* we'll use

$$\frac{W_{MWS}}{L_{MWS}} = \frac{1}{5} \cdot \frac{W}{L} \tag{20.59}$$

or using the same widths then

$$L_{MWS} = 5 \cdot L \tag{20.60}$$

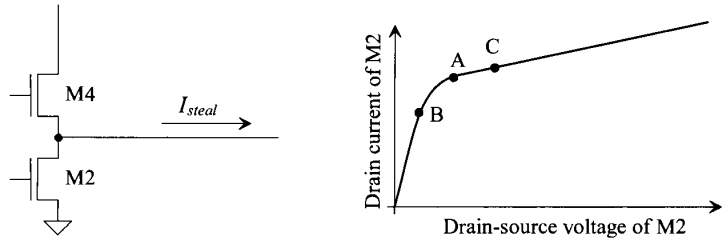


Figure 20.34 Showing how stealing current from M2 moves it into the triode region.

IPR2023-00698
Theta EX2019

Layout Concerns

When we go to layout the long L device seen in Fig. 20.32, we might simply layout a single MOSFET with the appropriate length. However, the threshold voltage can vary significantly with the length of the device. Figure 20.35 shows a method where MOSFETs connected in series with the same widths and their gates tied together behave like a single MOSFET with the sum of the individual MOSFET's lengths. Because each device is identical, changes in the threshold voltage shouldn't affect the biasing circuit.

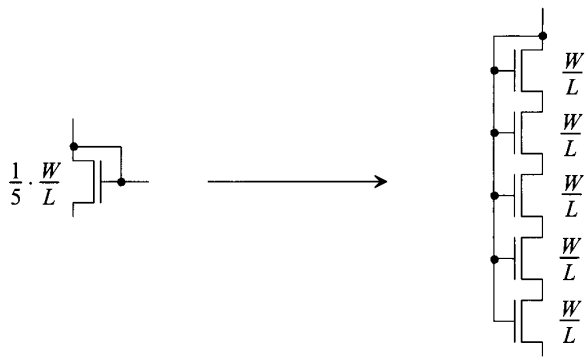


Figure 20.35 Using equal length devices to implement a long L MOSFET.

20.2.3 Wide-Swing, Short-Channel Design

When designing a wide-swing current mirror in a short-channel CMOS process, we know

$$V_{DS,sat} \neq V_{GS} - V_{THN} = V_{ovn} \tag{20.61}$$

Further, as seen in Fig. 9.32 in Ch. 9, the output resistance of a short-channel MOSFET depends on the drain-to-source voltage. Figure 20.36 shows a wide-swing current mirror designed in the short-channel CMOS process (see Table 9.2). Note that the current is approximately 4 μA when it should be, as seen in Fig. 20.23, 10 μA. Further, if we were to measure the output resistance, we would get approximately 600k, which is considerably less than the 4 MΩ seen in Table 20.1. As seen in Fig. 9.32, the output resistance at $V_{DS,sat}$ isn't 166k as indicated in Table 9.2 but much less. To increase the output resistance, let's attempt to bias M2 deeper into the saturation region. Let's do this by using a larger value of voltage on the gate of M4. For our short-channel design, let's try

$$\frac{W_{MWS}}{L_{MWS}} = \frac{1}{25} \cdot \frac{W}{L} \tag{20.62}$$

or a five times smaller ratio than what we used for the long-channel devices (Eq. [20.59]). The simulation results for the circuit in Fig. 20.36 are seen in Fig. 20.37 when MWS is sized 10/10. Notice that our output current is closer to 10 μA. Studying Fig. 20.36 and remembering our constant theme in this chapter that good matching requires both equal V_{GS} and V_{DS}, we see a concern. M1's drain-to-source voltage isn't the same as M2's. By

IPR2023-00698
Theta EX2019

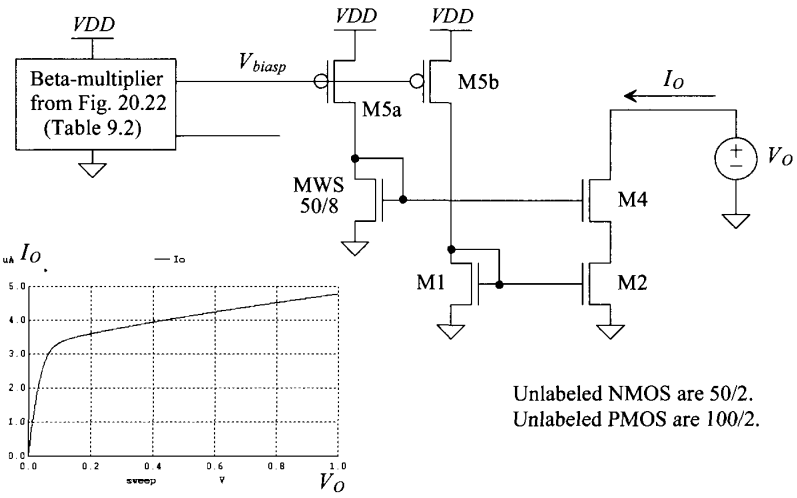
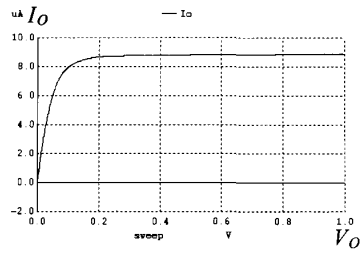


Figure 20.36 Wide-swing cascode current source in the short-channel process (bad).

making the gate voltage of M4 larger, we act to equalize the two drain-to-source voltages (we increase the drain voltage of M2) causing the output current to approach the current that flows in M1 (10 μA). Let's attempt to make the current mirrors more symmetrical in an effort to equilibrate the drain-to-source voltages of M1 and M2.

Figure 20.38 shows the addition of a MOSFET, M3, to lower M1's V_{DS} so that it matches M2's V_{DS}. The 10 μA current through M5b can now be mirrored accurately. M1's gate voltage increases until M1 (and thus M2) can sink the current supplied by M5b with the smaller V_{DS}. *For all wide-swing current mirrors the topology seen in Fig.* 20.38 *using M3 should be used.* Note also, in Fig. 20.38, the larger voltage needed across the current source to keep M2/M4 saturated. The drains of M1 and M2 are at 150 mV or considerably above $V_{DS,sat}$ (which is 50 mV from Fig. 9.32). M3 is moving close to the triode region. Equation (20.62) can be modified to move M3 away from triode by using a larger W/L device (say 1/10); however, the output resistance will decrease (again as seen in Fig. 9.32, the decrease in V_{DS} results in smaller output resistance, r_o).



MWS is 10/10 in Fig. 20.36.

Figure 20.37 Increasing the W/L of MWS to 1/25 the other W/Ls.

IPR2023-00698
Theta EX2019

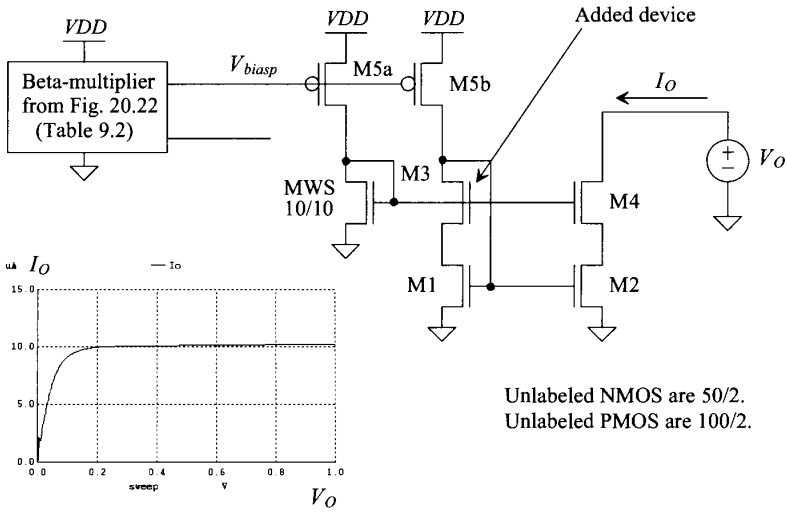


Figure 20.38 Wide-swing cascode current source in the short-channel process (good). Notice the drain-to-source voltages of M1 and M2 are the same.

Example 20.8

Regenerate the plot seen in Fig. 20.38 if MWS is sized using

$$\frac{W_{MWS}}{L_{MWS}} = \frac{1}{10} \cdot \frac{W}{L} \tag{20.63}$$

The simulation results are seen in Fig. 20.39 where MWS is sized 10/4. The output resistance is roughly four times smaller than what we got in Fig. 20.38 using Eq. (20.62). M3 is biased further away from the triode region. The point here is that there isn't any single equation that governs the selection of MWS size for all processes. Design trade-offs must be made.

Note that an interesting simulation to run at this point is to verify that the current mirror still functions correctly with small *VDD* (as seen in Fig. 20.23 with *VDD* down to 0.5 V). The output voltage, V_O, can still go to 1 V. ■

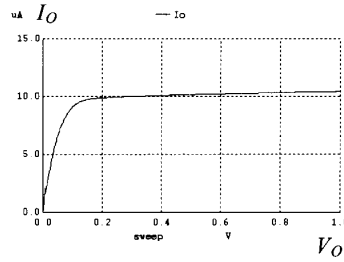


Figure 20.39 Resimulating the circuit in Fig. 20.39 using a 10/4 device for MWS.

IPR2023-00698
Theta EX2019

20.2.4 Regulated Drain Current Mirror

Consider adding an amplifier to the basic wide-swing current mirror as seen in Fig. 20.40. The purpose of the amplifier is to regulate or hold the drain of M2 at a fixed potential (in this case, the drain potential of M1). If we can hold the drain potential of M2 perfectly fixed, then M2's drain current won't vary and the output resistance of the current mirror will be infinite. In the next section, we'll discuss general biasing circuits. We'll use the drain voltage of M1 as a bias voltage simply for regulating the drain in current mirrors.

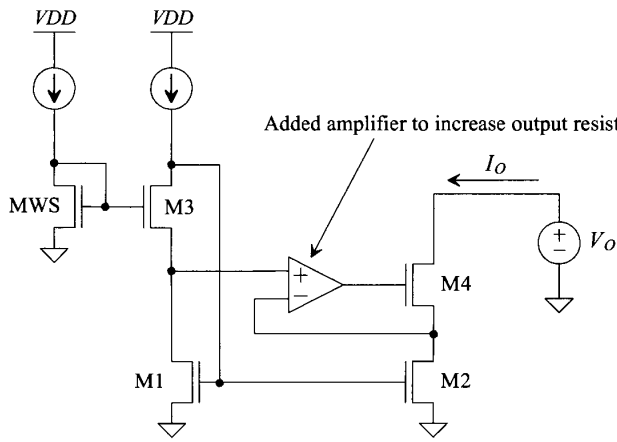


Figure 20.40 Regulating the drain of M2 using an amplifier.

We can estimate the enhancement in the output resistance of the current mirror by assuming the output of the added amplifier is related to its inputs using

$$v_{out} = A \cdot (v_+ - v_-) \tag{20.64}$$

where A is the gain of the added amplifier. Using Fig. 20.41, we can write, knowing the drain and gate of M1 are DC bias voltages (AC grounds),

$$v_{gs4} = -i_T r_o (A + 1) \tag{20.65}$$

and

$$i_T = g_m(-i_T r_o(A+1)) + \frac{v_T - i_T r_o}{r_o} \tag{20.66}$$

and finally,

$$R_o = \frac{v_T}{i_T} = 2r_o + g_m r_o^2 (A+1) \approx g_m r_o^2 \cdot A \tag{20.67}$$

When compared to Eq. (20.53), the output resistance is increased by the gain of the added amplifier. This result can be very useful (and we will use it often) in practical circuit design.

IPR2023-00698
Theta EX2019

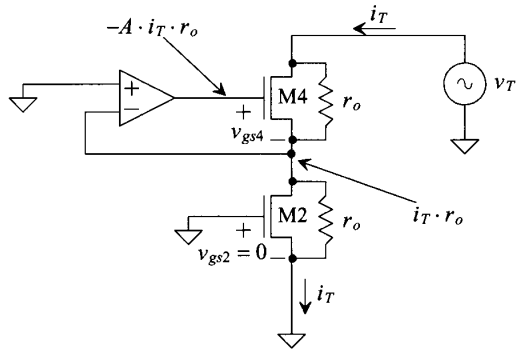


Figure 20.41 Determining the output impedance of a regulated drain current mirror.

A practical example of a current mirror that uses this technique is seen in Fig. 20.42. The drains of M1 and M2 are held at a V_{GS} by MA1 and MA2. If the drain potential of M2, for example, starts to decrease, MA2 starts shutting off, causing the gate voltage of M4 to increase, and pulling the drain potential of M2 back up. MA1 and MA2 are used to ensure that the circuit is symmetrical (M1 and M2 have the same V_{DS}). The drawback of this topology, as seen in the simulation results in the figure, is that the minimum voltage across the current mirror is $V_{DS,sat} + V_{GS}$ (just like the basic cascode, see Eq. [20.48]). The gain of the added amplifier (A) is, as we'll see in the next chapter, $g_m r_o/2$. We can estimate the output resistance as $g_m^2 r_o^3/2$ or 52 MΩ. Note that M3 is operating in the triode region.

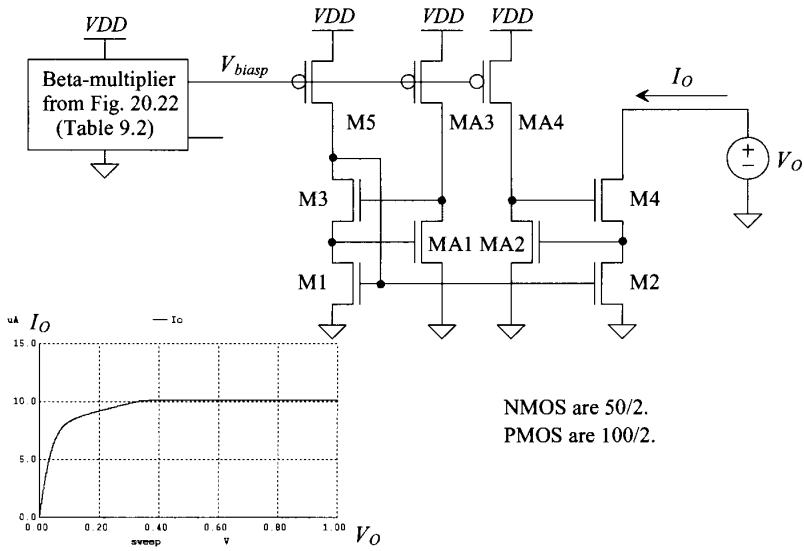


Figure 20.42 Using amplifiers to regulate the drain potentials in a current mirror.

IPR2023-00698
Theta EX2019

20.3 Biasing Circuits

In this section we provide biasing circuit examples for the chapters that follow based on the data in Tables 9.1 and 9.2.

20.3.1 Long-Channel Biasing Circuits

Figure 20.43 shows a general biasing circuit for the long-channel CMOS process based on the sizes and currents given in Table 9.1. The Beta-multiplier, self-referenced circuit from Fig. 20.15 is used for biasing wide-swing current mirrors (V_{bias1} to V_{bias4}). V_{high} and V_{low} will be used later with amplifiers for regulating the drain of a MOSFET, as seen in Fig. 20.40. V_{ncas} and V_{pcas} are used for a circuit called a "floating current source" (which is discussed later). Perhaps the best way to see how this circuit can be used is to give a couple of examples.

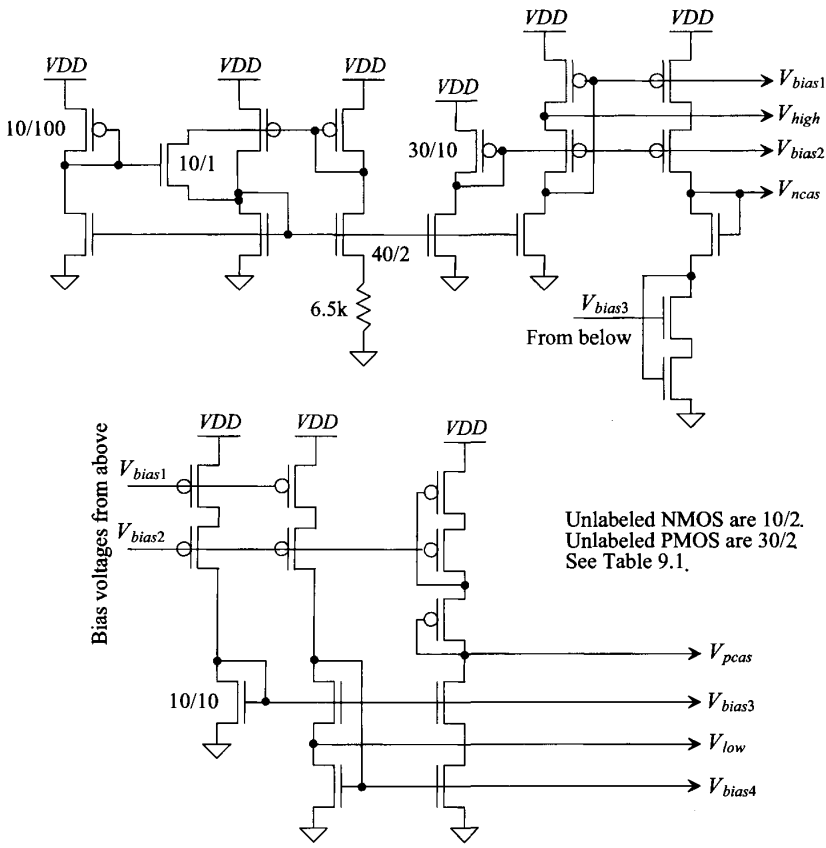
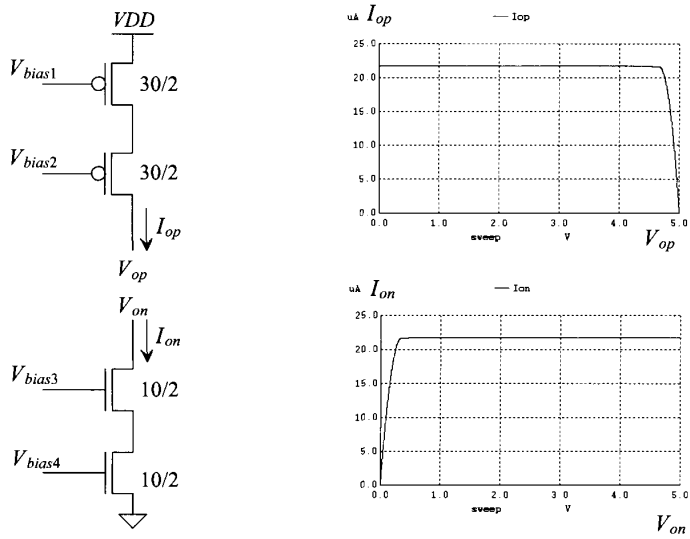


Figure 20.43 General biasing circuit for long-channel CMOS design using the data in Table 9.1

IPR2023-00698
Theta EX2019

Basic Cascode Biasing

Figure 20.44 shows how the bias voltages generated in Fig. 20.43 can be used to bias both PMOS and NMOS cascode current sources. As indicated in Table 9.1 and Fig. 20.31b, the minimum voltage across the current sources is $2V_{DS,sat}$ or 500 mV. The current that flows in the current sources is nominally 20 μA.



Bias voltages come from Fig. 20.43 (long-channel parameters in Table 9.1).

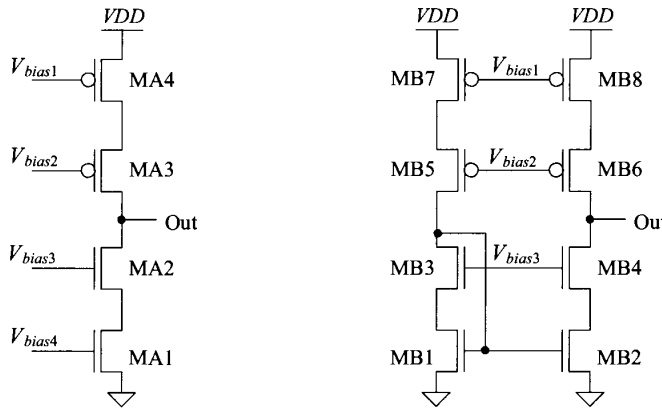
Figure 20.44 How cascode currents are biased and how they operate.

The Folded-Cascode Structure

Figure 20.45a shows a structure where an NMOS current source is connected to a PMOS current source. In this topology there is no way we can guarantee that the current flowing in the PMOS transistors is precisely equal to the current flowing in the NMOS transistors. This will cause the output of the circuit to move towards either *VDD* or ground (depending on which cascode structure is sourcing the most current). In (b) we fold the cascode structure over and diode-connect the folded NMOS devices (we could, just as well, have diode-connected the PMOS devices instead). Now, the gate potentials of MB1 and MB2 are set by the current that is sourced from MB5 and MB7. If the transistors are perfectly matched, the output voltage will equal the gate voltage of MB1 (the drain voltage of MB3) because of the circuit's symmetry. (We also saw this in Fig. 20.1.)

In Figure 20.46 we add some MOSFETs to steal 10 μA from MB7 and MB8. To make sure that the currents are equal, we size up MB7 and MB8 so that they source 30 μA. Note how we doubled the length of M1 and M2 so that they sink 10 μA from the same bias voltages.

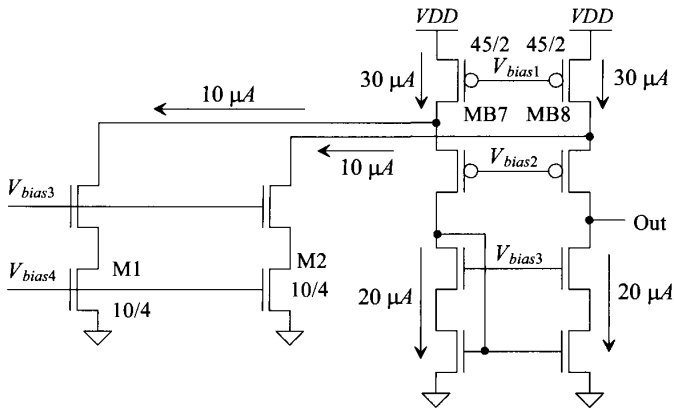
IPR2023-00698
Theta EX2019



(a) (b) Folded cascode structure

Bias voltages come from Fig. 20.43 (long-channel parameters in Table 9.1). All NMOS are 10/2, while all PMOS are 30/2.

Figure 20.45 Using a folded cascode structure to make sure that the current sourced by the PMOS equals the current sourced by the NMOS.



Bias voltages come from Fig. 20.43 (long-channel parameters in Table 9.1). All unlabeled NMOS are 10/2, while all unlabeled PMOS are 30/2.

Figure 20.46 Stealing current from the folded cascode structure.

IPR2023-00698
Theta EX2019

20.3.2 Short-Channel Biasing Circuits

Figure 20.47 shows a biasing circuit for a general analog design short-channel process. It uses the Beta-multiplier, self-referenced bias circuit from Fig. 20.22. As indicated in Sec. 20.1.4, the critical capacitance for stability is MCP. Notice that we are using the PMOS devices in the Beta-multiplier to bias the current mirrors, whereas in Fig. 20.43 we used the NMOS devices. We picked the PMOS here simply to increase the capacitance on V_{biasp} and so to further stabilize the circuit. Figure 20.48 shows the operation of the cascode current mirrors biased with the circuit in Fig. 20.47. It's important to realize that the minimum voltage across the current source is considerably above $2V_{DS,sat}$. Again, this was a design choice (see Eqs. [20.62] and [20.63]) to increase output resistance.

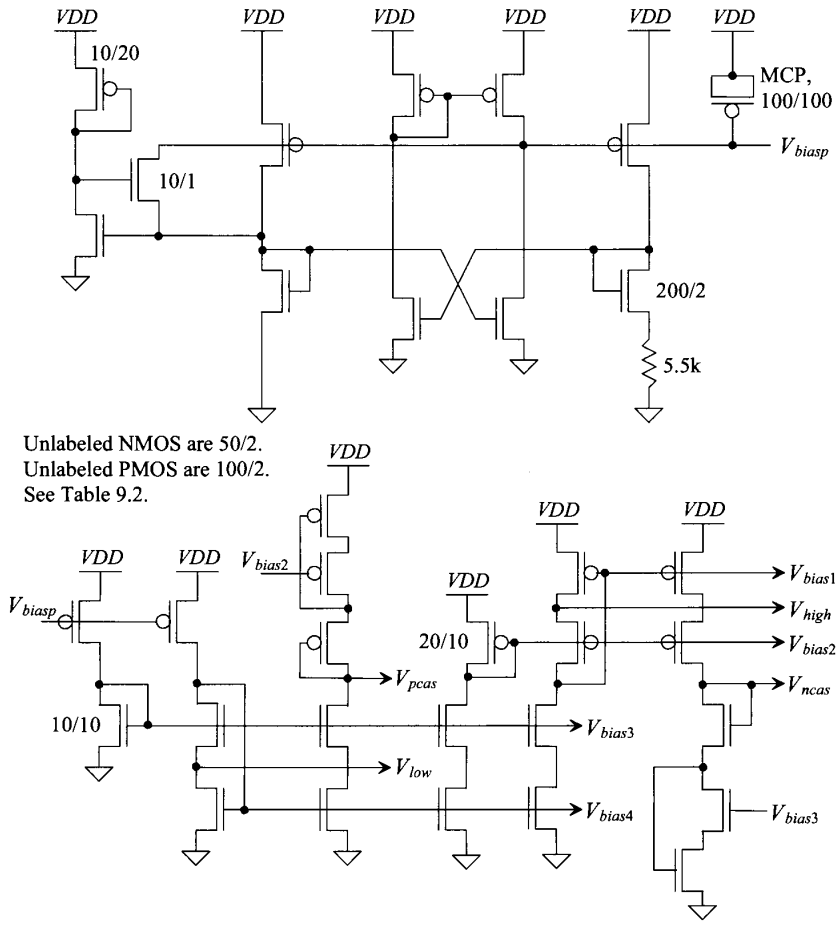
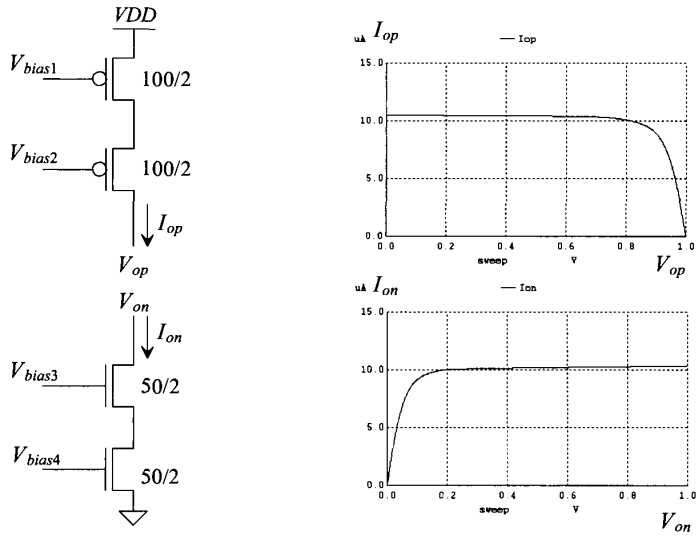


Figure 20.47 General biasing circuit for short-channel design using the data in Table 9.2.

IPR2023-00698
Theta EX2019



Bias voltages come from Fig. 20.47 (short-channel parameters in Table 9.2).

Figure 20.48 Cascode current sources operating in a short-channel process.

Floating Current Sources

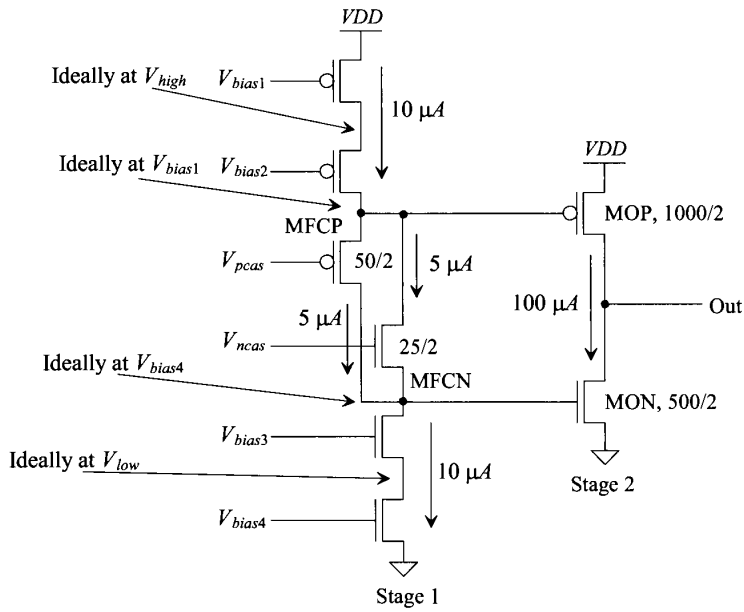
In Fig. 20.49 we add two MOSFETs, MFCP and MFCN, to the cascode structure to form a *floating current source*. The addition of MFCN, for example, allows the voltage across the PMOS cascode structure to become V_{bias1} (assuming that matching between the bias circuit and this circuit is perfect). The voltage across the NMOS cascode becomes V_{bias4}. This can be very useful for biasing the next stage in an amplifier. For example, if we connect an NMOS (MON in the figure) device to the source of MFCN, then, because this potential is ideally V_{bias4}, we treat the MOSFET as if it were biased with V_{bias4}. Since the width of MON in the figure is ten times the widths of the other NMOS, the current in the output stage will be 100 μA. Figure 20.50 shows the simulation results with currents that flow in stage 1 and stage 2 of the circuit.

Note that the circuit in Fig. 20.49 uses all of the bias voltages except for V_{low} and V_{high}. These voltages are used with circuits that employ drain regulation, as discussed in Sec. 20.2.4.

20.3.3 A Final Comment

The bias circuits in Figs. 20.43 and 20.47 will be used in the next several chapters to provide biasing for amplifier design examples. It's important to remember our constant theme when we design current mirrors, that is, if we can set both the gate-source and the drain-source voltages of a MOSFET, we set the drain current. Similarly if we can set the drain current and gate-source voltage, then we set the drain-source voltage.

IPR2023-00698
Theta EX2019



Bias voltages come from Fig. 20.47 (short-channel parameters in Table 9.2). Unlabeled NMOS are 50/2, while unlabeled PMOS are 100/2.

Figure 20.49 Biasing with a floating current source.

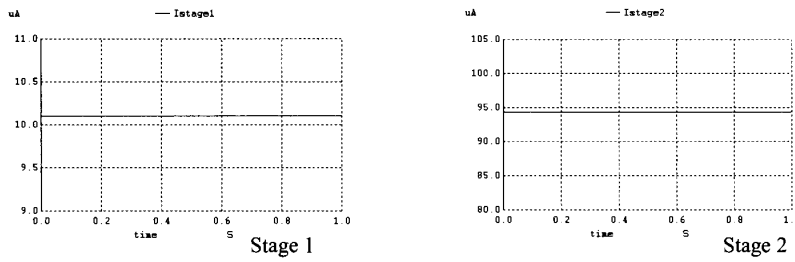


Figure 20.50 The simulated currents that flow in stages 1 and 2 of the circuit in Fig. 20.49.

ADDITIONAL READING

[1] S. J. Lovett, M. Welten, A. Mathewson, and B. Mason, "Optimizing MOS transistor mismatch," *IEEE Journal of Solid-State Circuits,* vol. 33, pp. 147–150, January 1998.

IPR2023-00698
Theta EX2019

[2] K. de Langen and J. H. Huijsing, "Compact low-voltage power-efficient operational amplifier cells for VLSI," *IEEE Journal of Solid-State Circuits,* vol. 33, pp. 1482–1496, October 1998.

[3] P. Larsson, "Parasitic resistance in an MOS transistor used as on-chip decoupling capacitance," *IEEE Journal of Solid-State Circuits,* vol. 32, pp. 574–576, April 1997.

[4] C. Yu and R. L. Geiger, "An automatic offset compensation scheme with ping-pong control for CMOS operational amplifiers," *IEEE Journal of Solid-State Circuits,* vol. 29, pp. 601–610, May 1994.

[5] R. G. Eschauzier, R. Hogervorst, and J. H. Huijsing, "A programmable 1.5 V CMOS class-AB operational amplifier with hybrid nested Miller compensation for 120 dB gain and 6 MHz UGF," *IEEE Journal of Solid-State Circuits,* vol. 29, pp. 1497–1504, December 1994.

[6] R. Hogervorst, J. P. Tero, R. G. H. Eschauzier, and J. H. Huijsing, "A Compact Power-Efficient 3 V CMOS Rail-to-Rail Input/Output Operational Amplifier for VLSI Cell Libraries," *IEEE Journal of Solid State Circuits,* vol. 29, pp. 1505–1513, December 1994.

[7] E. Säckinger and W. Guggenbühl, "A High-Swing, High-Impedance MOS Cascode Circuit," *IEEE Journal of Solid-State Circuits,* vol. 25, pp. 289–298, February 1990.

[8] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE Journal of Solid-State Circuits*, vol. 24, pp. 1433–1439, October 1989.

[9] K. Lakshmikumar, R. A. Hadaway, and M. A. Copeland, "Characterisation and Modeling of Mismatch in MOS Transistors for Precision Analog Design," *IEEE Journal of Solid-State Circuits*, vol. 21, pp. 1057–1066, December 1986.

[10] J. Shyu, F. Krummenacher, and G. C. Temes, "Random error effects in matched MOS capacitors and current sources," *IEEE Journal of Solid-State Circuits,* vol. 19, pp. 948–956, December 1984.

[11] P. R. Gray and R. G. Meyer, "MOS operational amplifier design - A tutorial overview," *IEEE Journal of Solid-State Circuits,* vol. 17, pp. 969–982, December 1982.

[12] J. L. McCreary, "Matching properties, and voltage and temperature dependence of MOS capacitors," *IEEE Journal of Solid-State Circuits,* vol. 16, pp. 608–616, December 1981.

[13] W. Steinhagen and W. L. Engl, "Design of integrated analog CMOS circuits - A multichannel telemetry transmitter," *IEEE Journal of Solid-State Circuits,* vol. 13, pp. 799–805, December 1978.

[14] E. Vittoz and J. Fellrath, "CMOS analog integrated circuits based on weak inversion operations," *IEEE Journal of Solid-State Circuits,* vol. 12, pp. 224–231, June 1977.

IPR2023-00698
Theta EX2019

PROBLEMS

20.1 Using the CMOS long-channel process (1 μm), determine the current flowing in the circuit seen in Fig. 20.51. Verify your answer with SPICE.

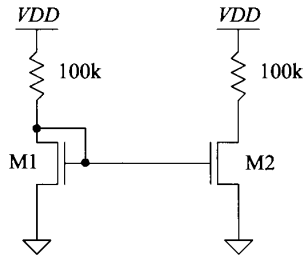


Figure 20.51 Current mirror used in Problem 20.1.

20.2 Repeat problem 20.1 for the circuit in Fig. 20.52. Can M1 and M2 be replaced with a single MOSFET? If so how and what size? If not why?

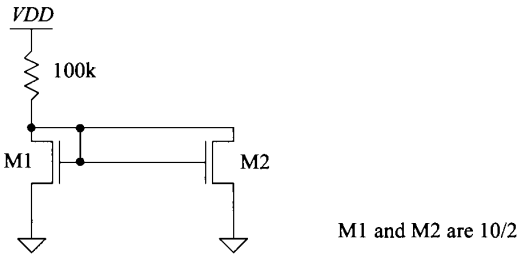


Figure 20.52 Circuit used in Problem 20.2.

20.3 Show the SPICE simulations for the PMOS devices in Ex. 20.1.

20.4 A threshold voltage mismatch in SPICE can be modeled by adding a small voltage source in series with the gate of one of the MOSFETs, as seen in Fig. 20.53. Show, using SPICE, that Eq. (20.8) is valid.

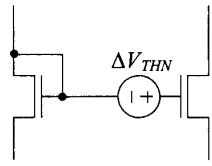


Figure 20.53 Modeling an offset voltage in SPICE (or for hand calculations).

IPR2023-00698
Theta EX2019

20.5 Show, using simulations and hand-calculations, that by using a larger value of $V_{DS,sat}$ when designing bias circuits, the MOSFETs enter the triode region earlier.

20.6 In Ex. 20.2, how does the gate voltage of M1/M2 change as *VDD* is decreased? How does the V_{SG} change? Use SPICE to verify your answers.

20.7 For the same bias current used in Table 9.1 (nominally 20 μA), regenerate Fig. 20.13 if minimum-length MOSFETs are used in the current mirror. Show the hand calculations leading to the selection of the MOSFET sizes.

20.8 Suppose that the bias circuit used to generate Fig. 20.16 shows a reference current change with *VDD,* as seen in Fig. 20.54. Knowing that the reference current should be constant with changes in *VDD,* what is wrong with the reference? (What important portion of the reference is causing the problem?)

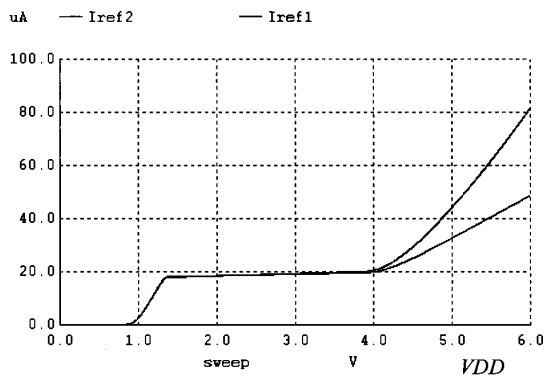


Figure 20.54 Problems with reference current increasing with VDD.

20.9 Using hand calculations, estimate the value of the voltage across *R* for the current reference in Fig. 20.15.

20.10 Using SPICE, determine if the reference circuit seen in Fig. 20.22 can become unstable with changes in *VDD*. Show that reducing the size of MCP and MCN affects the stability of the circuit.

20.11 Estimate the temperature behavior of the gate voltage of M1 in Fig. 20.10.

20.12 Design a voltage reference using the Beta-multiplier in the short-channel process and the data from Table 9.2. Generate a figure similar to Fig. 20.27 to show the temperature performance of your design.

20.13 *K* can be used to minimize *R* in Eq. (20.46) for a fixed-reference current. If M1/M2 experience a threshold voltage mismatch (see problem 20.4), discuss and show with simulations how much ΔV_{THN} can be tolerated and the error in the reference current that results.

IPR2023-00698
Theta EX2019

20.14 Using an AC test voltage, determine the output resistance of the MOSFET current mirror seen in Fig. 20.28.

20.15 Based on the data in Table 9.1, what are the voltages on the drain, gate, and source terminals of M1–M4 in Fig. 20.29, assuming that the devices are operating in the saturation region? Compare your hand calculations to simulations.

20.16 If the MOSFET in Fig. 20.55 is operating in the saturation region, determine the small-signal resistance looking into its drain.

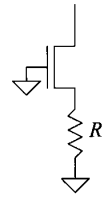


Figure 20.55 Resistance looking into the drain of a MOSFET with a source resistance. See Problem 20.16.

20.17 Suppose that the wide-swing current mirror seen in Fig. 20.38 is implemented in the long-channel CMOS process (Table 9.1). Estimate the size of MWS when M3 triodes. Verify your answer with simulations.

20.18 Label the voltages, based on the values in Table 9.1, in the schematic of the general biasing circuits seen in Fig. 20.43. Compare the tabulated values of V_{GS}, V_{DS}, etc., to the simulated values.

20.19 How would the current in Fig. 20.44 change if the width of the device connected to V_{bias2} were doubled? Were halved? Why? Explain what is going on in the circuit.

20.20 Should the voltage labeled "Out" in Fig. 20.49 be at a specific value? Why or why not?

IPR2023-00698
Theta EX2019

Chapter
21

Amplifiers

In this chapter we turn our attention towards amplifiers. Single-stage amplifiers are used in virtually every op-amp design. By replacing a passive load resistor with a MOSFET transistor (called an *active load*), significant amounts of chip area can be saved. An active load can also produce higher values of resistance when compared with a passive resistor, resulting in higher gains.

Several types of active loads are studied in this chapter. The *gate-drain load* consists of a MOSFET with the gate and drain shorted and provides large bandwidths and low-output impedance at the cost of reduced gain. The *current source load* amplifier has a higher gain and output impedance at the expense of lower bandwidth. Current source load amplifiers are preferred when external feedback is applied to the amplifier to set the gain. This chapter explores basic single-stage amplifiers with active loads, as well as the trade-offs associated with each. The cascode amplifier is examined in detail along with several configurations of output stages, including the push-pull amplifier.

21.1 Gate-Drain-Connected Loads

A gate-drain-connected MOSFET (load) is half of a current mirror as seen in Fig. 20.1 in the last chapter. Here we move the discussion from the DC operating conditions and biasing presented in the last chapter to AC small-signal analysis.

21.1.1 Common-Source (CS) Amplifiers

A gate-drain-connected MOSFET (operating with a nonzero drain current) can be thought of as a resistor of value $1/g_m$ (see Fig. 9.18 and Eq. [9.25]). The four possible common-source (CS) amplifiers using gate-drain-connected loads are seen in Fig. 21.1. In each configuration, M1 and M2 are assumed to be biased in the saturation region. Notice how the source of the amplifying MOSFET (i.e., not the load or gate-drain MOSFET) is common to both the input and the output in each configuration.

Examine the circuit shown in Fig. 21.1a. Since many MOSFET amplifiers are analyzed in this chapter and the next several chapters using small-signal analysis, an intuitive analysis approach will be used. This approach allows the designer to quickly analyze the gain of a circuit. To analyze the gain of the circuit in Fig. 21.1a, begin by

IPR2023-00698
Theta EX2019

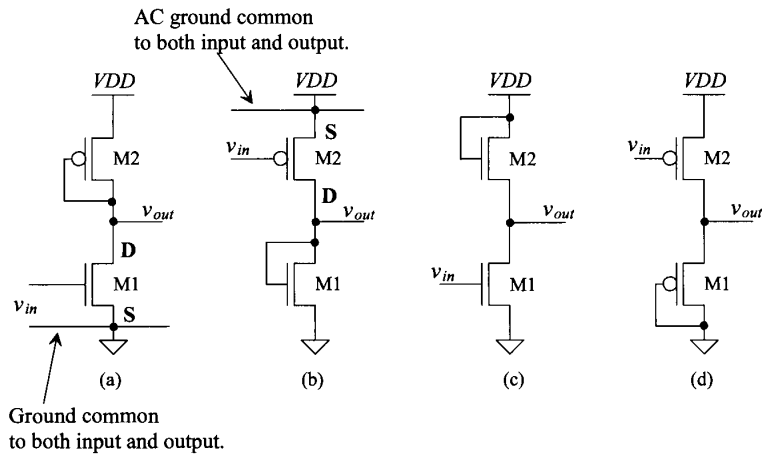


Figure 21.1 Four possible configurations of common-source amplifiers with gate-drain loads.

replacing M2 with a resistor of value $\frac{1}{g_{m2}}$ and replacing M1 with a current source of value $g_{m1} \cdot v_{in}$, provided $\frac{1}{g_{m2}} << r_{o1} || r_{o2}$. The equivalent circuit is shown in Fig. 21.2. Note that there is no body effect in either MOSFET and that only the low-frequency model is shown. Again, as seen in Ch. 9, a gate-drain-connected MOSFET has a small-signal resistance of value $\frac{1}{g_m}$. The small-signal gain of the **common-source amplifier** is given by

$$\frac{v_{out}}{v_{in}} = \frac{-i_d \cdot \frac{1}{g_{m2}}}{i_d \cdot \frac{1}{g_{m1}}} = -\frac{\frac{1}{g_{m2}}}{\frac{1}{g_{m1}}} = -\frac{\text{resistance in the drain}}{\text{resistance in the source}} = -\frac{g_{m1}}{g_{m2}} \qquad (21.1)$$

This result is very important in the intuitive analysis. It states that the small-signal gain of a common-source amplifier is simply the resistance in the drain of M1 divided by the resistance in the source (the resistance looking into the source of M1 added to any resistance from the source of M1 to ground). The actual value of the resistance in the

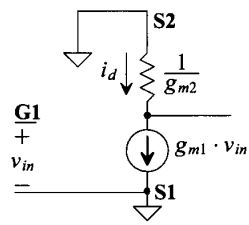


Figure 21.2 Simplified circuit of Fig. 21.1a.

IPR2023-00698
Theta EX2019

drain, as will be soon seen, is $\frac{1}{g_{m2}}||r_{o1}||r_{o2}$. In most cases, however, r_o will be much greater than $\frac{1}{g_m}$, and the approximation is an accurate one.

Example 21.1

Determine the small-signal AC gain of the circuit shown in Fig. 21.3 using the intuitive method presented in Eq. (21.1).

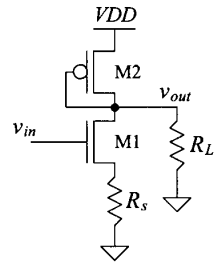


Figure 21.3 Amplifier analyzed in Ex. 21.1.

The small-signal gain of this circuit is given by

$$A_v = \frac{v_{out}}{v_{in}} = -\frac{\text{resistance in the drain}}{\text{resistance in source}}$$

The resistance in the drain is the parallel combination of all resistances connected to the drain of M1, given by

$$\text{Resistance in the drain} = \frac{1}{g_{m2}}||R_L \text{ for } r_{o1}||r_{o2} >> \frac{1}{g_{m2}}$$

The resistance in the source is the sum of the resistance looking into the source of M1 ($1/g_{m1}$) and the resistance connected from the source of M1 to ground. For the present problem, this resistance is given by

$$R_{\text{in the source}} = \frac{1}{g_{m1}} + R_s$$

The voltage gain is then given by

$$A_v = -\frac{\frac{1}{g_{m2}}||R_L}{\frac{1}{g_{m1}} + R_s} \text{ which for } R_L \to \infty \text{ and } R_s \to 0 \text{ reduces to } -\frac{g_{m1}}{g_{m2}} \blacksquare$$

We can determine the exact gain of the amplifier of Fig. 21.1a using the full small-signal model shown in Fig. 21.4. Using KCL at the output of the amplifier gives

$$g_{m1}v_{in} + \frac{v_{out}}{r_{o2}||r_{o1}} = -g_{m2}v_{out} \tag{21.2}$$

or

$$A_v = \frac{v_{out}}{v_{in}} = -\frac{g_{m1}}{g_{m2} + \frac{1}{r_{o1}||r_{o2}}} = -\frac{\frac{1}{g_{m2}}||r_{o1}||r_{o2}}{\frac{1}{g_{m1}}} = -\frac{\text{resistance in the drain}}{\text{resistance in the source}} \tag{21.3}$$

IPR2023-00698
Theta EX2019

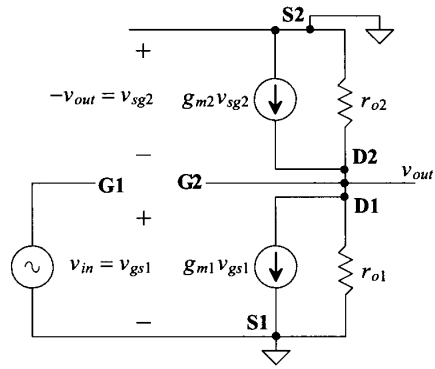


Figure 21.4 Small-signal model of the amplifier shown in Fig. 21.1(a).

When $\frac{1}{g_{m2}} << r_{o1} || r_{o2}$, this reduces to

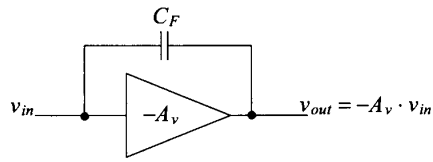
$$A_v = -\frac{\frac{1}{g_{m2}}}{\frac{1}{gm1}} = -\frac{g_{m1}}{g_{m2}} \tag{21.4}$$

which is the same form as Eq. (21.1).

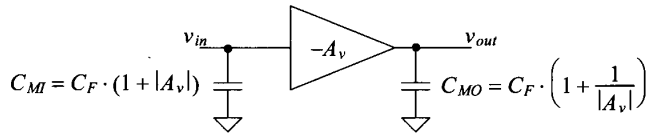
Miller's Theorem

Consider the inverting amplifier seen in Fig. 21.5a with a gain of $-A_v$ and a feedback capacitance between the output and the input of the amplifier. The current supplied by v_{in} or v_{out} to the capacitor is

$$i = \frac{v_{in} - v_{out}}{1/j\omega C_F} = j\omega C_F \cdot (1 + |A_v|) \cdot v_{in} = j\omega C_F \cdot \left(1 + \frac{1}{|A_v|}\right) \cdot v_{out} \tag{21.5}$$



(a)



(b)

Figure 21.5 The derivation of Miller's theorem.

IPR2023-00698
Theta EX2019

In Fig. 21.5b we replace the feedback capacitor with capacitors on the input and the output of the amplifier to ground. These (adjusted size) capacitors require the same currents from the input source and the output of the amplifier. Note that this substitution, Eq. (21.5), is often called *Miller's theorem,* and the feedback capacitor, C_F, is called the *Miller capacitance.*

Notice that the effective capacitance on the input of the amplifier is multiplied by the gain of the amplifier. For big gains, $|-A_v|$, this means that the input capacitance of the amplifier can be large and can result in slow circuits. Intuitively, this can be understood by realizing that the effective voltage across C_F is $A_v + 1$ times larger than the input signal itself. This large voltage drop across C_F (one side going up by v_{in} while the other side goes down by $A_v \cdot v_{in}$) results in a large displacement current, which makes the capacitor appear larger than it really is.

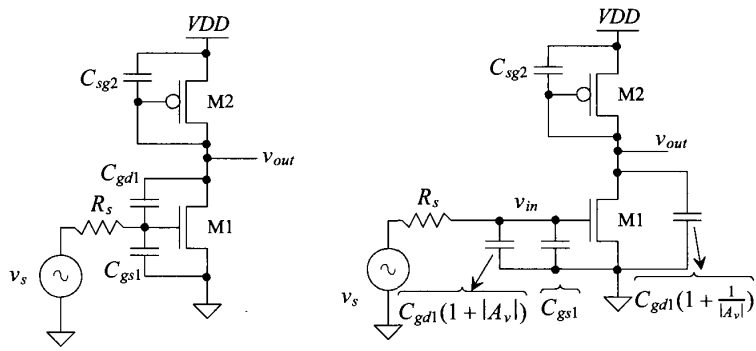
Frequency Response

Now consider the frequency response of the CS amplifier of Fig. 21.1a. The high-frequency (meaning that the device capacitances are included in the circuit) equivalent circuit is shown in Fig. 21.6a. The gate-drain capacitance of M2 isn't drawn because the gate and drain of M2 are shorted together. A source resistance is added to the circuit to model the effects of the driving source impedance. Miller's theorem is used to break C_{gd1} into two parts: a capacitance at the gate of M1 to ground and a capacitance from the drain of M1 to ground, as seen in Fig. 21.6b (note that the low-frequency gain of the amplifier in Fig. 21.1a is $-g_{m1}/g_{m2}$). Two *RC* time constants exist in this circuit: one on the input of the circuit and one on the output of the circuit. The input time constant is given by

$$\tau_{in} = R_s(C_{MI} + C_{gs1}) \tag{21.6}$$

where the Miller capacitance at the input, C_{MI}, is

$$C_{MI} = C_{gd1}\left(1 + \frac{g_{m1}}{g_{m2}}\right) \tag{21.7}$$



(a) (b) After applying Miller's theorem

Figure 21.6 Frequency response of the CS amplifier with gate-drain-connected load.

IPR2023-00698
Theta EX2019

The output time constant is found in a similar way and is given by

$$\tau_{out} = \frac{1}{g_{m2}} \cdot (C_{sg2} + C_{MO}) \tag{21.8}$$

The Miller capacitance which appears at the output, C_{MO}, now becomes

$$C_{MO} = C_{gd1}(1 + \frac{g_{m2}}{g_{m1}}) \tag{21.9}$$

The frequency response of the amplifier in Fig. 21.1a (Fig. 21.6a) is given by

$$A_v(f) = \frac{-\frac{g_{m1}}{g_{m2}}}{\left(1 + j\frac{f}{f_{in}}\right)\left(1 + j\frac{f}{f_{out}}\right)} \tag{21.10}$$

where the pole associated with the amplifier's input is located at

$$f_{in} = \frac{1}{2\pi\tau_{in}} \tag{21.11}$$

and the pole associated with the amplifier's output is located at

$$f_{out} = \frac{1}{2\pi\tau_{out}} \tag{21.12}$$

The Right-Half Plane Zero

It should be noted that when using Miller's theorem a zero is neglected. If we take a look at Fig. 21.5a, we see that at high frequencies C_F shorts the amplifier's output to its input. Our model in Fig. 21.5b doesn't show this shorting. What this means for the amplifier in Fig. 21.1a or its model in Fig. 21.6a is that C_{gd1} shorts the output to the input at high frequencies.

To characterize this high-frequency effect, where we can't use Miller's theorem, consider the circuit seen in Fig. 21.7. In this figure we model the amplifier in Fig. 21.6 (or Fig. 21.1) with a small-signal resistance on the output of the circuit called $R_o = \frac{1}{g_{m2}} || r_{o2} || r_{o1} \approx \frac{1}{g_{m2}}$ and a capacitance on the output, without including C_{gd1}, of $C_o = C_{sg2}$. We don't concern ourselves with the input time constant but rather just look at the output of the circuit. Summing the currents on the output of the model results in

$$\frac{v_{out} - v_{in}}{1/j\omega C_{gd1}} + \frac{v_{out}}{R_o || 1/j\omega C_o} + g_{m1} \cdot v_{in} = 0 \tag{21.13}$$

Solving for v_{out}/v_{in} gives

$$\frac{v_{out}}{v_{in}} = -g_{m1}R_o \cdot \frac{1 - j\omega\frac{C_{gd1}}{g_{m1}}}{1 + j\omega(C_{gd1} + C_o) \cdot R_o} \tag{21.14}$$

The pole should be recognized as what we wrote in Eq. (21.12) if the gain of the amplifier is small ($A_v = -g_{m1}/g_{m2} \approx 0$). In the numerator we see a right-half plane zero at

$$f_z = \frac{g_{m1}}{2\pi C_{gd1}} \tag{21.15}$$

A zero in the right-half plane (RHP) results in the same magnitude response as a zero in the left-half plane (LHP). However, the phase response is different. A zero in the RHP has the same influence on the phase of a system as a pole in the LHP. Understanding this

IPR2023-00698
Theta EX2019

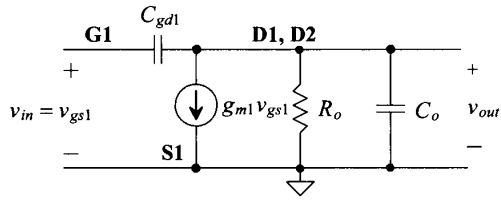


Figure 21.7 Small-signal model used to calculate the RHP zero for amplifiers in Fig. 22.1a–d.

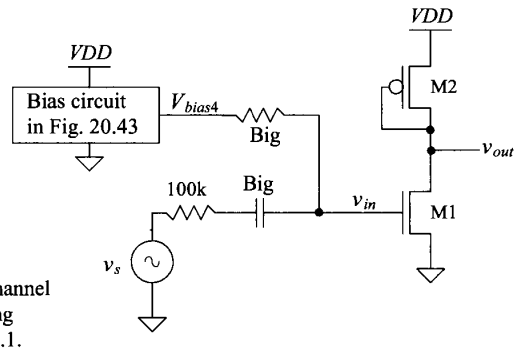
point is important. For the inverting amplifiers (which mean that at low frequencies their phase shift is 180°) in Fig. 21.1, for example, the input feeding directly to the output because of the RHP zero, without the inversion, can result in instability when the amplifier is used with feedback. The output can feed back and add to the input (positive feedback). The result is that the output of the amplifier will grow without limit (and, ultimately, the output of the amplifier will oscillate back and forth between some voltages). The magnitude response of the RHP zero is usually ignored because its value typically places it well beyond the pole frequency. However, in CMOS op-amp design, the phase shift of this RHP zero can, and will, result in stability and/or settling time issues.

Example 21.2

Determine the gain, magnitude, and phase shift of the amplifier shown in Fig. 21.8. Using an AC SPICE simulation, verify your hand calculations.

The bias circuit is used to bias the amplifier at the operating point seen in Table 9.1, that is, at a drain current of 20 μA. The big resistor and capacitor set the DC bias point but don't affect the AC operation of the circuit.

$$A_v = -\frac{\frac{1}{g_{m2}}}{\frac{1}{g_{m1}}} = -\frac{g_{m1}}{g_{m2}} = -\frac{150\ \mu A/V}{150\ \mu A/V} = -1$$



Use the long-channel sizes and biasing seen in Table 9.1.

Figure 21.8 Amplifier used in Ex. 21.2.

IPR2023-00698
Theta EX2019

From Table 9.1, C_{gs1} = 23.3 *f*F and C_{gd1} = 2 *f*F. The Miller input capacitance, because the gain of the amplifier is –1, is then 4 *f*F. The input pole is located at

$$f_{in} = \frac{1}{2\pi \cdot R_s(C_{MI} + C_{gs1})} = \frac{1}{2\pi \cdot 100k \cdot (27.3\,fF)} = 58\ MHz \qquad (21.16)$$

From Table 9.1, C_{sg2} = 70 *f*F. The Miller output capacitance is, again, 4 *f*F. The output pole is located at

$$f_{out} = \frac{1}{2\pi \cdot \frac{1}{g_{m2}} \cdot (C_{MO} + C_{sg2})} = \frac{1}{2\pi \cdot 6.5k \cdot (74\,fF)} = 331\ MHz \qquad (21.17)$$

The zero is located at

$$f_z = \frac{g_{m1}}{2\pi C_{gd1}} = \frac{150\ \mu A/V}{2\pi \cdot 2\,fF} = 11.9\ GHz$$

The transfer function of the amplifier is then

$$A_v(f) = \frac{v_{out}(f)}{v_s(f)} = -\frac{\left(1 - j\frac{f}{11.9\ GHz}\right)}{\left(1 + j\frac{f}{58\ MHz}\right)\left(1 + j\frac{f}{331\ MHz}\right)}$$

Figure 21.9 shows the AC analysis simulation results for the amplifier of Fig. 21.8. What would happen if we used v_{in} instead of v_s? We wouldn't see the input

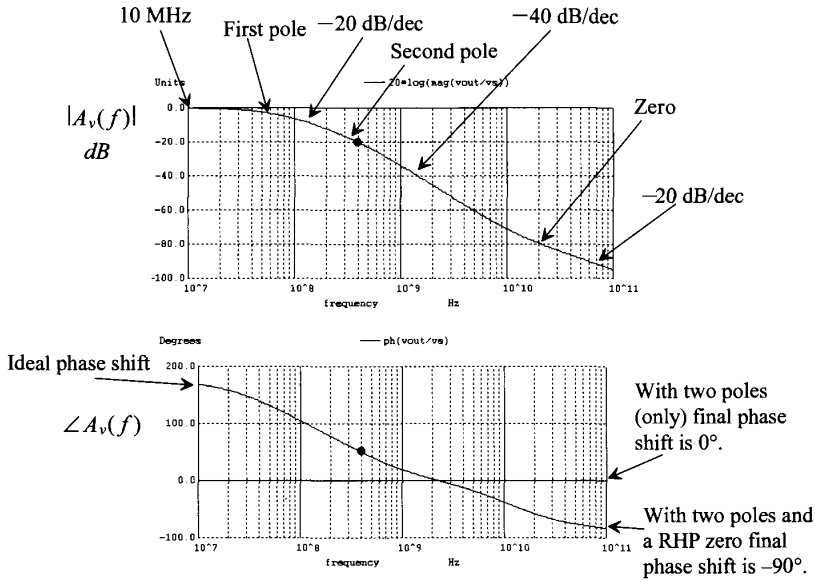


Figure 21.9 Frequency response of the amplifier in Fig. 21.8.

IPR2023-00698
Theta EX2019

pole's effects on the frequency response. In the magnitude response we see that right after the first pole the gain decreases by a factor of 10 for every increase in frequency by 10 (equivalent to saying –20 dB/decade). We could also say that the gain decreases by 2 for every increase in frequency by 2 (equivalent to saying –6 dB/octave). After the second pole, the gain decreases by 100 for every increase in frequency by 10. The zero causes the frequency response to go back to –20 dB/decade. (The output and input of M1 are shorted together so that all we see are the effects of R_s charging the input capacitance of the amplifier.) It is useful to look at, and understand, the behavior of the amplifier from v_{in} to v_{out} in the simulation.

The effects of the RHP zero occur one decade before f_z. This *excess phase* shift is the problem we discussed earlier. It makes the amplifier appear not like an inverting amplifier (phase shift of –180 degrees) but rather like a noninverting amplifier sooner (at lower frequencies).

Note that magnitude of voltage gains in the neighborhood of 1 are common when using gate-drain-connected loads. At this point it may appear that this amplifier is almost worthless; however, we will see that when this amplifier is used with additional circuitry (e.g., a current source load for a differential pair), it can be very useful. ■

Example 21.3

Determine the magnitude and phase of the time-domain signal v_{out}, at 400 MHz, for the amplifier in Ex. 21.2 based on the frequency plots seen in Fig. 21.9. Verify the results using SPICE.

The magnitude response at 400 MHz is

$$\left|\frac{v_{out}}{v_s}\right| = -20\ dB = 0.1$$

and the phase response is

$$\angle\frac{v_{out}}{v_{in}} = 50° = -310°$$

The positive value means that the output is leading, or occurring earlier in time, than the input. Remembering that phase shift is simply an indication of a time delay at a particular frequency, we can write

$$50° = 360 \cdot \frac{t_d}{T} = 360 \cdot t_d \cdot 400\ MHz \rightarrow t_d \approx 350\ ps$$

for a period of 2.5 ns. We'll use an input voltage of 1 mV to ensure small-signal operation. In the simulation that generated Fig. 21.9 our input AC signal was 1 V. Again, remember that SPICE replaces the active devices with linear models in an AC simulation so any value of amplitude can be used for the input signals. Using 1 V in an AC simulation is convenient because of the ratios often used when calculating gains. Figure 21.10 shows the SPICE simulation results. Notice how the peak amplitude of the input is 1 mV while the peak amplitude of the output (remembering the DC output voltage is $VDD - V_{SG}$ or, from Table 9.1, approximately 3.85 V) is 0.1 mV (–20 db down from the low-frequency value). In

IPR2023-00698
Theta EX2019

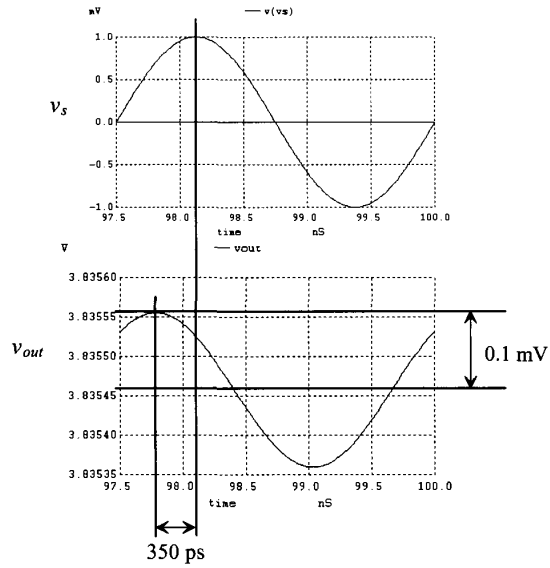


Figure 21.10 Time-domain behavior of the amplifier in Fig. 21.8 at 400 MHz.

a transient simulation we have to wait some time before the circuit settles to a stable operating point. In Fig. 21.10 we've waited 97.5 ns before looking at the output voltage. This gives the bias circuit, for example, time to turn on and generate the bias voltage needed in the amplifier (the value of the big capacitor was also reduced in this transient simulation so that its charging wouldn't increase the simulation time.) ■

A Common-Source Current Amplifier

The previous discussion was centered around CS voltage amplifiers. Next consider the circuit shown in Fig. 21.11. This amplifier can be thought of as a transimpedance amplifier, that is, voltage out and current in. The ideal input impedance of an amplifier with a current input is zero (all of the input current flows into the amplifier with this ideal input impedance). Note that the biasing of the amplifier is controlled by the current mirror formed between M3 and M1 (drawn in a different way to ensure that the reader recognizes the topology when it is seen).

The input impedance of the amplifier in Fig. 21.11 is given by

$$R_{in} = \frac{1}{g_{m3}} || r_{o3} || r_{o4} \approx \frac{1}{g_{m3}} \tag{21.18}$$

The input current, i_{in}, is determined by

$$i_{in} = g_{m3} v_{in} = g_{m3} v_{gs3} = g_{m3} v_{gs1} \tag{21.19}$$

The current through M1, and thus M2, is given by

$$i_d = g_{m1} v_{in} = \frac{g_{m1}}{g_{m3}} \cdot i_{in} \tag{21.20}$$

IPR2023-00698
Theta EX2019

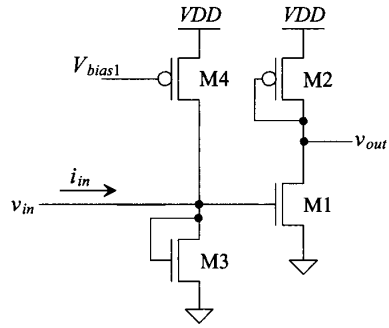


Figure 21.11 A transimpedance amplifier.

This equation can be rewritten, $g_{m3} = \beta_3(V_{GS3} - V_{THN})$ and $g_{m1} = \beta_1(V_{GS3} - V_{THN})$, as

$$\frac{i_d}{i_{in}} = \frac{\beta_1}{\beta_3} = \frac{W_1 L_3}{W_3 L_1} \tag{21.21}$$

That is, we can get a current gain simply by adjusting the size of M1 and M3. The current in M2 can then be mirrored out to a load. The transresistance gain of this configuration is given by

$$A_R = \frac{v_{out}}{i_{in}} = -\frac{i_d \frac{1}{g_{m2}}}{i_d \cdot \frac{g_{m3}}{g_{m1}}} = -\frac{g_{m1}}{g_{m2} g_{m3}} \tag{21.22}$$

Note that adding M3 and thus the resistor of value $\frac{1}{g_{m3}}$ to ground lowers the input time constant and increases the bandwidth of the amplifier.

Common-Source Amplifier with Source Degeneration

The amplifier in Fig. 21.3 is an example of a CS amplifier with source degeneration. The gain of the amplifier is reduced because the effective transconductance of the amplifier becomes

$$\frac{1}{g_{m,eff}} = \frac{1}{g_m} + R_s \tag{21.23}$$

as seen in Ex. 21.1. Adding the source resistance is useful, especially if $R_s >> 1/g_m$, to precisely set the transconductance of an amplifier. Note that we didn't, in Ex. 21.1, include the resistance looking into the drain of M1 when we wrote the "resistance in the drain." *Let's review, or again verify, how to determine the small-signal, low-frequency resistances looking into the drain and source of a MOSFET.* Using this information will greatly aid in our intuitive analysis of CMOS circuits.

Consider the test circuit shown in Fig. 21.12. Here the DC voltages ensure that the MOSFET is operating in the saturation region. We are concerned with the AC current that flows through v_t. This test voltage divided by i_d gives us the small-signal AC resistance looking into the drain of the MOSFET. In other words, this is the resistance connected to ground at the drain node.

IPR2023-00698
Theta EX2019

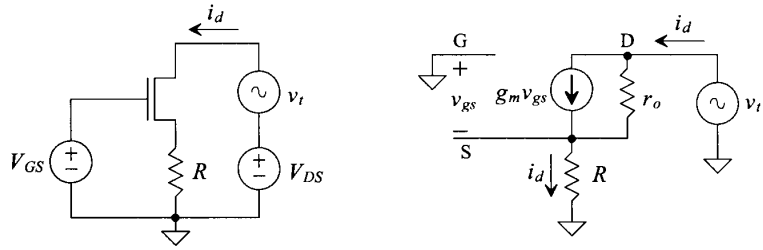


Figure 21.12 Circuit used to determine the resistance looking into the drain of a MOSFET.

Neglecting the body effect, we can write the test voltage as

$$v_t = (i_d - g_m v_{gs}) r_o + i_d R \tag{21.24}$$

where (noting $v_g = 0$ and $v_s = i_d R$)

$$v_{gs} = - i_d R \tag{21.25}$$

The resistance looking into the drain of the MOSFET is then given by

$$R_o = r_d = \frac{v_t}{i_d} = (1 + g_m R) r_o + R \approx (1 + g_m R) r_o \tag{21.26}$$

keeping in mind that $r_o \approx \frac{1}{\lambda I_D}$. For a cascoded MOSFET topology ($R = r_o$), this is approximately, as seen in Eq. (20.53), $g_m r_o^2$. For a *triple cascode*, three MOSFETs in series, ($R = g_m r_o^2$), the output resistance is

$$R_o \approx g_m^2 r_o^3 \tag{21.27}$$

noting that the output resistance goes up with each added MOSFET by the open-circuit gain, $g_m r_o$.

Example 21.4

Repeat Ex. 21.1 without neglecting the loading effects of the output resistances of M1 and M2.

The resistance looking into the drain of M2 is $\frac{1}{g_{m2}} || r_{o2}$, while the resistance looking into the drain of M1 is $r_{o1}(1 + g_{m1} R_s)$. The exact gain of the circuit is

$$A_v = \frac{v_{out}}{v_{in}} = - \frac{\frac{1}{g_{m2}} || r_{o2} || [r_{o1}(1 + g_{m1} R_s)] || R_L}{R_s + \frac{1}{g_{m1}}}$$

or when $r_{o1} || r_{o2} >> \frac{1}{g_{m2}}$, this reduces to

$$A_V = \frac{\frac{1}{g_{m2}} || R_L}{R_s + \frac{1}{g_{m1}}} \quad \blacksquare$$

IPR2023-00698
Theta EX2019

The resistance in the source of the MOSFET can be found using the circuit shown in Fig. 21.13. Neglecting the body effect and output resistance of the MOSFET gives

$$v_{in} = v_{gs} + i_d R \tag{21.28}$$

Since $g_m v_{gs} = i_d$, this equation can be written as

$$v_{in} = \frac{i_d}{g_m} + i_d R = i_d\left[\frac{1}{g_m} + R\right] \tag{21.29}$$

We can look at this in a simplified manner; that is, $\frac{1}{g_m}$ is the resistance looking into the source of a MOSFET, while R is the resistance connected from the source of the MOSFET to ground.

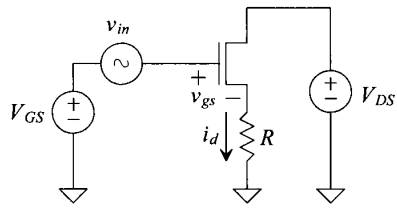


Figure 21.13 Circuit used to determine the resistance looking into the source of a MOSFET.

Noise Performance of the CS Amplifier with Gate-Drain Load

Figure 21.14 shows the CS amplifier of Fig. 21.1a with each MOSFET's model for noise power spectral density included (see Eq. [9.66]). Remembering that we only add the power from each noise source, we can write the output noise power spectral density as

$$V^2_{onoise}(f) = \left(\frac{1}{g_{m2}}\right)^2 \cdot \left(I^2_{M1} + I^2_{M2}\right) \tag{21.30}$$

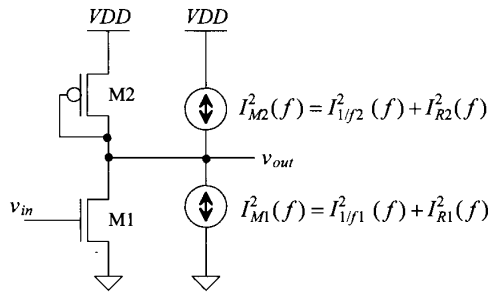


Figure 21.14 Noise modeling in a CS amplifier.

IPR2023-00698
Theta EX2019

The input-referred noise power spectral density is then

$$V_{inoise}^2(f) = \frac{V_{onoise}^2(f)}{A_v^2} = \left(\frac{g_{m2}}{g_{m1}}\right)^2 \cdot \left(\frac{1}{g_{m2}}\right)^2 \cdot \left(I_{M1}^2 + I_{M2}^2\right) \tag{21.31}$$

Minimizing the input-referred noise is accomplished by making g_{m1} large. Using the expressions in Eq. (9.66), we also see that minimizing g_{m2} reduces the input-referred noise. This is equivalent to saying large A_v results in minimum input-referred noise.

21.1.2 The Source Follower (Common-Drain Amplifier)

Source-follower configurations using a gate-drain-connected load are shown in Fig. 21.15. The drain of the amplifying device, that is, the MOSFET *not* gate-drain-connected (the load), is common to both the input and the output, so this topology is often called a common-drain amplifier. Setting the bias current in this topology can be challenging. In the following we don't concern ourselves with this issue.

A source follower implemented in CMOS has an asymmetric drive capability; that is, the ability of the follower to source current is not equal to the ability to sink current for a given bias condition and AC input signal. The small-signal gain of the NMOS source follower shown in Fig. 21.15 is simply determined by a voltage divider between the resistance looking into the source of M2 ($1/g_{m2}$) with the resistance of the gate-drain-connected load, M1 ($1/g_{m1}$). The gain of the **source follower amplifier** is

$$A_v = \frac{v_{out}}{v_{in}} = \frac{\text{resistance connected to the source}}{\text{resistance connected to the source + resistance looking into the source}} \tag{21.32}$$

or for the NMOS source follower in Fig. 21.15

$$v_{in} = \overbrace{v_{gs2}}^{i_d/g_{m2}} + \overbrace{v_{out}}^{i_d/g_{m1}} \text{ where } (v_{gs1} = v_{out}) \text{ and so } v_{out} = v_{in} \cdot \frac{\frac{1}{g_{m1}}}{\frac{1}{g_{m1}} + \frac{1}{g_{m2}}} \tag{21.33}$$

The output resistance of the source follower shown in this figure is

$$R_{out} = \frac{1}{g_{m1}} \| \frac{1}{g_{m2}} \tag{21.34}$$

Note that M2 can only source current, while M1 can only sink current, and the gain of the source follower is always less than one.

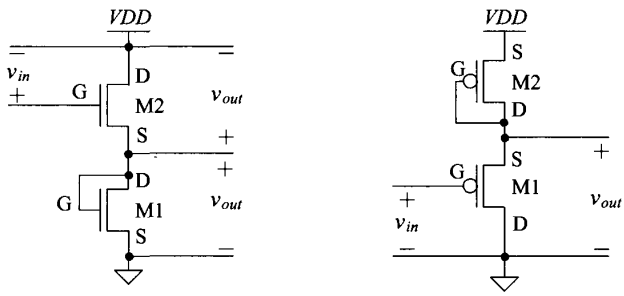


Figure 21.15 Source followers (common drain) using gate-drain loads.

IPR2023-00698
Theta EX2019

21.1.3 Common Gate Amplifier

The common gate amplifier with gate-drain load is shown in Fig. 21.16. The input resistance of this amplifier is simply the resistance looking into the source of M1, or

$$R_{in} = \frac{1}{g_{m1}} \tag{21.35}$$

The gain of this **common-gate amplifier** is

$$A_v = \frac{v_{out}}{v_{in}} = \frac{-i_d \cdot \frac{1}{g_{m2}}}{-v_{gs1}} = \frac{-i_d \cdot \frac{1}{g_{m2}}}{-i_d \cdot \frac{1}{g_{m1}}} = \frac{\text{resistance in the drain}}{\text{resistance in the source}} = \frac{\frac{1}{g_{m2}}}{\frac{1}{g_{m1}}} \tag{21.36}$$

or noting the positive gain

$$A_v = \frac{g_{m1}}{g_{m2}} = \sqrt{\frac{KP_n}{KP_p} \cdot \frac{W_1 L_2}{W_2 L_1}} \tag{21.37}$$

which is the same form as the gain for the CS amplifier with gate-drain load.

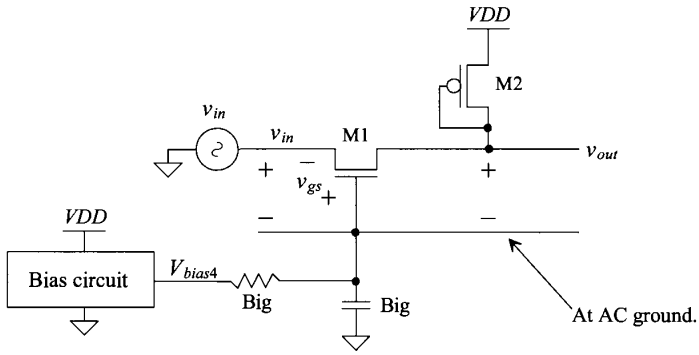


Figure 21.16 The common gate amplifier.

21.2 Current Source Loads

In the last section we talked about the gate-drain MOSFET part of a current mirror. In this section we talk about the current-source side of the mirror. The current source load provides an amplifier with the largest possible load resistance and thus the largest gain.

21.2.1 Common-Source Amplifier

Consider the CS amplifier with current source load shown in Fig. 21.17. The MOSFET M1 is the common-source component of the amplifier, while the MOSFET M2 is the current source load. The DC transfer characteristics are also seen in this figure. The amplifying portion of the curve occurs when the output is not close to *VDD* or ground (both M1 and M2 are saturated). The slope of the line when both transistors are saturated corresponds to the gain of the amplifier.

IPR2023-00698
Theta EX2019

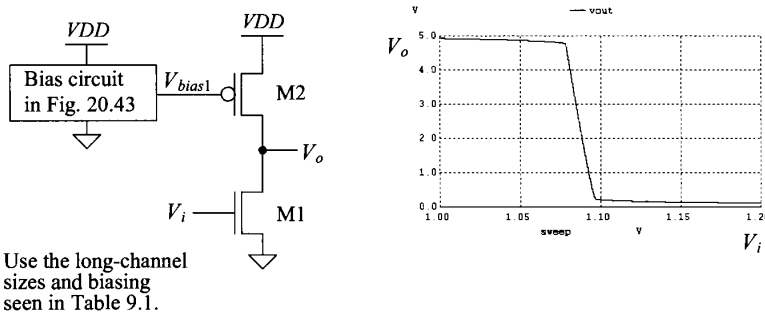


Figure 21.17 Common-source amplifier with current source load.

Class A Operation

We know from Table 9.1, and the bias circuit in Fig. 20.43, that M2 in Fig. 21.17 is biased to behave like a 20 μA current source. This means that if M1 shuts off, the maximum current we can supply to the output is 20 μA: this is important to understand. The CS amplifier in Fig. 21.17 is called a *class A* amplifier because, for proper operation, both MOSFETs are always conducting a current. In a *class B* amplifier, only one MOSFET is conducting a current at a given time. In *class AB* amplifiers (which we'll talk about later), both MOSFETs or a single MOSFET conduct a current at a given time.

Example 21.5

Estimate the maximum rate the load capacitor in Fig. 21.18 can be charged (estimate the slew rate across the load capacitor). Verify the estimate with SPICE.

The maximum current that the amplifier can source is 20 μA. The slew rate across the load is calculated using

$$I = C_L \cdot \frac{dV_{out}}{dt} \text{ or slew rate} = \frac{dV_{out}}{dt} = \frac{I}{C_L} \tag{21.38}$$

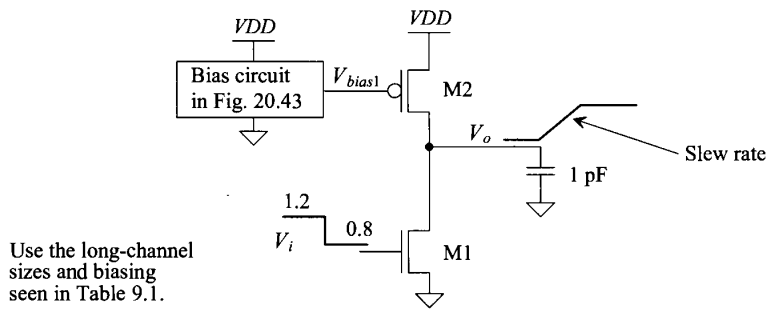


Figure 21.18 Slew rate limitations in a class A amplifier.

IPR2023-00698
Theta EX2019

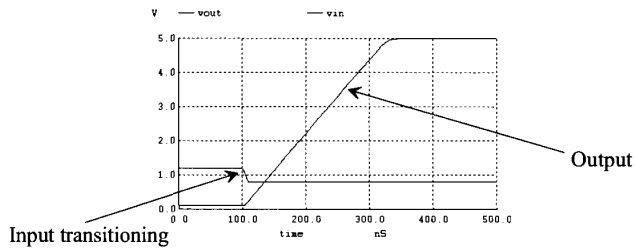


Figure 21.19 Verifying the results in Ex. 21.5

or

$$\frac{dV_{out}}{dt} = \frac{20\ \mu A}{1\ pF} = 20\ V/\mu s$$

For the output to transition 5 V requires 250 ns. ■

Small-Signal Gain

If we follow the intuitive method discussed in the last section, Eq. (21.1), the voltage gain of a common-source amplifier is the total parallel resistance at the drain of M1 divided by the resistance in the source of M1, or

$$A_v = \frac{v_o}{v_i} = -\frac{r_{o1} \| r_{o2}}{\frac{1}{g_{m1}}} = -g_{m1}(r_{o1} \| r_{o2}) = -\frac{g_{m1}}{g_{o1} + g_{o2}} \qquad (21.39)$$

where $r_{o1} = 1/g_{o1}$.

Open Circuit Gain

The gain of the amplifier can be increased by using a cascode current load in place of M2. The resistance looking into the drain of the cascode current source/load is much larger than the output resistance of M1. This situation is sometimes referred to as the *open circuit gain* of a common-source amplifier (see the specification in Tables 9.1 and 9.2). This specification indicates the maximum possible gain attainable with a single MOSFET (the MOSFET's load is its own output resistance r_o). The open circuit gain of M1 can be written as

$$\text{Open circuit gain} = -\frac{r_{o1}}{\frac{1}{g_{m1}}} = -g_{m1} r_{o1} = -\frac{\sqrt{2\beta_1 I_D}}{I_D \lambda_1} \qquad (21.40)$$

High-Impedance and Low-Impedance Nodes

In Fig. 21.17, the output node at the drains of M1 and M2 is termed a *high-impedance node*, that is, a node with only drain connections. The effective resistance at this node to ground is $r_{o1} \| r_{o2}$. A node connected to the source of a MOSFET or a MOSFET with its drain and gate connected is termed a *low-impedance node*. The small-signal resistance looking into the source of a MOSFET and the small-signal resistance of the gate-drain (diode)-connected MOSFET are both $1/g_m$. High-impedance nodes usually have a low-frequency pole associated with them that limits the speed of the amplifier.

IPR2023-00698
Theta EX2019

Frequency Response

Figure 21.20 shows the AC frequency response circuit for the CS amplifier with current source load in Fig. 21.17. *VDD* in this figure is drawn as an AC ground. The source-gate capacitance of M2 isn't shown because both sides of it are at AC ground (DC voltages), so it doesn't affect the frequency behavior of the circuit. The pole associated with the input node of the circuit is

$$f_{in} = \frac{1}{2\pi(C_{gs1} + C_{gd1}(1 + |A_v|)) \cdot R_s} \tag{21.41}$$

while the pole associated with the output node of the circuit is located at

$$f_{out} = \frac{1}{2\pi\left(C_{dg2} + C_{gd1}\left(1 + \frac{1}{|A_v|}\right)\right) \cdot r_{o1} || r_{o2}} \tag{21.42}$$

where $A_v = -g_{m1} \cdot r_{o1} || r_{o2}$ (see Eq. [21.39]). Using Eq. (21.42) to calculate the output pole, results in a *wrong estimate* when A_v is large. As we'll see in a moment, the fact that A_v decreases above f_{in} causes v_{in} to "flatten out" because the Miller capacitance, $C_{gd1}(1 + |A_v|)$, gets smaller. The effect of v_{in} ceasing to decrease causes the output pole to appear much larger than what is predicted by Eq. (21.42). This important phenomenon is called *pole splitting*.

The zero in the transfer function (see Eq. (21.15)) is located at

$$f_z = \frac{g_{m1}}{2\pi C_{gd1}} \tag{21.43}$$

The transfer function of the amplifier is then estimated as

$$A_v(f) = -g_{m1} \cdot (r_{o1} || r_{o2}) \cdot \frac{\left(1 - j\frac{f}{f_z}\right)}{\left(1 + j\frac{f}{f_{in}}\right)\left(1 + j\frac{f}{f_{out}}\right)} \tag{21.44}$$

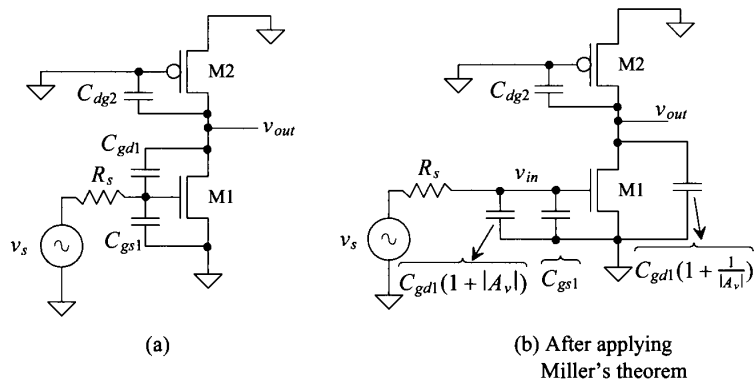


Figure 21.20 Frequency response of the CS amplifier with current source load.

IPR2023-00698
Theta EX2019

We can then write the magnitude as

$$|A_v(f)| = \frac{g_{m1} \cdot (r_{o1} || r_{o2}) \cdot \sqrt{1 + \left(\frac{f}{f_z}\right)^2}}{\left(\sqrt{1 + \left(\frac{f}{f_{in}}\right)^2}\right) \cdot \left(\sqrt{1 + \left(\frac{f}{f_{out}}\right)^2}\right)} \tag{21.45}$$

and the phase shift through the amplifier as

$$\angle A_v = 180 - \tan^{-1}\left(\frac{f}{f_z}\right) - \tan^{-1}\left(\frac{f}{f_{in}}\right) - \tan^{-1}\left(\frac{f}{f_{out}}\right) \tag{21.46}$$

Example 21.6

Determine the gain, magnitude, and phase shift of the amplifier shown in Fig. 21.21. Using an AC SPICE simulation, verify your hand calculations.

For the AC small-signal analysis, the big resistor appears like an open and the big capacitor a short. These components are added to ensure that M1 is biased to sink the current supplied by M2.

The low-frequency gain of the circuit is, from Eq. (21.39) and Table 9.1,

$$|A_v| = (150\ \mu A/V) \cdot (5M\Omega || 4M\Omega) = 333\ V/V \rightarrow 50\ dB$$

The input capacitance is

$$C_{in} = (C_{gs1} + C_{gd1}(1 + |A_v|)) = (23.3 + 2 \cdot 333)fF = 691\,fF\,! \tag{21.47}$$

The output capacitance is

$$C_{out} = \left(C_{dg2} + C_{gd1}\left(1 + \frac{1}{|A_v|}\right)\right) \approx 6 + 2\,fF = 8\,fF \tag{21.48}$$

Using Eqs. (21.41)–(21.43), we get $f_{in} = 2.3$ MHz, $f_{out} = 9$ MHz, and $f_z = 11.9$ GHz. Figure 21.22 shows the simulation results. As the discussion following Eq. (21.42) indicated, the estimate for f_{out} is wrong. To determine the correct value, let's discuss pole splitting (see Eq. [21.62]). ■

Use the long-channel sizes and biasing seen in Table 9.1.

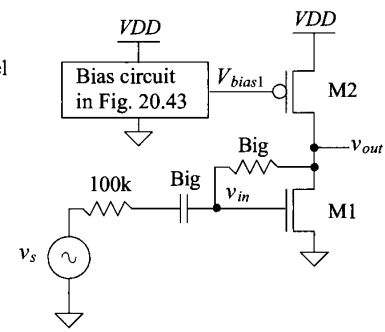


Figure 21.21 Amplifier used in Ex. 21.6.

IPR2023-00698
Theta EX2019

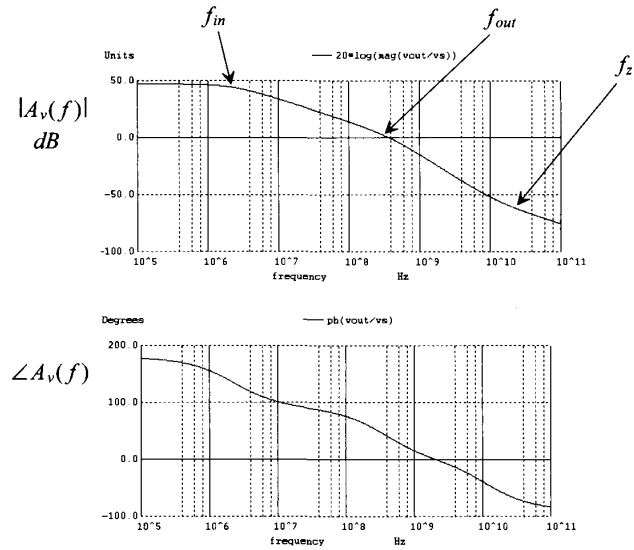


Figure 21.22 Frequency response of the amplifier in Fig. 21.21.

Pole Splitting

As discussed after Eq. (21.42), using Miller's theorem to calculate the location of the output pole when the gain of the amplifier is large results in error. Clearly, as just demonstrated in Ex. 21.6, the calculated value of f_{out} is 9 MHz while the SPICE simulation in Fig. 21.22 shows that f_{out} is approximately 200 MHz. The source of this discrepancy can be traced to the fact that above f_{in} the gain of the amplifier, A_v, decreases causing the effective input capacitance of the amplifier to decrease. As seen in Fig. 21.23, this causes the input voltage to "flatten out" instead of continuing to fall at –20 dB/decade. The result is the second, or output pole, appears to split from (leave) the first,

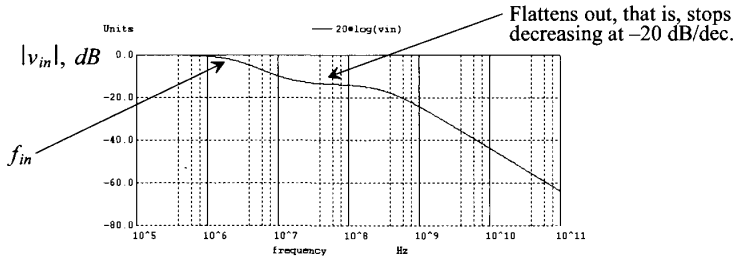


Figure 21.23 Showing how the amplifier's, in Fig. 21.21, input voltage flattens out as the gain decreases.

IPR2023-00698
Theta EX2019

or input pole, and move to a higher frequency. Note that if we resimulate the amplifier in Fig. 21.21 with the source resistance set to zero ohms (so there isn't an input pole), pole splitting isn't present and Eq. (21.42) can be used to calculate the location of the output pole.

To characterize this effect, let's use the circuit in Fig. 21.7 with the addition of a source resistance, R_s, as seen in Fig. 21.24. In this figure, $R_o = r_{o1} \| r_{o2}$ and $C_o = C_{dg2}$. From Eq. (21.14), we can write

$$\frac{v_{out}}{v_{in}} = -g_{m1}R_o \cdot \frac{1 - j\omega\frac{C_{gd1}}{g_{m1}}}{1 + j\omega(C_{gd1} + C_o) \cdot R_o} \tag{21.49}$$

At the gate of M1, we can write

$$\frac{v_{in} - v_s}{R_s} + \frac{v_{in}}{1/j\omega C_{gs1}} + \frac{v_{in} - v_{out}}{1/j\omega C_{gd1}} = 0 \tag{21.50}$$

or

$$v_{in} = \frac{\frac{v_s}{R_s} + v_{out} \cdot j\omega C_{gd1}}{\frac{1}{R_s} + j\omega C_{gs1} + j\omega C_{gd1}} \tag{21.51}$$

Substituting Eq. (21.51) into Eq. (21.49) gives

$$v_{out} = -g_{m1}R_o \cdot \frac{1 - j\omega\frac{C_{gd1}}{g_{m1}}}{1 + j\omega(C_{gd1} + C_o) \cdot R_o} \cdot \frac{\frac{v_s}{R_s} + v_{out} \cdot j\omega C_{gd1}}{\frac{1}{R_s} + j\omega C_{gs1} + j\omega C_{gd1}} \tag{21.52}$$

or, with $s = j\omega$,

$$\frac{v_{out}}{v_s} = \frac{-g_{m1}R_o \cdot \left(1 - s\frac{C_{gd1}}{g_{m1}}\right)}{s^2[R_oR_s(C_{gd1}C_{gs1} + C_oC_{gs1} + C_oC_{gd1})] + s[(C_{gd1} + C_o)R_o + (C_{gs1} + C_{gd1})R_s + C_{gd1}g_{m1}R_oR_s] + 1} \tag{21.53}$$

The zero is still, as indicated in Eqs. (21.15) and (21.43), located at

$$f_z = \frac{g_{m1}}{2\pi C_{gd1}} \tag{21.54}$$

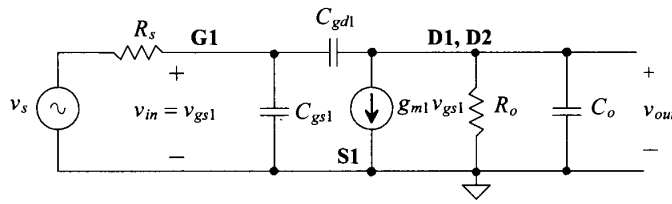


Figure 21.24 Circuit used to calculate the transfer function of the amplifier in Fig. 21.21.

IPR2023-00698
Theta EX2019

At low frequencies the denominator is approximately (the coefficient of the s^2 term is small)

$$1 + j\omega \cdot [(C_{gd1} + C_o)R_o + (C_{gs1} + C_{gd1})R_s + C_{gd1}g_{m1}R_oR_s] \tag{21.55}$$

and so the low-frequency pole, using $|A_v| = g_{m1}(r_{o1} || r_{o2}) = g_{m1}R_o$, $C_o = C_{dg2}$, is located at

$$f_1 \approx \frac{1}{2\pi[(C_{gd1} + C_{dg2}) \cdot r_{o1} || r_{o2} + (C_{gs1} + C_{gd1}(1 + |A_v|)) \cdot R_s]} \tag{21.56}$$

If $C_{gd1}(1 + |A_v|)$ is much larger than the other capacitances,

$$f_1 \approx \frac{1}{2\pi C_{gd1}(1 + |A_v|) \cdot R_s} \tag{21.57}$$

Notice the similarity to Eq. (21.41), where the pole is associated with the input of the amplifier. Further notice that if we had used this result in Ex. 21.6, the input pole would have changed very little. Also notice that if $R_s = 0$, then

$$f_1 \approx \frac{1}{2\pi(C_{gd1} + C_{dg2}) \cdot r_{o1} || r_{o2}} \tag{21.58}$$

Notice the similarity to Eq. (21.42), where the pole is associated with the output of the amplifier. To determine the location of the second root of the denominator (the second pole location, f_2), let's factor Eq. (21.55) from the denominator of Eq. (21.53) or

$$[1 + s[(C_{gd1} + C_o)R_o + (C_{gs1} + C_{gd1})R_s + C_{gd1}g_{m1}R_oR_s]] \cdot \left(1 + \frac{s^2[R_oR_s(C_{gd1}C_{gs1} + C_oC_{gs1} + C_oC_{gd1})]}{s[(C_{gd1} + C_o)R_o + (C_{gs1} + C_{gd1})R_s + C_{gd1}g_{m1}R_oR_s] + 1}\right) \tag{21.59}$$

where this equation is in the form

$$\left(1 + j \cdot \frac{f}{f_1}\right) \cdot \left(1 + j \cdot \frac{f}{f_2}\right) \tag{21.60}$$

Note again that if $R_s = 0$, the second term in Eq. (21.59) goes to unity. Looking at the second term of Eq. (21.59), let's divide the numerator and denominator by sR_oR_s

$$1 + j \cdot \frac{2\pi f \cdot (C_{gd1}C_{gs1} + C_oC_{gs1} + C_oC_{gd1})}{(C_{gd1} + C_o)/R_s + (C_{gs1} + C_{gd1})/R_o + C_{gd1}g_{m1} + 1/sR_oR_s} \tag{21.61}$$

In any practical MOSFET amplifier, $g_m >> 1/r_o$ (if not, then the magnitude of the open circuit gain, Eq. (21.40), is too small for the MOSFET to be of practical value as an amplifying device). Also, the R_s is assumed to be large (if not, then f_2 is not used), that is, the amplifier's transfer function has a single-pole response, where the pole is associated with the output of the amplifier as seen in Eq. (21.58). We can therefore write

$$f_2 \approx \frac{g_{m1}C_{gd1}}{2\pi \cdot (C_{gd1}C_{gs1} + C_oC_{gs1} + C_oC_{gd1})} \tag{21.62}$$

Calculating the location of the second pole in Ex. 21.6 with this equation results in $f_2 =$ 240 MHz (which is very close to the simulated location).

IPR2023-00698
Theta EX2019

Why the name pole splitting? If we increase the effective size of C_{gd1} by placing a capacitor, C_c, from the amplifier's input to its output (so that C_c is in parallel with C_{gd1}, that is, the effective value is $C_c + C_{gd1}$), then, as Eq. (21.56) shows, the low-frequency location of the pole, f_1, decreases. At the same time, increasing the effective size of C_{gd1} causes the location of the high-frequency pole f_2, as seen in Eq. (21.62), to increase. Thus the name *pole splitting*.

Pole Splitting Summary

Because of the ubiquity of pole splitting in CMOS amplifier design, let's summarize our discussion using a generic model, Fig. 21.25. In terms of the parameters seen in this figure the location of the zero is now

$$f_z = \frac{g_{m2}}{2\pi \cdot C_c} \tag{21.63}$$

The first pole is located at

$$f_1 \approx \frac{1}{2\pi[(C_c + C_2) \cdot R_2 + (C_1 + C_c(1 + g_{m2}R_2)) \cdot R_1]} \tag{21.64}$$

For large gain, $g_{m2}R_2$, we can simplify this to

$$f_1 \approx \frac{1}{2\pi g_{m2} R_2 R_1 C_c} \tag{21.65}$$

The second pole is located at

$$f_2 \approx \frac{g_{m2} C_c}{2\pi \cdot (C_c C_1 + C_1 C_2 + C_c C_2)} \tag{21.66}$$

The transfer function is then written as

$$A_v(f) = \frac{v_{out}(f)}{v_s(f)} = g_{m1}R_1 g_{m2}R_2 \cdot \frac{\left(1 - j \cdot \frac{f}{f_z}\right)}{\left(1 + j \cdot \frac{f}{f_1}\right) \cdot \left(1 + j \cdot \frac{f}{f_2}\right)} \tag{21.67}$$

where $A_v = g_{m1}R_1 g_{m2}R_2$ (the gain is now positive because of the defined direction of the current source in Fig. 21.25).

Again note that if C_c is increased, f_1 moves downwards and f_2 moves upwards (pole splitting). Also note that if the capacitor, C_c, is made large (quite common), then f_1 is much lower than the location of the other pole or the zero (the pole associated with f_1 is

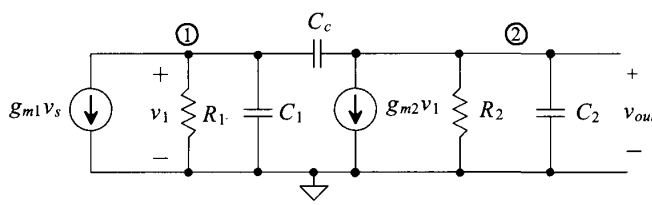


Figure 21.25 Generic model used to estimate bandwidth in a CMOS amplifier.

IPR2023-00698
Theta EX2019

said to be the *dominant pole*). The amplifier's transfer function, with the help of Eq. (21.65), can be written as

$$A_v(f) = \frac{v_{out}(f)}{v_s(f)} \approx \frac{g_{m1}R_1 g_{m2}R_2}{\left(1 + j \cdot \frac{f}{f_1}\right)} = \frac{g_{m1}R_1 g_{m2}R_2}{(1 + j \cdot 2\pi f \cdot g_{m2}R_2 R_1 C_c)} \tag{21.68}$$

This equation is in the form

$$A_v(f) = \frac{v_{out}(f)}{v_s(f)} = \frac{A_{DC}}{1 + j \cdot \frac{f}{f_{3dB}}} \tag{21.69}$$

where

$$A_{DC} = g_{m1}R_1 g_{m2}R_2 \text{ and } f_{3dB} = \frac{1}{2\pi g_{m2}R_2 R_1 C_c} \tag{21.70}$$

The magnitude and phase responses of a generic CMOS amplifier are seen in Fig. 21.26. Note that at frequencies much larger than the amplifier's 3 dB frequency, that is, $f/f_{3dB} >> 1$, Eq. (21.68) can be written as

$$A_v(f) = \left|\frac{v_{out}(f)}{v_s(f)}\right| \approx \frac{g_{m1}}{2\pi f \cdot C_c} \tag{21.71}$$

To determine the frequency, f_{un}, where the transfer function is unity, we set Eq. (21.71) to one and solve to yield

$$f_{un} = \frac{g_{m1}}{2\pi C_c} \tag{21.72}$$

We've derived some very useful equations, so let's apply them in some examples.

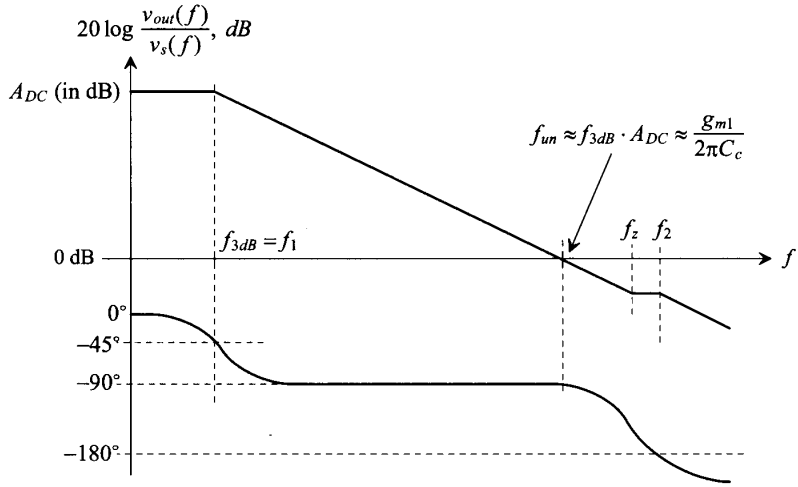


Figure 21.26 Magnitude and phase responses of a generic CMOS amplifier.

IPR2023-00698
Theta EX2019

Example 21.7
Repeat Ex. 21.6 using the model in Fig. 21.25.

Using the model parameters in Fig. 21.25 and looking at Figs. 21.20a and 21.21, we can write (with the help of Table 9.1)

$$R_1 = R_s = 100\text{k},\ C_1 = C_{gs1} = 23.3\,fF,\ C_c = C_{gd1} = 2\,fF,$$

$$C_2 = C_{dg2} = 6\,fF,\ R_2 = r_{o1} \| r_{o2} = 2.22\ \text{M}\Omega$$

$$g_{m2}\ (\text{in Fig. 21.25}) = g_{m1}\ (\text{in Fig. 21.21}) = 150\ \mu A/V$$

$$g_{m1}\ (\text{in Fig. 21.25}) = 1/R_s\ (\text{in Fig. 21.21}) = 10\ \mu A/V$$

Using Eq. (21.67) or (21.70), we have $|A_v| = 333\ V/V \rightarrow 50\ dB$. Using Eq. (21.63), we get $f_z = 11.9\ GHz$; Eq. (21.65), we get $f_1 = 2.3\ MHz$; and finally using Eq. (21.66), we get $f_2 = 240\ MHz$. These calculations should be compared to the simulation results seen in Fig. 21.22. ■

Example 21.8
Repeat Ex. 21.7 if R_s is zero. Verify the hand calculations using SPICE.

Again we use the model seen in Fig. 21.25. From Ex. 21.7, $g_{m1}R_1$ is unity (and so R_s doesn't affect the gain). Further, the location of the zero doesn't change. Since R_s is zero ($R_1 = 0$ so the current source, $g_{m1}v_s$, and the resistor, R_1 in Fig. 21.25 are replaced with a voltage source, v_s), the location of f_2 is moved to an infinite frequency (as seen in Eq. [21.59]). Using Eq. (21.64), we get

$$f_1 \approx \frac{1}{2\pi(C_{gd1} + C_{dg2}) \cdot r_{o1} \| r_{o2}} = \frac{1}{2\pi(8\,fF) \cdot 2.22M} = 9\ MHz$$

The simulation results are seen in Fig. 21.27. Note the small perturbation in the frequency response above f_1. The response is not perfectly decreasing at –20 dB/decade. This imperfection can be traced to the bias voltage, V_{bias1}, used on the gate of M2. It is not a perfect AC ground as was assumed in our hand calculations. By placing a capacitor on the gate of M2 to ground, the variation in V_{bias1} is reduced and the amplifier behaves as expected. This is an *important point.* It is often a good idea to "bypass" the bias voltages to AC ground to reduce their impedance in practical design. This also keeps circuits that are biased from the same bias circuit from interacting. ■

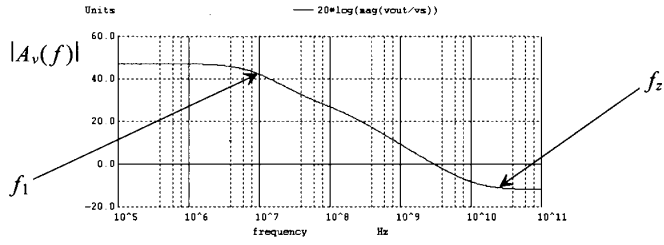


Figure 21.27 Simulation results for Ex. 21.8.

IPR2023-00698
Theta EX2019

Example 21.9
Determine the frequency response of the amplifier seen in Fig. 21.28. Verify the hand calculations using simulations.

Use the long-channel sizes and biasing seen in Table 9.1.

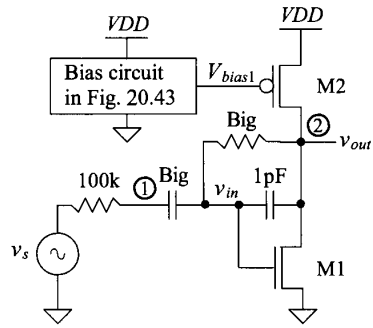


Figure 21.28 Amplifier used in Ex. 21.9.

The only thing that is different between this amplifier and the one in Ex. 21.7 is the value of C_c. In Fig. 21.28

$$C_c = 1\ pF + C_{gd1} \approx 1\ pF$$

We can then write, using Eq. (21.65)

$$f_1 \approx \frac{1}{2\pi \cdot (150\ \mu A/V) \cdot 100k \cdot 2.2MEG \cdot 1pF} = 4.8\ kHz$$

using Eq. (21.66)

$$f_2 \approx \frac{(150\ \mu A/V) \cdot 1p}{2\pi \cdot (1p \cdot 23.3f + 23.3f \cdot 6f + 1p \cdot 6f)} = 811\ MHz$$

At the risk of stating the obvious, these two poles are split apart by a great distance with the addition of the 1 pF capacitor. The zero, from Eq. (21.63), is located at

$$f_z = \frac{150\ \mu A/V}{2\pi \cdot 1p} = 23\ MHz$$

The unity-gain frequency is found using Eq. (21.72) with $R_s = 1/g_{m1}$.

$$f_{un} = \frac{1}{2\pi R_s C_c} = 1.59\ MHz$$

The simulation results are seen in Fig. 21.29. The value of the gain is 47 dB (we calculated 50 dB). The value of the first pole is approximately 6 kHz (we calculated 4.8 kHz). The small discrepancies are most likely the result of differences in the actual circuit parameters compared to the values we used in the formulas. ■

IPR2023-00698
Theta EX2019

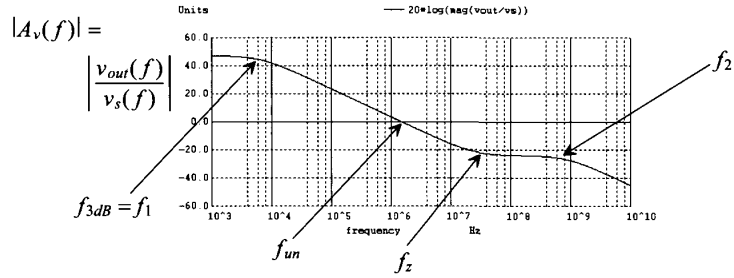


Figure 21.29 Frequency response of the amplifier in Fig. 21.28.

Example 21.10

Estimate the frequency response of the amplifier seen in Fig. 21.30. Verify the estimates with SPICE.

MA's gate and drain, at DC, are at the same potential. This biases MB to sink the right amount of current from MD.

Using the model in Fig. 21.25, we see that the source voltage, v_s, modulates the drain current in MA. We can therefore write

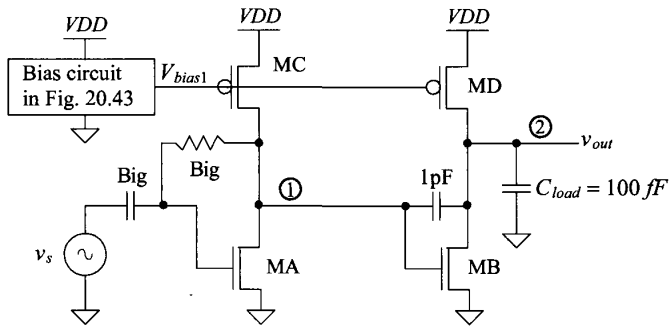
$$g_{mA}v_s = g_{m1}v_s \rightarrow g_{m1} \text{ (in Fig. 21.25)} = g_{mA} \text{ (in Fig. 21.30)} = 150\ \mu A/V$$

$$g_{m2} \text{ (in Fig. 21.25)} = g_{mB} \text{ (in Fig. 21.30)} = 150\ \mu A/V$$

$$R_1 = r_{oA} \| r_{oC} = \ R_2 = r_{oB} \| r_{oD} = 2.22\ M\Omega$$

To determine the capacitance on nodes 1 and 2, we can use Fig. 21.31. At node 1

$$C_1 = C_{gdA} + C_{dgC} + C_{gsB} = 31.3\ fF$$



Use the long-channel sizes and biasing seen in Table 9.1.

Figure 21.30 Amplifier used in Ex. 21.10.

IPR2023-00698
Theta EX2019

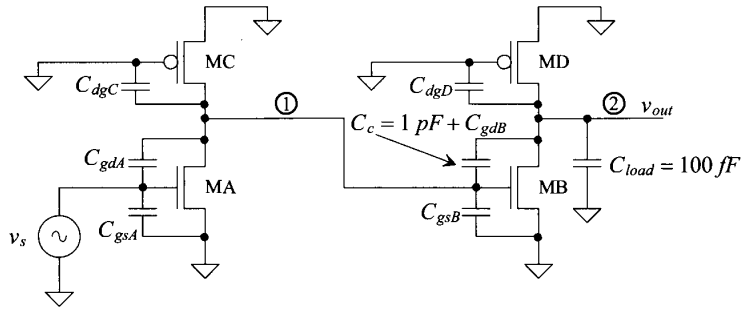


Figure 21.31 The capacitances for the amplifier in Fig. 21.30.

where C_{gdA} is the output Miller capacitance of MA ($C_{gdA} \approx C_{gdA}\left(1+\frac{1}{|A|}\right)$ when the gain of MA is large). At node 2

$$C_2 = C_{load} + C_{dgD} = 106\,fF$$

The value of C_c is

$$C_c = 1\,pF + C_{gdB} \approx 1\,pF \text{ (just the added capacitor)}$$

The gain of the topology is estimated as

$$A_{DC} = g_{m1}R_1 g_{m2}R_2 = (150\ \mu A/V)^2(2.22\ M\Omega)^2 = 110,889\ V/V \rightarrow 101\ dB$$

The poles and zero can be calculated using Eqs. (21.63), (21.65), and (21.66)

$$f_z = \frac{150\ \mu A/V}{2\pi \cdot 1.002\,pF} = 23\ MHz$$

$$f_1 \approx \frac{1}{2\pi \cdot (150\ \mu A/V)(2.22\ M\Omega)^2(1.002\,pF)} = 214\ Hz$$

$$f_2 \approx \frac{(150\ \mu A/V) \cdot 1.002}{2\pi \cdot (1.002 \cdot 0.0313 + 0.0313 \cdot 0.106 + 1.002 \cdot 0.106)\,pF} = 170\ MHz$$

The location of the unity gain frequency is calculated using Eq. (21.72) as 23 MHz. The simulation results are seen in Fig. 21.32. ■

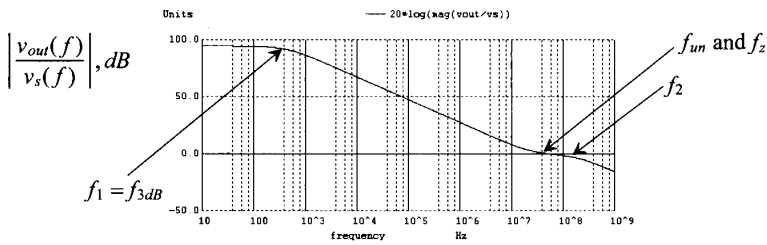


Figure 21.32 Frequency response of the amplifier in Fig. 21.30.

IPR2023-00698
Theta EX2019

Canceling the RHP Zero

Notice in Ex. 21.10 that the unity-gain frequency (the frequency where the gain is one) is equal to the frequency of the zero. As mentioned earlier, the input of the amplifier feeds directly through C_c to the amplifier's output without the phase inversion. If the gain of the amplifier is larger than one (or even slightly less than one), then the lack of inversion, when the amplifier uses feedback, can result in an unstable amplifier. The amplifier's output signal feeds back and adds to its input signal without the inversion through the amplifier. To avoid this situation, we might add an amplifier in series with C_c that allows the output to feed back to the input through C_c but not vice versa, Fig. 21.33a. A simpler solution is to add a resistor in series with the capacitor, Fig. 21.33b, to attenuate the higher frequency signals (where the zero occurs) and push the zero out to a higher frequency. Adding the resistor moves the zero to a frequency (see Eq. [21.63])

$$f_z = \frac{1}{2\pi \cdot C_c \cdot \frac{1}{g_{m2}}} \xrightarrow{\text{With a resistor}} f_z = \frac{1}{2\pi \cdot C_c \cdot \left(\frac{1}{g_{m2}} - R_z\right)} \qquad (21.73)$$

If $R_z = 1/g_{m2}$, the zero disappears (is pushed to an infinite frequency). If $R_z > 1/g_{m2}$, the zero is pushed back into the LHP (phase shift is opposite from the poles). *Any practical design where pole splitting is used should include the zero-nulling resistor* R_z. Consider the following example.



Figure 21.33 Removing the zero in the amplifier's transfer function.

Example 21.11

Using a zero-nulling resistor, show that the location of the zero in the frequency response of the amplifier in Fig. 21.30 can be eliminated.

In this amplifier, $R_z = 1/g_{m2} = 6.5\ k\Omega$. By adding this resistor in series with the 1 pF capacitor, we get the frequency response seen in Fig. 21.34. Note that the zero is still seen but at a considerably higher frequency. ■

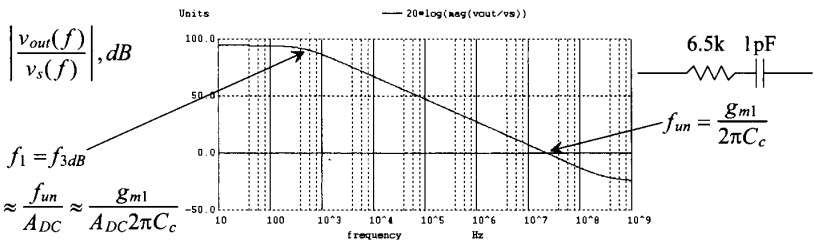


Figure 21.34 Pushing the zero in Fig. 21.32 to a higher frequency.

IPR2023-00698
Theta EX2019

Noise Performance of the CS Amplifier with Current Source Load

To determine the noise performance of the CS amplifier with current source load, we can use the schematic in Fig. 21.14 with the gate of M2 at AC ground. The output noise power spectral density is then

$$V_{onoise}^2(f) = (r_{o1} || r_{o2})^2 \cdot \left(I_{M1}^2 + I_{M2}^2\right) \tag{21.74}$$

The input-referred noise is then

$$V_{inoise}^2(f) = \frac{V_{onoise}^2(f)}{A_v^2} = \frac{1}{g_{m1}^2} \cdot \left(I_{M1}^2 + I_{M2}^2\right) \tag{21.75}$$

showing once again that to minimize the input-referred noise we need to make the transconductance of M1 large (or, in other words, make A_v large).

21.2.2 The Cascode Amplifier

Consider a CS amplifier with current source load as seen Fig. 21.17 but implemented using the short-channel CMOS process. Using the parameters in Table 9.2, the gain of the topology, Eq. (21.39), is only 16.7 (it was 333 using the long-channel process). To boost the gain of the single-stage amplifier and to eliminate, or more correctly to reduce, the Miller effect, consider the cascode amplifier seen in Fig. 21.35. The resistance looking into the drain of M3 is $g_{mp} \cdot r_{op}^2$ (see Table 20.1) with a value of 16.6 MΩ. The resistance looking into the drain of M2 is 4.2 MΩ (again from Table 20.1). The gain of the cascode amplifier is the resistance in the drain divided by the resistance in the source of the amplifying device (M1) or

$$\frac{v_{out}}{v_{in}} = -\frac{g_{mn}r_{on}^2 || g_{mp}r_{op}^2}{1/g_{mn}} = -g_{mn} \cdot R_{ocas} \tag{21.76}$$

Using the bias circuit from Fig. 20.47 and the sizes in Table 9.2 (and thus the small-signal parameters in this table).

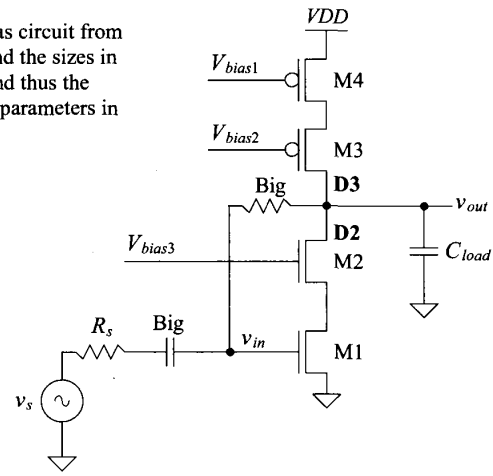


Figure 21.35 A cascode amplifier.

IPR2023-00698
Theta EX2019

where

$$R_{ocas} = g_{mn} r_{on}^2 || g_{mp} r_{op}^2 \tag{21.77}$$

Using the parameters from Table 9.2, the gain of the cascode amplifier in Fig. 21.35 is, roughly, 500.

Frequency Response

The resistance seen in the drain of M1 is approximately $r_o/2$ (see Fig. 21.37 and Eq. [21.82]). The gain from the gate of M1 (the amplifier's input) to the drain of M1 is then

$$A_{v1} = \frac{v_{d1}}{v_{in}} = \frac{-v_{gs2}}{v_{gs1}} \approx \frac{-i_d \cdot r_o/2}{i_d/g_{m1}} = -g_{m1} \cdot \frac{r_o}{2} \tag{21.78}$$

If we calculate the Miller capacitance, Eq. (21.5), we see that the lower the gain of M1 the less loading (from the Miller capacitance) on the input of the amplifier. From Table 9.2, g_{m1} = 150 μA/V and, from Fig. 9.32 (knowing M1 is biased close to a $V_{DS,sat}$ of 70 mV), r_o is 50k then $A_{v1} = -g_{m1} \cdot \frac{r_o}{2} = -3.75$ (let's use this result in the following example).

Example 21.12

Estimate the frequency response of the cascode amplifier in Fig. 21.35 if R_s is 100k and $C_L = 100\,fF$. Verify the hand calculations with simulations.

The input time constant is calculated as

$$\tau_{in} = R_s \cdot (C_{gs1} + (1 + |A_{v1}|) \cdot C_{gd1})$$

or

$$\tau_{in} = 100k \cdot (4.17\,fF + 4.75 \cdot 1.56\,fF) = 1.16\,ns$$

The pole associated with this input time constant is then

$$f_{in} = \frac{1}{2\pi\tau_{in}} = 137\,MHz$$

Since the load capacitance is large, $C_{load} >> C_{gd2} + C_{dg3}$, we can write

$$\tau_{out} = R_{ocas} \cdot (C_{load} + C_{gd2} + C_{dg3}) \approx g_{mn} r_{on}^2 || g_{mp} r_{op}^2 \cdot C_{load} = 335\,ns$$

and so the pole associated with the output node is located at

$$f_{out} = \frac{1}{2\pi\tau_{out}} = 475\,kHz$$

Again, the gain of the topology is 500 (54 dB). The simulation results are seen in Fig. 21.36. The magnitude response is

$$\left|\frac{v_{out}}{v_s}\right| = \frac{500}{\sqrt{1 + \left(\frac{f}{475kHz}\right)^2} \cdot \sqrt{1 + \left(\frac{f}{137\,MHz}\right)^2}}$$

and the phase response (in degrees) is

$$\angle\frac{v_{out}}{v_{in}} = 180 - \tan^{-1}\frac{f}{475\,kHz} - \tan^{-1}\frac{f}{137\,MHz} \quad \blacksquare$$

IPR2023-00698
Theta EX2019

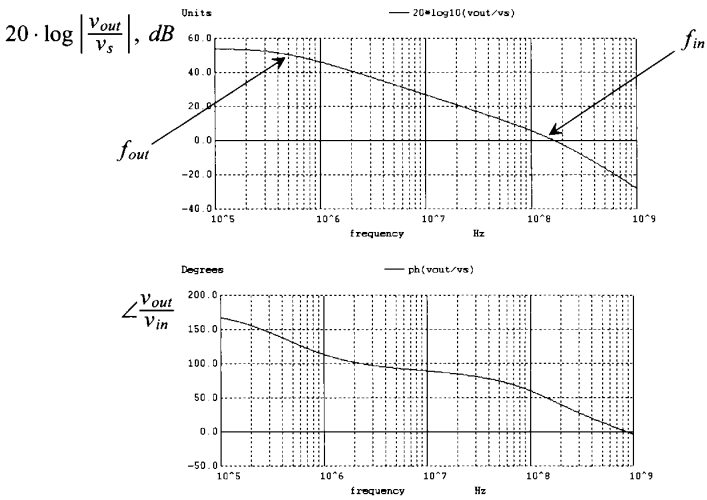


Figure 21.36 The simulation results for Ex. 21.12.

Class A Operation

The cascode amplifier in Fig. 21.35 is another example of a class A amplifier. When the input of the amplifier goes sufficiently negative, M1 shuts off. M3 and M4 are a current source and so the maximum rate that the load capacitance can be charged is given by Eq. (21.38). For the amplifier in Fig. 21.35, this is 10 μA/100 *fF* or 100 mV/ns.

Noise Performance of the Cascode Amplifier

The cascode's output noise power spectral density is given by

$$V_{onoise}^2(f) = \left(g_{mn} r_{on}^2 \| g_{mp} r_{op}^2\right)^2 \left(I_{M1}^2 + I_{M4}^2\right) \qquad (21.79)$$

where the noise contributions from M2/M3 are negligible, see Fig. 21.44. The input-referred noise power is then ($g_{m1} = g_{mn}$)

$$V_{inoise}^2(f) = \frac{V_{onoise}^2(f)}{A_v^2} = \frac{\left(I_{M1}^2 + I_{M4}^2\right)}{(g_{m1})^2} \qquad (21.80)$$

Again, maximizing the transconductance of the amplifying device (equivalent to saying maximizing the gain of the amplifier) reduces the input-referred noise.

Operation as a Transimpedance Amplifier

Figure 21.37 shows a transimpedance amplifier (current input and voltage output). Note that the AC voltages between the gates and sources of M1 and M4 are zero (both the gates and the sources are at AC ground). From the figure the input resistance is

$$R_{in} = \frac{-v_{in}}{i_{in}} = \frac{1 + \frac{R_{ocasp}}{r_{on}}}{g_{mn} + \frac{2}{r_{on}} + \frac{R_{ocasp}}{r_{on}^2}} \qquad (21.81)$$

IPR2023-00698
Theta EX2019

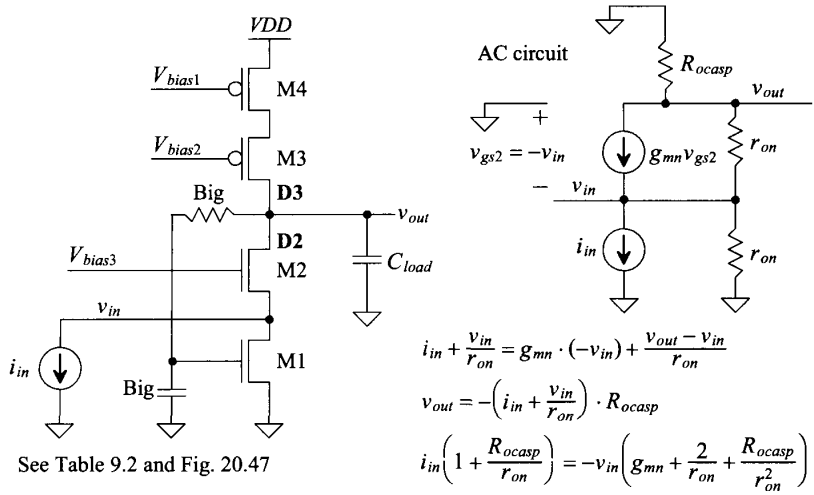


Figure 21.37 A transimpedance amplifier.

noting that if the drain of M2 is at AC ground or a low impedance then R_{in} is approximately $1/g_{mn}$. If $r_{on} \approx r_{op} \approx r_o$, $g_m = g_{mn} \approx g_{mp}$, and $R_{ocasp} \approx R_{ocasn} \approx g_m r_o^2$ then

$$R_{in} \approx \frac{\frac{R_{ocasp}}{r_{on}}}{2g_{mn}} \approx \frac{r_o}{2} \text{ and thus } v_{in} = -\frac{r_o}{2} \cdot i_{in} \tag{21.82}$$

To calculate the gain we can write

$$v_{out} = -\left(i_{in} + \frac{v_{in}}{r_{on}}\right) \cdot R_{ocasp} \rightarrow \frac{v_{out}}{i_{in}} = -\frac{R_{ocasp}}{2} \tag{21.83}$$

In terms of the NMOS and PMOS cascodes we can write

$$\frac{v_{out}}{i_{in}} = -g_{mn}r_{on}^2 \| g_{mp}r_{op}^2 = -R_{ocasn} \| R_{ocasp} \tag{21.84}$$

In other words the output voltage is simply the product of the input current with the cascode amplifier's output resistance.

21.2.3 The Common-Gate Amplifier

M2 in Fig. 21.37 is an example of a common-gate (CG) amplifier. We can redraw this circuit, as seen in Fig. 21.38 with a voltage source input, to show how the gate of the amplifying device, M2, is common to both the input and the output of the amplifier. Though the gate of M2 is at a DC voltage of V_{bias3}, we think of it as being at AC ground. M1 is simply an ordinary current source while M3 and M4 are a cascode current source load. The input resistance of this amplifier is given by Eq. (21.81); however, by connecting the input to a voltage source the output resistance and gain change. The voltage gain can be calculated by first writing

$$\frac{v_{out}}{R_{ocasp}} + g_{mn} \cdot (-v_{in}) + \frac{v_{out} - v_{in}}{r_{on}} = 0 \tag{21.85}$$

IPR2023-00698
Theta EX2019

See Table 9.2 and Fig. 20.47

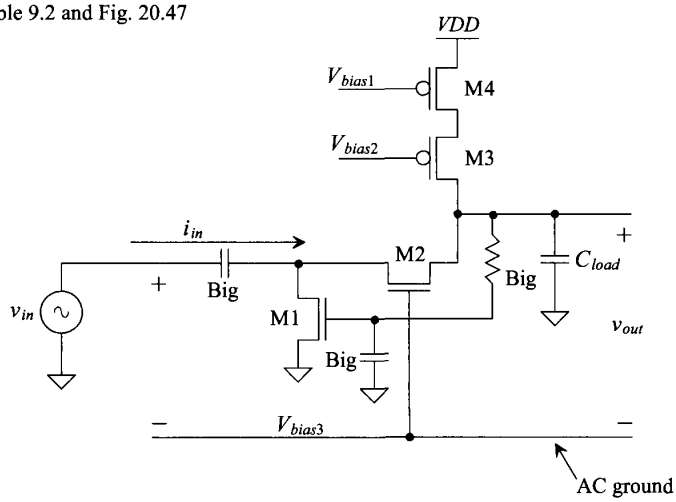


Figure 21.38 A common-gate amplifier.

or

$$A_v = \frac{v_{out}}{v_{in}} = \frac{g_{mn} + \frac{1}{r_{on}}}{\frac{1}{R_{ocasp}} + \frac{1}{r_{on}}} = \frac{R_{ocasp} || r_{on}}{\frac{1}{g_{mn}} || r_{on}} \approx g_{mn} \cdot r_{on} \qquad (21.86)$$

where, because the source of M2 is connected to a voltage source, the amplifier's output resistance is $R_{ocasp} || r_{on} \approx r_{on}$.

21.2.4 The Source Follower (Common-Drain Amplifier)

The source follower (SF) with current source load is seen in Fig. 21.39. Looking at the NMOS SF, we can write, for AC small signals,

$$v_{in} = v_{gs2} + v_{out} \qquad (21.87)$$

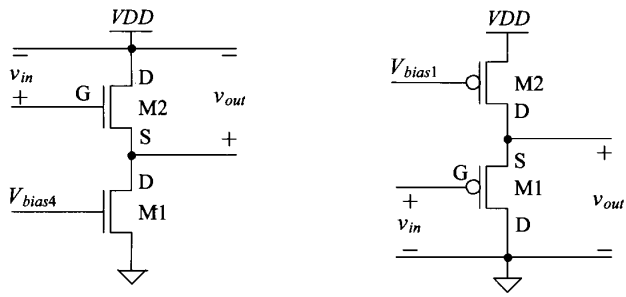


Figure 21.39 Source followers (common drain) using current source loads.

IPR2023-00698
Theta EX2019

Knowing the resistance looking into the drain of M1 is r_o and the current that flows in M1 is i_d, we can write (neglecting M2's output resistance)

$$v_{out} = i_d \cdot r_o \text{ and } v_{gs2} = \frac{i_d}{g_{m2}} \tag{21.88}$$

Solving for the gain, we get

$$A_v = \frac{v_{out}}{v_{in}} = \frac{r_o}{r_o + 1/g_{m2}} = \frac{g_{m2} r_o}{g_{m2} r_o + 1} \tag{21.89}$$

If we use a current source for the load with a very large small-signal output resistance, the gain goes to one. (Limitations concerning the body effect will be discussed in a moment.) Looking at Eq. (21.89), we see that the gain is a voltage divider between the resistance looking into the source of a MOSFET ($1/g_m$) and the output resistance of the current source (here using the simple, single MOSFET r_o) as seen in Eq. (21.32).

Note that the maximum current the NMOS SF can sink is limited by the size of the current flowing in the current source load M1. The SF is a class A amplifier where, when discharging a load capacitance, the current through M1 limits the rate at which the capacitance can discharge (as indicated in Eq. [21.38]). M2 only sources current and M1 only sinks current. When the SF is sourcing a current, M2 provides the current to both the amplifier's output and to M1. When the SF is sinking a current, the current in M2 decreases while the current in M1 is constant (resulting in a net sinking of current on the amplifier's output). Finally, note that the maximum output voltage of an NMOS SF occurs when the gate of M2 goes to *VDD*. The maximum output voltage then goes to $VDD - V_{GS2}$. The minimum output voltage is set by the minimum voltage across the current source load (to keep the MOSFETs in the saturation region).

Body Effect and Gain

We might think, after looking at Eq. (21.89), that by using a cascode current source load for the source follower we can get an AC small-signal gain very close to unity. However, the body effect ultimately limits the gain of the amplifier. Figure 21.40 shows an SF with an ideal current source load. If we use the AC small-signal model seen in the figure where the output resistance, r_o, is infinite, we see that the AC output voltage, v_{out}, equals the AC source to bulk potential v_{sb}. Further, we can write

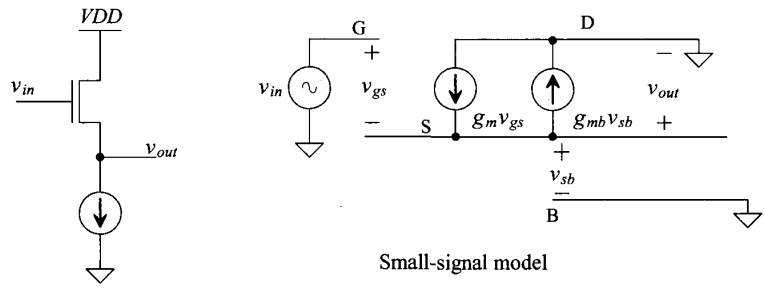


Figure 21.40 SF gain with body effect

IPR2023-00698
Theta EX2019

$$v_{gs} = v_{in} - v_{out} \tag{21.90}$$

and, summing the currents at the output node,

$$g_m v_{gs} = g_{mb} v_{sb} = g_{mb} v_{out} \tag{21.91}$$

Solving for the gain of the SF with body effect and knowing $g_{mb} = \eta \cdot g_m$ (Eq. (9.28))

$$\frac{v_{out}}{v_{in}} = \frac{g_m}{g_m + g_{mb}} = \frac{1}{1+\eta} \tag{21.92}$$

If $\eta = 0.25$, the gain is 0.8. The obvious way of eliminating the body effect (and thus move the gain closer to one) is to put the common-drain device in its own well, Fig. 21.41.

See Table 9.2 and Fig. 20.47

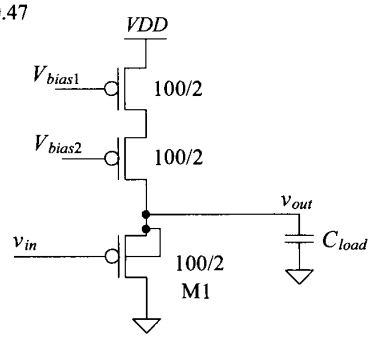


Figure 21.41 PMOS SF without body effect.

Level Shifting

One of the important uses of an SF is to provide a DC level shift to an input voltage. For example, M1 in Fig. 21.41 will remain in saturation as long as

$$v_{SD} = v_{OUT} \geq v_{SG} - V_{THP} = \overbrace{v_{OUT} - v_{IN}}^{v_{SG}} - V_{THP} \tag{21.93}$$

or

$$v_{IN} \geq -V_{THP} \tag{21.94}$$

As long as the input voltage is greater than $-V_{THP}$, M1 operates in the saturation region. This result is practically important. Note the output and input of the SF in Fig. 21.41 are related by

$$v_{OUT} = v_{IN} + V_{SG} \tag{21.95}$$

Using the parameters in Table 9.2 where *VDD* is 1 *V*, V_{SG} = 350 mV, and V_{THP} = 280 mV, the input voltage is shifted upwards, on the output of the amplifier, by 350 mV. The input signal can go down to –280 mV before M1 triodes. The minimum voltage across the current source, as seen in Fig. 20.48, is approximately 150 mV. This means that the range of the input signal is from *VDD* – 150 mV – 350 mV (= 500 mV = $V_{in,max}$) down to – 280 mV. Simulation results are seen in Fig. 21.42.

IPR2023-00698
Theta EX2019

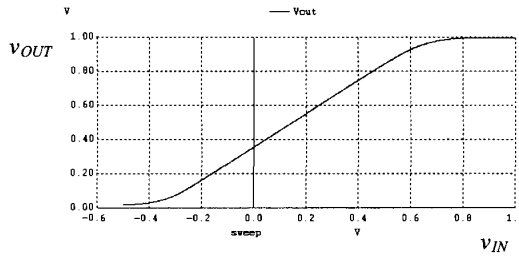


Figure 21.42 How the SF in Fig. 21.41 can shift negative input voltages upwards. This circuit is very useful when input signals are centered around ground.

Input Capacitance

Consider the partial schematic of an SF seen in Fig. 21.43. The capacitance on the input can be determined by looking at how much charge is supplied by the input source for an input voltage change Δv_{IN}

$$Q_{IN} = \Delta v_{IN} \cdot C_{dg} + (\Delta v_{IN} - A_v \cdot \Delta v_{IN}) \cdot C_{sg} \tag{21.96}$$

where $A_v \leq 1$. The input capacitance of an SF is then

$$C_{IN} = \frac{Q_{IN}}{\Delta v_{IN}} = C_{dg} + C_{sg}(1 - A_v) \approx C_{dg} \tag{21.97}$$

As the gain, A_v , of the SF approaches one, the input capacitance approaches the drain-gate capacitance of the MOSFET. In other words the source-gate capacitance doesn't affect the input capacitance (unlike the CS amplifier, which can have a very large input capacitance due to the Miller effect; see Ex. 21.6). Intuitively we can understand why C_{sg} doesn't affect the input capacitance by realizing that the displacement current through it is zero when the gate and source potentials move at the same rate. When the input voltage change and the output voltage changes are equal, the current through C_{sg} is zero. The SF is often used on the input of amplifiers that must have low input capacitance (like in charge-amplification applications). Unfortunately, the noise performance of the SF, because it doesn't have gain, is poorer than the CS or CG amplifiers.

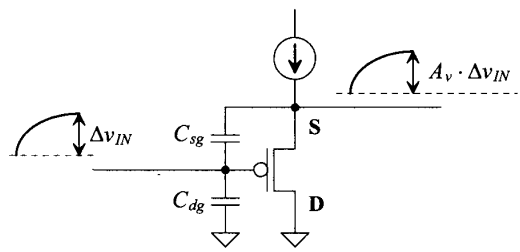


Figure 21.43 Input capacitance of a SF.

IPR2023-00698
Theta EX2019

Noise Performance of the SF Amplifier

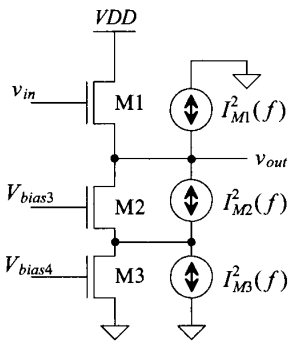
The SF amplifier with MOSFET noise sources is seen in Fig. 21.44. The noise current from M3 flows through M2 to the output node. The resistance on the output node (the output resistance of the SF) is the resistance looking into the source of M1 in parallel with the resistance looking into the drain of M2 or

$$R_{oSF} = \frac{1}{g_{m1}} || R_o \approx \frac{1}{g_{m1}} \tag{21.98}$$

The output noise power spectral density is

$$V^2_{onoise}(f) = \frac{I^2_{M1}(f) + I^2_{M3}(f)}{g^2_{m1}} \tag{21.99}$$

Noting that the SF's gain is close to one, the input-referred noise is nearly equal to the output noise PSD. Again, by using a large value of g_{m1}, we can minimize both the input-referred and output noise.



Neglect noise current from M2 since, $I_{M2}(f) + i_{D2} = 0$. The noise current that leaves the source must equal the noise current that enters the drain, that is, none of M2's noise current flows into the source of M1 and thus varies the output voltage.

Figure 21.44 Noise model of the source follower amplifier.

Frequency Behavior

Using Eq. (21.97), we see that if we drove the SF with an input whose source resistance was R_s, then we would have a pole associated with the input at

$$f_{in} = \frac{1}{2\pi \cdot R_s C_{dg}} \tag{21.100}$$

a very high frequency. On the output of the SF driving a capacitive load, we can write, using Eq. (21.98),

$$f_{out} = \frac{1}{2\pi \cdot R_{oSF} C_{load}} = \frac{g_{m1}}{2\pi C_{load}} \tag{21.101}$$

However, both Eqs. (21.100) and (21.101) assume that the transconductance doesn't vary with frequency. At very high frequencies, the input and output of the SF are shorted together through the gate-source capacitance. To calculate the variation of the transconductance with frequency, consider the circuit seen in Fig. 21.45. We can write

IPR2023-00698
Theta EX2019

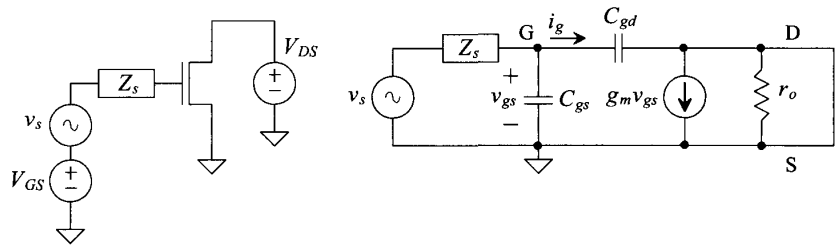


Figure 21.45 Determining the variation of the transconductance with frequency.

$$v_{gs}(f) = v_s \cdot \frac{\frac{1}{j\omega\cdot(C_{gs}+C_{gd})}}{\frac{1}{j\omega\cdot(C_{gs}+C_{gd})}+Z_s} = \frac{v_s}{1+Z_s\cdot j\omega\cdot(C_{gs}+C_{gd})} \tag{21.102}$$

and

$$i_d(f) = g_m \cdot v_{gs}(f) = \frac{g_m v_s}{1+Z_s\cdot j\omega\cdot(C_{gs}+C_{gd})} \tag{21.103}$$

The effective transconductance as a function of frequency is then

$$g_m(f) = \frac{g_m}{1+j\omega\cdot Z_s(C_{gs}+C_{gd})} \tag{21.104}$$

The impedance looking into the source of a MOSFET as a function of frequency is then

$$R_{\text{into source}} = \frac{1}{g_m(f)} = \frac{1}{g_m}\cdot(1+j\omega\cdot Z_s(C_{gs}+C_{gd})) \tag{21.105}$$

We note that at low frequencies the impedance is resistive and has a value of $1/g_m$. However, at higher frequencies and with a resistive source impedance, that is, $Z_s = R_s$, the impedance looking into the source appears to be the series connection of a resistor with a value of $1/g_m$ and an inductor of value

$$L_s = \frac{R_s(C_{gs}+C_{gd})}{g_m} \tag{21.106}$$

An SF driving a capacitive load can exhibit ringing because of the effective RLC circuit formed by its output impedance driving a load capacitance. If the impedance driving an SF is inductive (as we would have in the cascade of two SFs), the output impedance can become negative. For example, if $Z_s = j\omega L$, then substituting into Eq. (21.105), we get

$$R_{\text{into source}} = \frac{1}{g_m} - \omega^2\cdot L\cdot(C_{gs}+C_{gd}) \tag{21.107}$$

A battery or voltage source is an example of a circuit with a negative resistance (any source of power has a negative resistance). At higher frequencies the resistance looking into the source of the MOSFET becomes negative. This means that the circuit will oscillate or simply have a poor step response (the voltage across the capacitive load will ring). Often, to implement a microwave frequency oscillator using a MOSFET, an inductor is added from the gate of the MOSFET to AC ground to create the negative resistance.

IPR2023-00698
Theta EX2019

SF as an Output Buffer

One of the common uses of an SF is as an output buffer. Connecting a resistive load to a high-impedance node kills the gain of a single-stage CMOS amplifier. A buffer amplifier is often inserted between the high-impedance node and the load resistance to keep the gain high. Consider the following example.

Example 21.13

Suppose that the cascode amplifier in Fig. 21.35, with R_s = 100k, must drive a 1 pF capacitor in parallel with a 10 kΩ resistor. Design an SF buffer to ensure that the 10 kΩ resistor doesn't kill the gain of the amplifier. Estimate the frequency response of the amplifier. Verify your design with SPICE simulations.

Looking at Eq. (21.76) and the associated discussion, we see that the output resistance of the cascode amplifier is in the megaohms region. Connecting a 10 kΩ load resistor on the amplifier's output would cause the cascode gain to drop from 500 to less than 1. The DC output voltage of the cascode amplifier in Fig. 21.35 is the V_{GS} of M1 (which, from Table 9.2 is 350 mV). If we use an NMOS SF, this DC voltage isn't enough to ensure that all of the MOSFETs in the SF operate in the saturation region. Looking at the NMOS SF in Fig. 21.39, a voltage of 350 mV on the gate of M2 would be just enough to turn it on but leave little voltage to drop across the current source M1. We'll use a PMOS source follower in this design. The schematic of the design is seen in Fig. 21.46. The current source (M5 and M6) in the SF must drive 1 V across the resistor of 10 kΩ (100 μA), so we've bumped up their size. Knowing, from Table 9.2, that when using the bias circuit in Fig. 20.47 a 100/2 PMOS conducts 10 μA of current, we increased the widths of M5–M7 so that they conduct 125 μA of current with the same bias voltages (the current source, M5/M6, sources 125 μA of current).

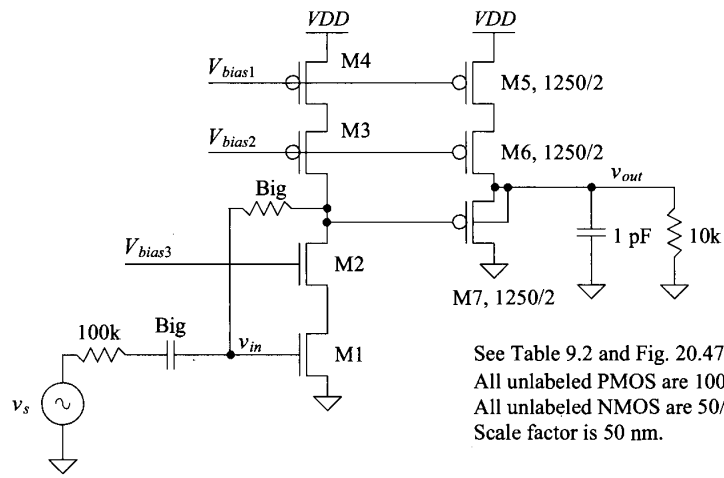


Figure 21.46 A cascode amplifier with SF output buffer.

IPR2023-00698
Theta EX2019

The location of the input pole, f_{in} , is still (from Ex. 21.12) located at 218 MHz. The SF won't have much affect on the overall frequency response (because of its low input capacitance and small output resistance). However, now the cascode amplifier's load capacitance changes. Using the result in Eq. (21.97) and the data from Table 9.2, we can estimate the capacitance on the output of the cascode amplifier as the gate-drain capacitance of M7 ($3.7\,fF \cdot 12.5 \approx 50\,fF$). The factor of 12.5 comes from the increase in the PMOS's width by 12.5. Again, using the results from Ex. 21.12, the output pole, f_{out} , (output of the cascode amplifier) is at approximately 1 MHz (double what it was in Ex. 21.12).

The gain of the cascode amplifier is approximately 500. The transconductance of the wider PMOS devices is estimated using (see Eq. [9.22])

$$g_{m,wide} = K \cdot g_{m,Table\ 9.2} = 1.875\ mA/V \text{ where } K = 12.5 \quad (21.108)$$

The gain of the SF driving a resistive load, see Eq. (21.89), is

$$A_{v,SF} = \frac{R_{load}}{R_{load} + 1/g_{m,wide}} = \frac{10k}{10k + 533} = 0.95 \quad (21.109)$$

The overall gain from the input of the cascode amplifier to the output of the SF is then estimated as $500 \cdot 0.95 = 475 \rightarrow 53.5\ dB$. The simulation results are seen in Fig. 21.47. Note that the bandwidth of this amplifier can be increased by reducing the width of M7. This reduces the input capacitance of the SF output amplifier and thus pushes the pole location out to a higher frequency. ■

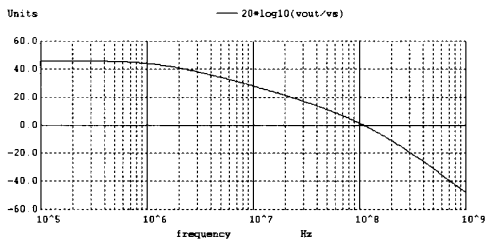


Figure 21.47 Simulating the operation of the amplifier in Fig. 21.46.

A Class AB Output Buffer Using SFs

In order to drive the 10k resistor in the previous example, we had to increase the size of the current source used in the SF. The practical problem with this approach is that as the load resistor gets small, the current source must source a significant current. The current source is sourcing this current all of the time (class A operation) either to the load or, if the load capacitance is being discharged, to the common-drain MOSFET (M7 in Fig. 21.46). What we need, for better drive capability, is for one side of the output buffer to shut off if the other side is supplying a large amount of current. For example, if, in Fig. 21.46, M7 starts to pull a large amount of current from the load, we would want M5 and M6 to turn off so that power wasn't wasted (class AB operation).

IPR2023-00698
Theta EX2019

A class AB output buffer using SFs is seen in Fig. 21.48. The output buffer is comprised of M1–M4. M5 and M6 form a common-source amplifier with a current source load. When the circuit is in steady state, the AC input, v_{in}, is zero and the gate of M6 is at a DC potential of V_{bias1}. The current through M3 and M4 is mirrored by M1 and M2. This sets the DC current in M1/M2 to a known value (important). As v_{in} goes up, M6 shuts off. However, the current in M5 is constant, so the gates of M3 and M4 move towards ground. This turns M1 on and shuts M2 off (thus the class AB action). Similarly if v_{in} decreases, M6 turns on and pulls the gates of M3 and M4 up. The result is that M2 turns on and M1 shuts off.

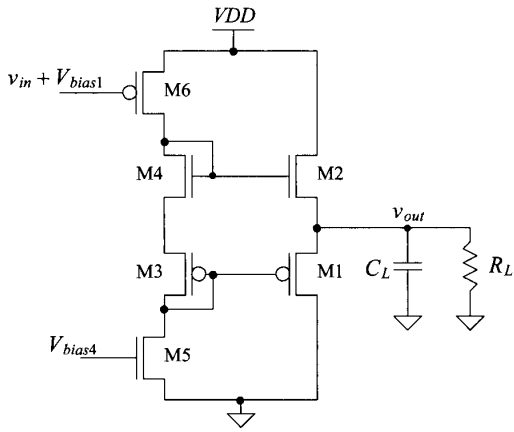


Figure 21.48 Class AB output buffer (M1–M4) using source followers M1 and M2.

The *practical problem* with this buffer is that the output can't swing very close to the power supply voltages (rails). The lowest potential we can get on the gate of M1 is ground, and the highest potential we can get on the gate of M2 is *VDD*. Knowing that M1 has to have a source-to-gate voltage greater than V_{THP} and M2 has to have a gate-source voltage greater than V_{THN}, we can write

$$VDD - V_{THN} \geq v_{out} \geq V_{THP} \tag{21.110}$$

For our short-channel CMOS process (neglecting body effect which will make things worse), the range of output voltages is half the power supply voltage. For example, with *VDD* = 1 V, our output voltage may swing up to (roughly) 750 mV and down to 250 mV. Losing half of the supply voltage often makes this output buffer impractical.

21.3 The Push-Pull Amplifier

What is needed for an output buffer is a topology like an inverter where the output can swing from rail (*VDD*) to rail (ground). Figure 21.49 shows a possible topology with biasing circuitry. The current source, I_{bias}, is a floating current source as seen in Fig. 20.49 of the last chapter. In Fig. 21.49, with zero AC input current, i_{in}, M1 and M2 mirror

IPR2023-00698
Theta EX2019

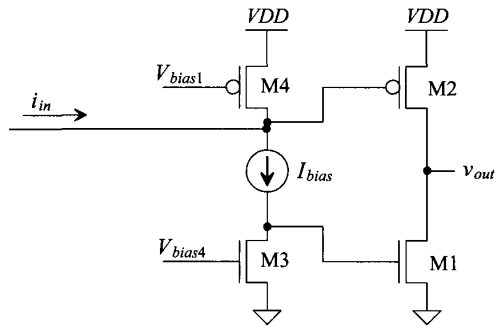


Figure 21.49 Class AB amplifier using an inverter output structure (push-pull).

the current in M3 and M4 (I_{bias}). This precisely sets the current in the output stage (again, important). A positive AC input current causes both the gates of M1 and M2 to go up (shutting M2 off and turning M1 on). M1 and M2 are pushing or pulling a current to/from the output (and so this topology is often called a *push-pull amplifier*). Because the output can swing very close to ground and *VDD* before M1 and M2 triode or shut off, this topology is very useful in modern CMOS output buffer design.

Note that the AC input current can be connected to the gate of M2, as shown in Fig. 21.49, or the gate of M1 (or both).

21.3.1 DC Operation and Biasing

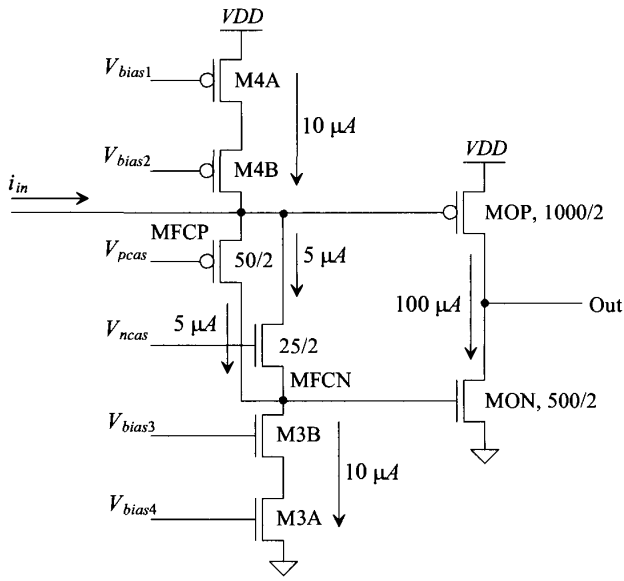
The biasing of the push-pull amplifier can be accomplished, as seen in Fig. 21.50 (see also Fig. 20.49). MFCP and MFCN form the floating current source. M3 (A and B) with M4 (A and B) are cascode current sources. If the input current is positive, the gate of MOP is charged up and thus MOP shuts off. At the same time, MFCP turns on (more), causing the gate of MON to go up, turning it on further. If the input current is negative, MOP turns on and MON shuts off.

Figure 21.51 shows the amplifier in Fig. 21.50 with a voltage input (V_{bias4} replaced with an input voltage, V_{in}). M3 and M4 now form a common-source amplifier with a current source load. Shown in Fig. 21.51 is a plot of V_{out} for varying V_{in} with no load. The slope of the curve in this figure is the gain. To show the slope (the gain) as a function of V_{in}, we can take the derivative of the output voltage (also shown in the figure). The gain of this topology (without a load) is roughly 5,000 V/V. If we place a 1k resistor on the output of the amplifier, the gain drops to 1,500 V/V. Using a 100-ohm resistor results in a gain of only 150. The push-pull amplifier, MON/MOP, has a gain less than one when the load resistor is only 100 ohms. If the amplifier did need to drive such a heavy load (small resistance), it would be a good idea to increase the widths of MON and MOP.

Power Conversion Efficiency

The power conversion efficiency (PCE) is defined as the ratio of the power supplied to the load to the power delivered to the amplifier and load by a power supply. In other words, a perfectly efficient amplifier doesn't dissipate any power; rather, all of the power

IPR2023-00698
Theta EX2019



Bias voltages come from Fig. 20.47 (short-channel parameters in Table 9.2). Unlabeled NMOS are 50/2, while unlabeled PMOS are 100/2.

Figure 21.50 Biasing the push-pull amplifier.

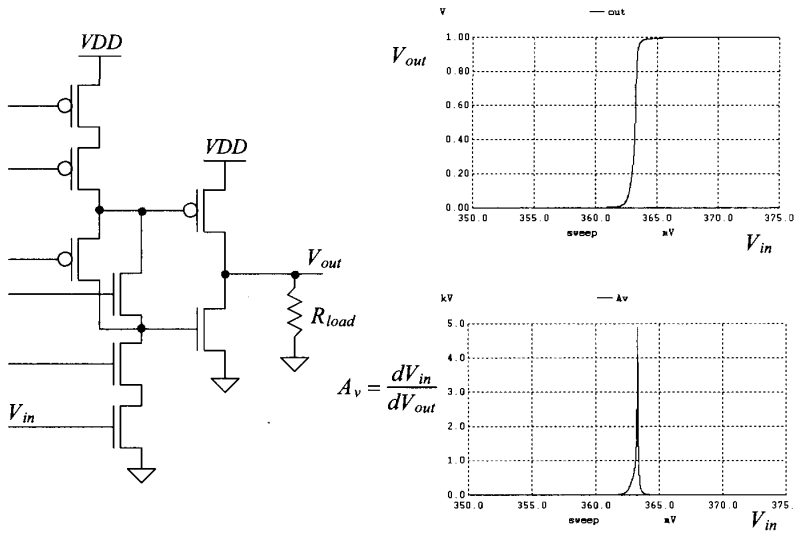


Figure 21.51 Simulating the amplifier in Fig. 21.50 with a voltage input.

IPR2023-00698
Theta EX2019

from the power supply (*VDD*) is supplied to the load. The PCE is usually written as a percentage

$$\%\ \text{PCE} = \frac{\text{Load power, } P_{load}}{\text{Supply power, } P_{supply}} \times 100\ \% \tag{21.111}$$

For example, the source follower seen in Fig. 21.41 (a class A amplifier) pulls a fixed current of I_{bias} (= 10 μA for the sizes seen in this figure and Table 9.2). The power supplied (pulled or delivered) from the power supply is

$$P_{supply} = VDD \cdot I_{bias} \tag{21.112}$$

If the SF is driving a resistive load, R_{load}, the peak value (under ideal conditions) of a sinewave voltage applied to this resistor is *VDD*/2 (the sinewave can go up or down by *VDD*/2). The RMS value of this sinewave is then $VDD/\left(2\sqrt{2}\right)$ and the power supplied to the load is then

$$P_{load} = \frac{\left(VDD/\left(2\sqrt{2}\right)\right)^2}{R_{load}} \tag{21.113}$$

The current source, I_{bias}, when the output voltage goes to *VDD,* drives the load resistor directly (M1 in Fig. 21.41 shuts off), that is

$$VDD = I_{bias} \cdot R_{load} \tag{21.114}$$

This is the best efficiency because no power is wasted in the amplifier (M1 is off and all of the supply power is delivered to the load). Calculating the PCE of the SF gives

$$\%\ \text{PCE} = \frac{1}{VDD \cdot I_{bias}} \cdot \frac{\left(VDD/\left(2\sqrt{2}\right)\right)^2}{VDD/I_{bias}} \times 100\ \% \tag{21.115}$$

So the PCE % is, at best, 12.5%. This is the ideal efficiency of a class A amplifier. If the output voltage swing is reduced, the PCE goes down as well.

For the class AB amplifier, the PCE can approach 100%. (Typical values for general designs with low distortion are approximately 75%.) If the current supplied to the load is much larger than the current burned in the amplifier, the PCE is large (much larger than the ideal 12.5% of the class A amplifier). Consider the following example.

Example 21.14

Suppose the amplifier in Figs. 21.50 and 21.51 drives a 1 kΩ load. Using SPICE plot the current supplied to the amplifier and load, the current supplied to only the load, and the current supplied to only the amplifier as a function of the input voltage, V_{in}. Using the results estimate the PCE.

The simulation results are seen in Fig. 21.52. The current supplied to the amplifier includes the bias circuit current. Note how, when the output is 1 V (= *VDD*), the current supplied to the load is 1 mA. The total current supplied by *VDD* is approximately 1.2 mA. At the risk of stating the obvious, the push-pull amplifier is much more power-efficient than the SF (and has a wider output swing).

IPR2023-00698
Theta EX2019

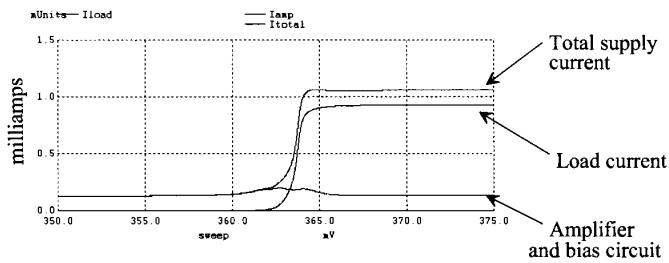


Figure 21.52 Supply, load, and total currents when the amplifier in Fig. 21.51 drives a 1k load resistor.

The PCE efficiency can be estimated by writing the power dissipated by the load

$$P_{load} = \left(\frac{VDD}{2\sqrt{2}}\right)^2 \cdot \frac{1}{R_{load}} = \left(\frac{1}{2\sqrt{2}}\right)^2 \cdot \frac{1}{1k} = 125\ \mu W$$

The power supplied by *VDD* (assuming the amplifier and bias circuit pull a constant 200 μA of current) is then

$$P_{supply} = P_{load} + VDD \cdot 200\ \mu A = 325\ \mu W$$

The PCE is then

$$\%\ \text{PCE} = \frac{125}{325} \times 100\% = 38\%$$

Note that the bias circuit pulls 140 μA. If we recalculate the PCE without the bias circuit current included, we get

$$\%\ \text{PCE} = \frac{125}{185} \times 100\% = 68\% \quad \blacksquare$$

21.3.2 Small-Signal Analysis

The simplified schematic of the push-pull amplifier is seen in Fig. 21.53. The resistance on the output of the amplifier (the drains of MOP and MON) is $r_{op}||r_{on}||R_{load} \approx R_{load}$. The drain current of MON is $g_{mon}v_{in}$ and the drain current of MOP is $-g_{mop}v_{in}$. The output voltage is then

$$v_{out} = -(g_{mon} + g_{mop}) \cdot v_{in} \cdot R_{load} \tag{21.116}$$

Note the "resistance in the source" in this amplifier is the parallel combination of the resistance looking into the sources of MON and MOP (both sources are connected to AC ground). We can rewrite Eq. (21.116) as

$$A_{v,push-pull} = \frac{v_{out}}{v_{in}} = -\frac{R_{load}}{\frac{1}{g_{mon}}||\frac{1}{g_{mop}}} \tag{21.117}$$

If the load resistance is 1k and the sum of the transconductances (using wider devices than what is indicated in Tables 9.1 or 9.2, as seen in Fig. 21.50) is 1 mA/V, then the gain of the push-pull amplifier is −1 V/V. Note that if the load resistance changes, the gain of the output stage push-pull amplifier changes as well.

IPR2023-00698
Theta EX2019

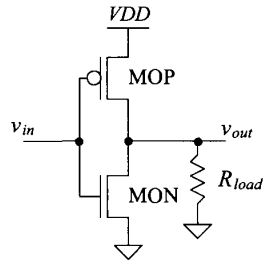


Figure 21.53 Small-signal analysis of the push-pull amplifier.

The gain of the CS amplifier with the cascode load portion of the amplifier in Figs. 21.50 and 21.51 (M3 and M4) is given by Eq. (21.76). The overall small-signal gain of the topology in Fig. 21.51 is given by the product of the two gains or

$$A_v = g_{mn} \cdot \left(g_{mn} r_{on}^2 || g_{mp} r_{op}^2\right) \cdot (g_{mon} + g_{mop}) \cdot R_{load} \qquad (21.118)$$

We might wonder how the floating current source, MFCP and MFCN, affects the gain. If we look at Fig. 21.50, we see that one or the other has a source connection to the drains of either M3 (gate of MON) or M4 (gate of MOP). Further we might think that this source connection would load the gates with a $1/g_m$ small-signal resistance. However, because MFCP and MFCN form a feedback loop, their addition doesn't reduce the output resistance of the CS amplifier (M3 and M4). Consider the following.

Figure 21.54 shows a test circuit used for determining the small-signal resistance that the PMOS current source (M4) sees on its output. To determine this resistance, we apply a test voltage and look at the current that flows. Looking at the figure and noting the current that flows in $r_{on} || r_{op}$ is $(i_t - i_{dn} - i_{dp})$, we can write

$$v_t = (i_t - i_{dn} - i_{dp}) \cdot r_{on} || r_{op} + i_t \cdot R_{ncas} \qquad (21.119)$$

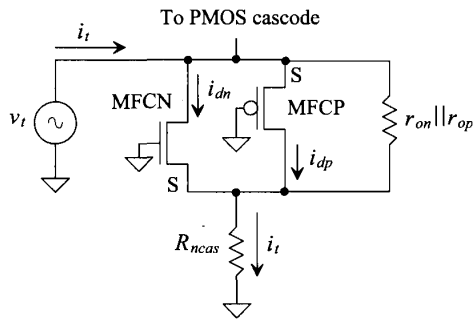


Figure 21.54 Determining the loading affects of the floating current source.

IPR2023-00698
Theta EX2019

where R_{ncas} is the output resistance of the cascode stack. The gate-source AC voltage of MFCN is $-i_t R_{ncas}$ and so

$$i_{dn} = g_{mn} \cdot (-i_t R_{ncas}) \tag{21.120}$$

Similarly, the drain current through MFCP is

$$i_{dp} = g_{mp} \cdot v_t \tag{21.121}$$

Substituting these equations into Eq. (21.119) gives

$$v_t = (i_t + g_{mn} \cdot i_t R_{ncas} - g_{mp} \cdot v_t) \cdot r_{on} || r_{op} + i_t \cdot R_{ncas} \tag{21.122}$$

or

$$\overbrace{v_t \cdot (1 + g_{mp} \cdot r_{on} || r_{op})}^{\approx g_{mp} \cdot r_{on} || r_{op}} = i_t \cdot \overbrace{(1 + g_{mn} R_{ncas} \cdot r_{on} || r_{op} + R_{ncas})}^{\approx g_{mn} \cdot r_{on} || r_{op} \cdot R_{ncas}} \tag{21.123}$$

The resistance the PMOS cascode sees is then

$$\frac{v_t}{i_t} \approx R_{ncas} \tag{21.124}$$

In other words, the floating current source doesn't load the cascode structure.

21.3.3 Distortion

Small-signal analysis works remarkably well for all of the internal nodes in an op-amp. However, the output buffer must, ideally, swing from rail-to-rail. To illustrate the problem with this, consider the basic gain of a CS amplifier with current source load, Eq. (21.39),

$$|A_v| = g_{m1} \cdot (r_{o1} || r_{o2}) = \frac{\sqrt{2\beta_1 (I_D + i_d)}}{2(\lambda_n + \lambda_p)(I_D + i_d)} = \frac{\sqrt{2\beta_1}}{2(\lambda_n + \lambda_p)\sqrt{(I_D + i_d)}} \tag{21.125}$$

Normally, the AC component of the drain current, i_d, is assumed to be much less than the DC component of the drain current, I_D, and the amplifier gain is essentially constant (small-signal approximation). If the AC component is comparable to the DC component, noticeable distortion results. The voltage gain for large inputs depends on the input signal amplitude.

Characterizing an amplifier begins by applying a pure single-tone sinusoid of the form

$$V_{in}(t) = V_p \sin 2\pi f \cdot t \tag{21.126}$$

to the input of the amplifier. The output of the amplifier is a series of tones at an integer multiple of the input tone given by

$$V_{out}(t) = a_1 V_p \sin(\omega t) + a_2 V_p \sin(2\omega t) + \cdots + a_n V_p \sin(n\omega t) \tag{21.127}$$

The magnitude of the fundamental or wanted signal is $a_1 V_p$. Ideally, a_2 through a_n are zero, and the amplifier is free of distortion. The nth term harmonic distortion is given by

$$HD_n = \frac{a_n}{a_1}, \text{ for } n > 1 \tag{21.128}$$

The *total harmonic distortion* (THD) is given by

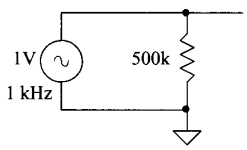
IPR2023-00698
Theta EX2019

$$THD = \sqrt{\frac{a_2^2 + a_3^2 + a_4^2 + \cdots + a_n^2}{a_1^2}} \qquad (21.129)$$

Again, output buffers, amplifiers used to drive a large load capacitance or low resistance, are examples of amplifiers where low THD is important. If the output amplifier is biased so that the DC component of the drain current is large compared to the transient or time-varying current, the buffer will dissipate too much power for most applications. Therefore, in almost all situations, the biasing current in an output buffer is comparable, or even smaller for class B operation, than the time-varying current. We reduce the distortion by employing feedback around the amplifier. If the open-loop gain of an op-amp varies from 1,000 to 10,000 depending on the amplitude of the input signal, then the feedback around the amplifier reduces the gain sensitivity. In fact, it is nearly impossible to design a linear output amplifier with low distortion without using feedback.

Modeling Distortion with SPICE

SPICE can be used to simulate distortion using a transient analysis and the .FOUR (Fourier) statement. The general form of this statement is .FOUR FREQ OV1 <OV2 OVI . . . where FREQ is the frequency of the fundamental and OV1 . . . are the outputs of the circuit (the voltage or current outputs for which SPICE will calculate distortion). As a simple example, consider the circuit and netlist shown in Fig. 21.55. The input sine wave must have at least one full period for the .FOUR statement to calculate distortion. In the case where there's more than one period, SPICE uses the output over the last full period. Also, the maximum transient step size should be less than the period of the input divided by 100. For the example of Fig. 21.55 with a 1 kHz input (a period of 1 ms), the maximum print size should be 10 μs.




```
*** SPICE Fourier Analysis Example ***
.four 1k Vinout
R1 Vinout 0 500k
Vinout Vinout 0 DC 0 AC 0 0 SIN(0 1 1khz 0 0)
.tran 10u 2m 0 10u
.end
```

Figure 21.55 Simple circuit to demonstrate the use of the .FOUR statement.

The resulting simulation output follows.

```
Fourier analysis for vinout:
 No. Harmonics: 10, THD: 3.17012e-06 %, Gridsize: 200, Interpolation Degree: 1
```

Harmonic	Frequency	Magnitude	Phase	Norm. Mag	Norm. Phase
0	0.000000e+00	-5.60197e-09	0.000000e+00	0.000000e+00	0.000000e+00
1	1.000000e+03	9.996317e-01	-9.39744e-06	1.000000e+00	0.000000e+00
2	2.000000e+03	1.120393e-08	-8.64000e+01	1.120806e-08	-8.64000e+01
3	3.000000e+03	1.120393e-08	-8.46000e+01	1.120806e-08	-8.46000e+01
4	4.000000e+03	1.120394e-08	-8.28000e+01	1.120806e-08	-8.28000e+01
5	5.000000e+03	1.120394e-08	-8.10000e+01	1.120806e-08	-8.10000e+01
6	6.000000e+03	1.120394e-08	-7.92000e+01	1.120806e-08	-7.92000e+01

IPR2023-00698
Theta EX2019

```
7    7.000000e+03   1.120393e-08  -7.74000e+01   1.120806e-08  -7.74000e+01
8    8.000000e+03   1.120393e-08  -7.56000e+01   1.120806e-08  -7.56000e+01
9    9.000000e+03   1.120393e-08  -7.38000e+01   1.120806e-08  -7.38000e+01
```

Notice that, as we would expect, the output of this circuit doesn't show any harmonic distortion. SPICE calculates the magnitude and phase of the first nine harmonics and DC. Also note that SPICE automatically calculates the THD for a circuit.

As a more practical example, consider the push-pull amplifier shown in Fig. 21.56. The SPICE netlist for simulating the distortion performance of this output amplifier is shown below. Note that to avoid a long start-up transient we started the big capacitors at the DC bias voltages (V_{bias1} for the PMOS and V_{bias4} for the NMOS). When the input sinewave has an amplitude of 2 mV, the THD is 0.6%. When the input drops to 100 μV (small-signal approximation is more valid), the THD drops to 0.046%. If the input amplitude increases to 10 mV (so the output of the push-pull amplifier swings close to the rails), the THD increases to 9.2%.

```
*** Figure 21.56 CMOS: Circuit Design, Layout, and Simulation ***

.option scale=50n
.tran 10n 2u UIC
.four 1MEG Vout
VDD   VDD    0      DC    1
Vin   Vin    0      DC    0      sin 0 2m 1MEG
RL    out    0      1k
Rbigp Vbias1 vgp    1G
Cbigp vgp    vin    1 IC=0.643
Rbign Vout   vgn    1G
Cbign Vgn    vin    1 IC=0.362
Xbias VDD Vbias1 Vbias2 Vbias3 Vbias4 Vhigh Vlow Vncas Vpcas bias
MON   Vout   vgn    0      0      NMOS L=2 W=500
MOP   Vout   vgp    VDD    VDD    PMOS L=2 W=1000
(subcircuit and SPICE models not shown)
```

Note that another useful SPICE tool to evaluate the distortion introduced into a signal by a circuit is a Discrete Fourier Transform (DFT). The **spec** command is used in SPICE (spectral analysis) to determine the output spectrum of an amplifier. Using this command is covered in depth in the book entitled, *CMOS Mixed Signal Circuit Design.*

Bias circuit from Fig. 20.47 (Table 9.2).

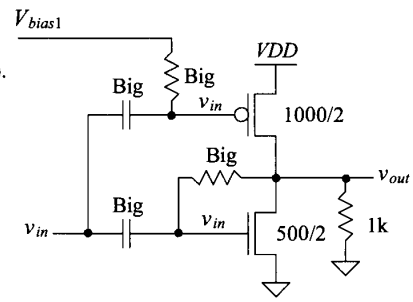


Figure 21.56 Simulating the distortion in a push-pull amplifier.

IPR2023-00698
Theta EX2019

ADDITIONAL READING

[1] K. N. Leung, P. K. T. Mok, W. Ki, and J. K. O. Sin, "Three-stage large capacitive load amplifier with damping-factor-control frequency compensation," *IEEE Journal of Solid-State Circuits,* vol. 35, pp. 221–230, February 2000.

[2] A. B. Dowlatabadi, "A robust, load-insensitive pad driver," *IEEE Journal of Solid-State Circuits,* vol. 35, pp. 660–665, April 2000.

[3] B. Razavi, "CMOS technology characterization for analog and RF design," *IEEE Journal of Solid-State Circuits,* vol. 34, pp. 268–276, March 1999.

[4] K. de Langen and J. H. Huijsing, "Compact low-voltage power-efficient operational amplifier cells for VLSI," *IEEE Journal of Solid-State Circuits,* vol. 33, pp. 1482–1496, October 1998.

[5] P. E. Allen, B. J. Blalock, and G. A. Rincon, "A 1V CMOS Opamp using Bulk-Driven MOSFETs," *IEEE International Solid-State Circuits Conference,* vol. 38, pp. 192–193, February 1995.

[6] R. G. Eschauzier, R. Hogervorst, and J. H. Huijsing, "A programmable 1.5 V CMOS class-AB operational amplifier with hybrid nested Miller compensation for 120 dB gain and 6 MHz UGF," *IEEE Journal of Solid-State Circuits,* vol. 29, pp. 1497–1504, December 1994.

[7] R. Hogervorst, J. P. Tero, R. G. H. Eschauzier, and J. H. Huijsing, "A Compact Power-Efficient 3 V CMOS Rail-to-Rail Input/Output Operational Amplifier for VLSI Cell Libraries," *IEEE Journal of Solid State Circuits,* vol. 29, pp. 1505–1513, December 1994.

[8] A. A. Abidi, "On the operation of cascode gain stages," *IEEE Journal of Solid-State Circuits,* vol. 23, pp. 1434–1437, December 1988.

[9] S. L. Wong and C. A. T. Salama, "An efficient CMOS buffer for driving large capacitive loads," *IEEE Journal of Solid-State Circuits,* vol. 21, pp. 464–469, June 1986.

[10] D. B. Ribner and M. A. Copeland, "Design techniques for cascoded CMOS op amps with improved PSRR and common-mode input range," *IEEE Journal of Solid-State Circuits,* vol. 19, pp. 919–925, December 1984.

[11] V. R. Saari, "Low-power high-drive CMOS operational amplifiers," *IEEE Journal of Solid-State Circuits,* vol. 18, pp. 121–127, February 1983.

[12] B. K. Ahuja, "An Improved Frequency Compensation Technique for CMOS Operational Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 18, pp. 629–633, December 1983.

[13] P. R. Gray and R. G. Meyer, "MOS operational amplifier design - A tutorial overview," *IEEE Journal of Solid-State Circuits,* vol. 17, pp. 969–982, December 1982.

[14] B. J. Hosticka, "Dynamic CMOS amplifiers," *IEEE Journal of Solid-State Circuits,* vol. 15, pp. 887–894, October 1980.

IPR2023-00698
Theta EX2019

PROBLEMS

21.1 Using simulations, show that the small-signal resistance of a gate-drain-connected PMOS device, Fig. 21.57, behaves like a resistor with a value of $1/g_m$.

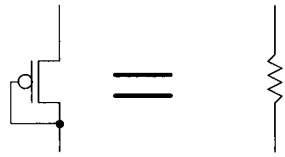
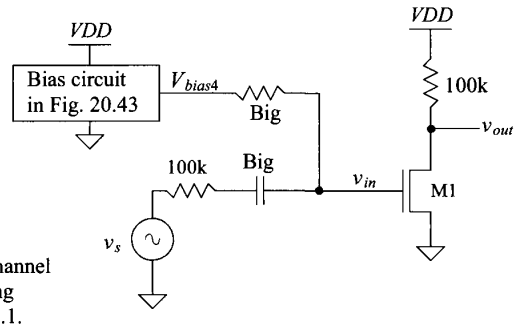


Figure 21.57 The small-signal behavior of a diode-connected MOSFET.

21.2 Estimate the frequency response of the circuit seen in Fig. 21.58. Verify your hand calculations using SPICE.



Use the long-channel sizes and biasing seen in Table 9.1.

Figure 21.58 Amplifier used in Problem 21.2.

21.3 Repeat problem 21.2 if the amplifier drives a 100 *fF* load capacitance.

21.4 For the circuit in Fig. 21.12, estimate the effective transconductance of the circuit, g_{meff}, that relates the MOSFET drain current to the AC input voltage, v_{in}. Verify your solution with SPICE.

21.5 Simulate the operation of the common-gate amplifier in Fig. 21.16 using the bias circuit in Fig. 20.43 and the device sizes seen in Table 9.1. Compare the simulation results to hand calculations.

21.6 Estimate the transfer function of the amplifier in Ex. 21.6 if it drives a 100 fF load. Verify your hand calculations using simulations.

21.7 Determine the frequency response of the amplifier seen in Fig. 21.59. Verify your hand calculations using simulations.

IPR2023-00698
Theta EX2019

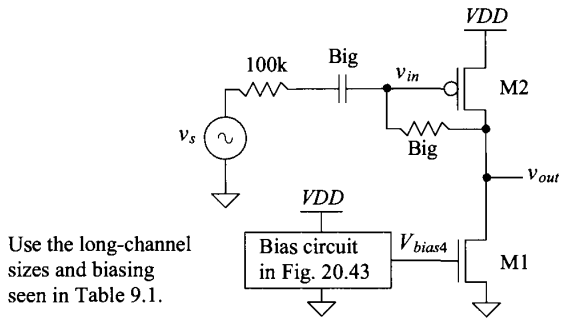


Figure 21.59 Amplifier for Problem 21.7.

21.8 Repeat Ex. 21.9 using the biasing circuit from Fig. 20.47 and the sizes in Table 9.2.

21.9 Repeat Ex. 21.10 using the biasing circuit from Fig. 20.47 and the sizes in Table 9.2.

21.10 Repeat Ex. 21.12 using the biasing circuit from Fig. 20.43 and the sizes in Table 9.1.

21.11 Show, using SPICE simulations and an ideal current source of 10 μA (device size from Table 9.2), the difference in the voltage gains with and without body effect for the circuit seen in Fig. 21.60.

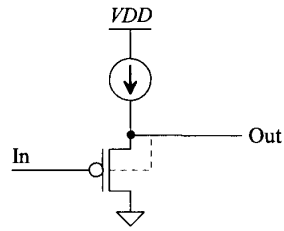


Figure 21.60 Gain of an SF with and without body effect for Problem 21.11.

21.12 Does the SF in Fig. 21.46 exhibit slew-rate limitations if it needs to discharge the 1-pF load capacitance quickly? Why or why not? Verify your answer using simulations.

21.13 Simulate the operation of the class AB output buffer in Fig. 21.48 using the bias circuit from Fig. 20.47 and the sizes in Table 9.2. If the buffer is driving a load resistor of 10k, plot v_{in} against v_{out}. What is the linear output range?

IPR2023-00698
Theta EX2019

21.14 In the class AB output buffer in Fig. 21.50 or 21.51, a load resistor connected to ground causes the PMOS device to conduct more current then the NMOS. Why? Resimulate the buffer in Fig. 21.51 if it drives a 1k resistor and MON is reduced in size to 50/2. Is this a better design? Why or why not?

21.15 Using simulations, show the problem of using an inverter output buffer without a floating current source as seen in Fig. 21.53. (The quiescent current that flows in the MOSFETs is huge and not accurately controlled as it is in Fig. 21.50.)

21.16 Is the distortion the output buffer in Fig. 21.56 introduces into a signal a function of the load resistance? Verify your answer with simulations and show some time-domain waveforms with and without distortion.

IPR2023-00698
Theta EX2019

Chapter
22

Differential Amplifiers

In the last chapter big resistors and capacitors were used to bias the circuits to the correct operating point, as seen in Fig. 21.21. The DC operating voltage on the gate of M1 (in Figs. 21.17 or 21.21) is extremely important when biasing the amplifier. If it's not at precisely the correct value, then the current sourced by M2 won't equal the current in M1 when both MOSFETs are operating in the saturation region.

The differential amplifier (*diff-amp*) is used on the input of an amplifier to allow input voltages to move around so that biasing of the gain stages isn't affected (that is, so it isn't a function of the input voltage). The diff-amp is a fundamental building block in CMOS analog integrated circuit design, and an understanding of its operation and design is extremely important. In this chapter we discuss three basic types of differential amplifiers: the source-coupled pair, the source cross-coupled pair, and the current differential amplifier.

22.1 The Source-Coupled Pair

The source-coupled pair comprised of M1 and M2 is shown in Fig. 22.1. When M1 and M2 are used in this configuration they are sometimes called a *diff-pair*.

22.1.1 DC Operation

The diff-pair in Fig. 22.1 is biased with a current source so that

$$I_{SS} = i_{D1} + i_{D2} \tag{22.1}$$

If we label the input voltages at the gates of M1 and M2 as v_{I1} and v_{I2}, we can write the input difference as

$$v_{DI} = v_{I1} - v_{I2} = v_{GS1} - v_{GS2} \tag{22.2}$$

or in terms of the AC and DC components of the differential input voltage, v_{DI},

$$v_{DI} = V_{GS1} + v_{gs1} - V_{GS2} - v_{gs2} \tag{22.3}$$

When the gate potentials of M1 and M2 are equal, then (assuming both are operating in the saturation region)

IPR2023-00698
Theta EX2019

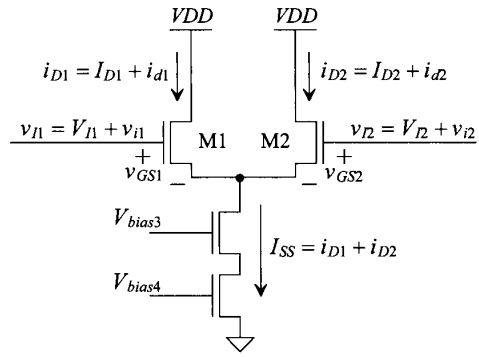


Figure 22.1 A basic NMOS source-connected pair made using the diff-pair M1 and M2.

$$I_{D1} = I_{D2} = \frac{I_{SS}}{2} \tag{22.4}$$

Maximum and Minimum Differential Input Voltage

Since we know that a saturated MOSFET follows the relation

$$i_D = \frac{\beta_n}{2}(v_{GS} - V_{THN})^2 \tag{22.5}$$

the difference in the input voltages may be written as

$$v_{DI} = \sqrt{\frac{2}{\beta_n}}\left(\sqrt{i_{D1}} - \sqrt{i_{D2}}\right) \tag{22.6}$$

The maximum difference in the input voltages, v_{DIMAX} (maximum differential input voltage), is found by setting i_{D1} to I_{SS} (M1 conducting all of the tail bias current) and i_{D2} to 0 (M2 off)

$$v_{DIMAX} < v_{I1} - v_{I2} = \sqrt{\frac{2 \cdot L \cdot I_{SS}}{KP_n \cdot W}} \tag{22.7}$$

The minimum differential input voltage, v_{DIMIN}, is found by setting i_{D2} to I_{SS} and i_{D1} to 0

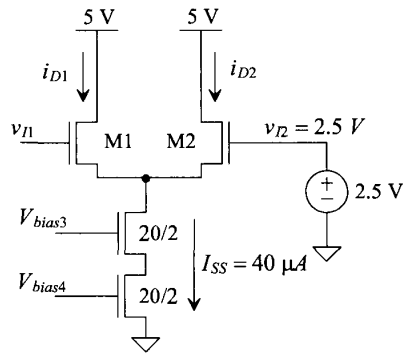
$$v_{DIMIN} = -v_{DIMAX} = -(v_{I1} - v_{I2}) = -\sqrt{\frac{2 \cdot L \cdot I_{SS}}{KP_n \cdot W}} \tag{22.8}$$

Example 22.1

Estimate the maximum and minimum voltage on the gate of M1 in Fig. 22.2 that ensures that neither M1 or M2 shut off. Verify your hand calculations with SPICE simulations.

Note that the diff-amp tail current, I_{SS}, is 40 μA (the widths of the MOSFETs were doubled from the sizes indicated in Table 9.1). When $v_{I1} = v_{I2} = 2.5$ V, the differential input voltage, v_{DI}, is 0 and the current flowing in M1 and M2 is 20 μA ($= i_{D1} = i_{D2}$).

IPR2023-00698
Theta EX2019



Parameters from Table 9.1
Bias circuit from Fig. 20.43

Figure 22.2 Diff-amp used in Ex. 22.1.

Using Eq. (22.7), we can estimate the maximum voltage on the gate of M1 as

$$v_{I1MAX} = \sqrt{\frac{2 \cdot L \cdot I_{SS}}{KP_n \cdot W}} + v_{I2} = \sqrt{\frac{2 \cdot 2 \cdot 40}{120 \cdot 10}} + 2.5 = 2.865\ V$$

or v_{I1} is 365 mV above v_{I2}. The minimum voltage allowed on the gate of M1 (to keep some current flowing in M1) is then 365 mV below v_{I2} or 2.135 V. The simulation results are seen in Fig. 22.3. Notice how, when the inputs are equal (both are 2.5 V), the currents in M1 and M2 are equal and approximately 20 μA ($= I_{SS}/2$). ■

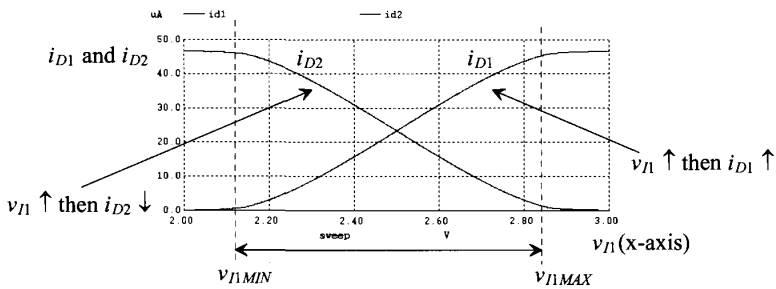


Figure 22.3 Simulating the operation of the diff-amp in Fig. 22.2.

Maximum and Minimum Common-Mode Input Voltage

When the diff-amp is used on the input of an op-amp, the inputs are forced, via feedback around the op-amp, to the same values (or very nearly the same values). This value (or more precisely the average of the two inputs) is called the *common-mode voltage*. Figure 22.4 shows the diff-amp with both inputs tied together. We're interested in the maximum and minimum voltage that will keep both M1 and M2 operating in the saturation region

IPR2023-00698
Theta EX2019

(that is, not off or in the triode region). We'll call the maximum common-mode voltage, v_{CMMAX} and the minimum common-mode voltage, v_{CMMIN}. Note that when the common-mode voltage (the potential on the gates of M1 and M2) is too high, both M1 and M2 triode (behave like resistors). When the common-mode voltage is too low, M1 and M2 shut off.

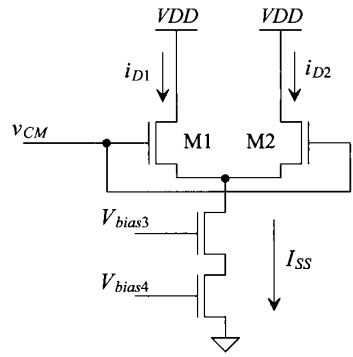


Figure 22.4 Diff-amp used to calculate the minimum and maximum diff-amp input voltages.

To determine the maximum input voltage, we know that for M1 and M2 to remain in the saturation region

$$V_{DS} \geq V_{GS} - V_{THN} \rightarrow V_D \geq V_G - V_{THN} \tag{22.9}$$

Because $V_D = VDD$, we can write

$$V_{CMMAX} = VDD + V_{THN} \tag{22.10}$$

The input common-mode voltage can actually be higher than *VDD* before M1/M2 move into the triode region.

For the minimum input voltage, we can write

$$V_{CMMIN} = V_{GS1,2} + 2 \cdot V_{DS,sat} \tag{22.11}$$

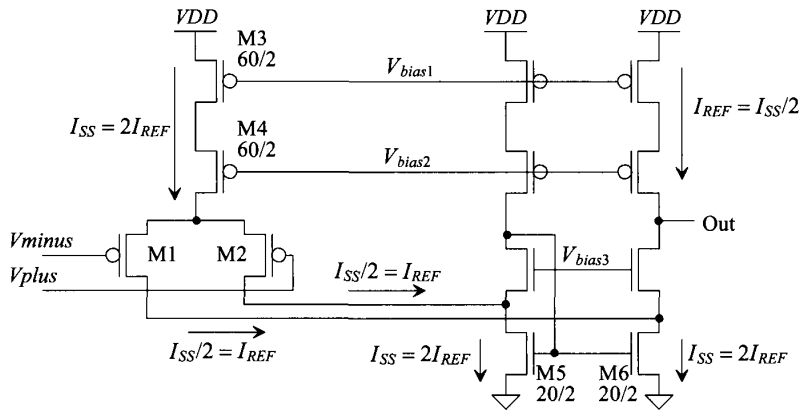
where the minimum voltage across the current source is assumed to be $2V_{DS,sat}$. For the long-channel process, from Table 9.1, the minimum input voltage for an NMOS diff-amp is 1.55 V.

Example 22.2

Figure 22.5 shows a folded cascode amplifier (see Fig. 20.45). Assuming that all MOSFETs are operating in the saturation region, estimate the minimum and maximum input voltage of the amplifier.

Note how the widths of M5–M6 are doubled to sink the additional current from the diff-amp. It's important to understand how the sizes of the MOSFETs are adjusted in the current mirrors so that the currents sum correctly. The currents, as seen in Fig. 22.5, are labeled assuming *Vplus* = *Vminus*.

IPR2023-00698
Theta EX2019



Parameters from Table 9.1
Unlabeled PMOS are 30/2
Unlabeled NMOS are 10/2
Bias circuit from Fig. 20.43

Figure 22.5 Folded cascode amplifier.

The maximum allowable input common-mode voltage is determined by

$$V_{CMMAX} = VDD - V_{SG} - 2V_{SD,sat} = 5 - 1.15 - 0.5 = 3.35\ V$$

The minimum input common-mode voltage is determined by noting the drain voltage of M5 and M6 (or M1 and M2) is a $V_{DS,sat}$

$$V_{SD} \geq V_{SG} - V_{THP} \rightarrow V_D \leq V_G + V_{THP}$$

and so

$$V_{CMMIN} = V_D - V_{THP} = 0.25 - 0.9 = -0.65$$

In other words, the input common-mode voltage can actually be below ground and the amplifier will still function correctly. This result is *very useful*. The folded-cascode amplifier with PMOS diff-amp can be used when the input voltage swings around ground. Note, however, that the output voltage of the amplifier will not swing below ground but is, rather, limited to $2V_{DS,sat}$ above ground and $2V_{SD,sat}$ below *VDD* if the MOSFETs are to remain in saturation (the high-gain region of operation). ■

Current Mirror Load

The previous uses of the diff-amp have exploited the fact that the output of the diff-amp is a current controlled by a differential input voltage, v_{DI}. In some applications, it is desirable to have a diff-amp with a voltage output. Towards this goal, consider the diff-amp with a current mirror load, as seen in Fig. 22.6. An imbalance in the drain currents of M1 and M2 causes the output of the diff-amp to swing either towards *VDD* or

IPR2023-00698
Theta EX2019

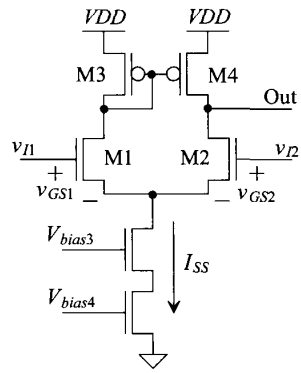


Figure 22.6 Diff-amp with a current mirror load.

ground. The minimum input common mode voltage is, once again, given by Eq. 22.11. The maximum input common mode voltage is determined knowing that the drain voltage of M2 is the same as the drain voltage of M1 (when both diff-amp inputs are the same potential), that is, $VDD-V_{SG}$ of the PMOS. We can therefore write

$$V_{DS} \geq V_{GS} - V_{THN} \rightarrow V_D \geq V_G - V_{THN} \rightarrow V_{CMMAX} = VDD - V_{SG} + V_{THN} \quad (22.12)$$

For the parameters in Table 9.1, $V_{CMMAX} = 5 - 1.15 + 0.8 = 4.65\ V$.

The voltage output swing can be determined by noting that the maximum voltage output is limited by keeping M4 in saturation. Therefore,

$$V_{OUTMAX} = VDD - V_{SD,sat} \quad (22.13)$$

The minimum output voltage is determined by the voltage on the gate of M2 (M2 must remain in saturation).

$$V_D \geq V_G - V_{THN} \rightarrow V_{OUTMIN} = V_{I2} - V_{THN} \quad (22.14)$$

Example 22.3

Suppose the diff-amp in Fig. 22.6 is implemented with the sizes and bias currents seen in Table 9.1 (I_{SS} = 40 μA). If the gate of M2 is held at 4 V, estimate the diff-amp's output swing. Verify the answers with SPICE.

From Eq. (22.13), the output can swing up to 4.75 V before M4 triodes. From Eq. (22.14), the output can swing down to 4–0.8 or 3.2 V before M2 triodes. The simulation results are seen in Fig. 22.7. Looking at the results, we see the output goes down to approximately 2.8 V or 400 mV less than what we predicted. This can be attributed to the body effect in M1 and M2. With body effect, the threshold voltage is 1.2 V (instead of 0.8 V). ■

IPR2023-00698
Theta EX2019

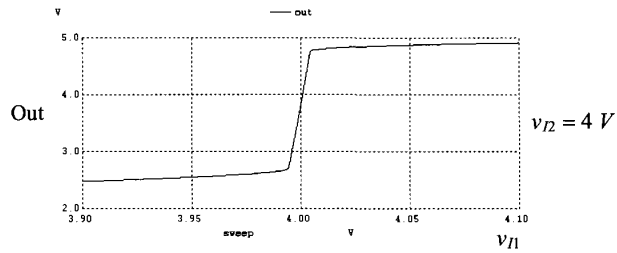


Figure 22.7 Simulation results for Ex. 22.3

Biasing from the Current Mirror Load

Consider the connection of the common-source amplifier, M7, to the output of the diff-amp in Fig. 22.8. When the inputs to the diff-amp are at the same potential, the currents that flow in M3 and M4 are equal (= $I_{SS}/2$). We know from Ch. 20 that the drain of M4 is then at the same potential as its gate. This means, for biasing purposes, that the gate of M7 can be treated as if it were tied to the gate of M3 (M4).

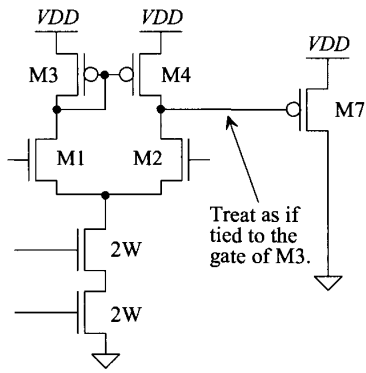


Figure 22.8 Using a diff-amp to bias the next stage amplifier.

Minimum Power Supply Voltage

Looking at the diff-amp in Fig. 22.6, we can estimate the minimum allowable power supply voltage, VDD_{min}, as

$$VDD_{min} = V_{SG3} + \overbrace{V_{DS,sat1}}^{\text{minimum voltage across M1}} + \overbrace{2V_{DS,sat}}^{\text{minimum voltage across Iss}} \qquad (22.15)$$

Looking at this equation, we see that the V_{SG} of M3 is the "weak link" in the minimum power supply voltage. Using the parameters from Table 9.2 (the short-channel devices with a *VDD* of 1 V), we get a VDD_{min} of 500 mV. We can lower this value by using a single MOSFET to bias the diff-amp (instead of a cascode circuit).

IPR2023-00698
Theta EX2019

22.1.2 AC Operation

To describe the AC small-signal operation of the diff-amp in Fig. 22.1, consider the AC schematic seen in Fig. 22.9. We've replaced the current source with an open and the DC supply (*VDD*) with a short (to ground). In the following discussion, it's important to remember that a negative AC current simply means that the overall current (AC + DC) is decreasing (see Ex. 9.5 in Ch. 9). We can write

$$v_{di} = v_{i1} - v_{i2} = v_{gs1} - v_{gs2} = \frac{i_{d1}}{g_m} - \frac{i_{d2}}{g_m} \tag{22.16}$$

Reviewing Fig. 22.9, we see that

$$i_{d1} = -i_{d2} = i_d \tag{22.17}$$

and thus

$$v_{gs1} = -v_{gs2} = \frac{v_{di}}{2} \tag{22.18}$$

so finally

$$g_m v_{gs1} = g_m(-v_{gs2}) = i_d = g_m \cdot \frac{v_{di}}{2} \tag{22.19}$$

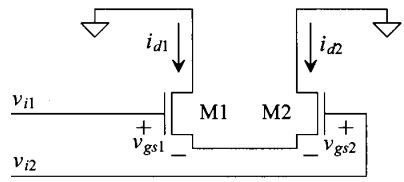


Figure 22.9 AC circuit for a diff-amp.

Example 22.4
Estimate the AC components of the drain currents of M1 and M2 for the circuit in Fig. 22.2 if $v_{I1} = 2.5 + 1mV \cdot \sin(2\pi \cdot 1kHz \cdot t)$. Verify your hand calculations using SPICE.

From Table 9.1 we know the g_m of the NMOS devices used in the diff-pair is 150 $\mu A/V$. The AC component of v_{I2} is zero so we know

$$v_{gs1} = -v_{gs2} = 0.5\ mV$$

and thus

$$i_d = g_m v_{gs} = (150 \times 10^{-6})(500 \times 10^{-6}) = 75\ nA$$

This is the AC component of the drain currents. When i_{d1} = 75 nA, then i_{d2} = –75 nA. The overall drain currents are written as

$$i_{D1} = 20\ \mu A + (75\ nA) \cdot \sin(2\pi 1k \cdot t) \text{ and } i_{D2} = 20\ \mu A - (75\ nA) \cdot \sin(2\pi 1k \cdot t)$$

The simulation results are seen in Fig. 22.10. ■

IPR2023-00698
Theta EX2019

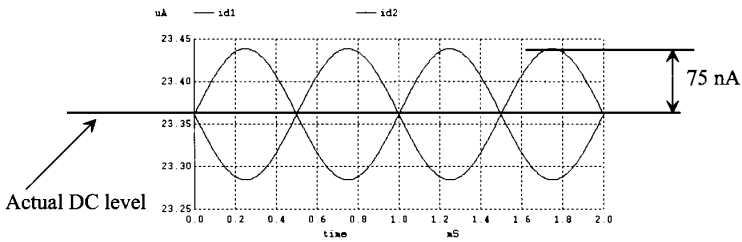


Figure 22.10 Simulation results for Ex. 22.4.

AC Gain with a Current Mirror Load

To determine the AC gain of the diff-amp with a current mirror load, Fig. 22.6, consider the small signal model seen in Fig. 22.11. We've replaced the diode-connected MOSFET, M3, with a $1/g_{m3}$ resistor. Also the resistance looking into the drain of M4 (r_{o4}) and the resistance looking into the drain of M2 (r_{o2}) are drawn explicitly. We assumed that the small-signal resistance looking into the drain of M2 is simply r_{o2} (for all practical design cases this is a good assumption). Note how, because of the current mirror action, the AC current flowing in M4 is i_{d1}. The current in M3 is mirrored by M4. The output voltage can be written as

$$v_{out} = (i_{d1} - i_{d2}) \cdot (r_{o2} || r_{o4}) \tag{22.20}$$

Knowing $i_{d1} = -i_{d2} = i_d$

$$v_{out} = 2i_d \cdot (r_{o2} || r_{o4}) \tag{22.21}$$

The differential mode gain is then written with the help of Eq. (22.19) as

$$A_d = \frac{v_{out}}{v_{di}} = \frac{v_{out}}{v_{i1} - v_{i2}} = g_m \cdot (r_{o2} || r_{o4}) \tag{22.22}$$

Notice that as the voltage on the gate of M1 increases, the current in M1 (and M3/M4) increases. This causes the output voltage to increase. When the gate potential of M2 increases, so does its drain current, causing the output voltage to decrease. Sometimes the gate potential of M1 is called the *noninverting* input and the gate potential of M2 is called the *inverting* input.

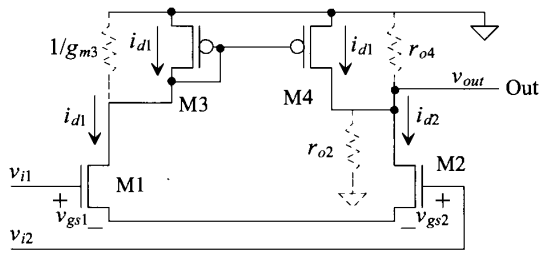


Figure 22.11 AC circuit of the diff-amp with a current-source load.

IPR2023-00698
Theta EX2019

Example 22.5
Determine the output voltage for the diff-amp seen in Fig. 22.12. Verify your hand calculations using SPICE.

The tail current of the diff-amp conducts 20 μA so M1 and M2, when the inputs are equal, each conduct 10 μA. Using the parameters from Table 9.2 and Eq. (22.22), the gain of the diff-amp is

$$A_d = g_m \cdot (r_{on} || r_{op}) = (150 \times 10^{-6}) \cdot \left(\frac{167 \cdot 333}{167 + 333} \times 10^3 \right) = 16.7 \; V/V$$

This means that the 1 mV AC input will appear as a 16.7 mV AC output. The DC level on the output is $VDD - V_{SG} = 650\,mV$. The output voltage is then

$$v_{out}(t) = 0.65 + (16.7\,mV) \cdot \sin(2\pi 10MHz \cdot t)$$

The simulation results are seen in Fig. 22.13. ■

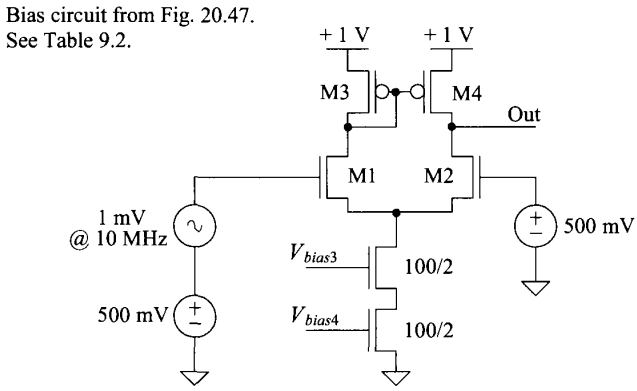


Figure 22.12 Circuit used in Ex. 22.5.

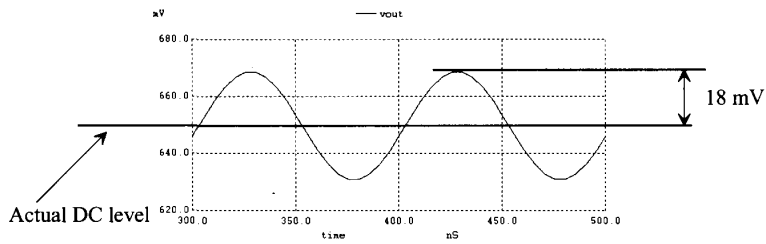


Figure 22.13 The output of the circuit in Fig. 22.12. Note that the transient analysis doesn't show the beginning of the simulation where the reference circuit is starting up.

IPR2023-00698
Theta EX2019

Example 22.6

Estimate the f_{3dB} of the diff-amp in Ex. 22.5 if it drives a load capacitance of 1 pF. Verify your hand calculations with SPICE.

The resistance on the output node is $r_{o2}||r_{o4} = 111\ k\Omega$; therefore, since this is the *high-impedance node*, we can estimate the circuit's 3-dB frequency in the circuit as

$$f_{3dB} = \frac{1}{2\pi \cdot 111k \cdot 1\ pF} = 1.4\ MHz$$

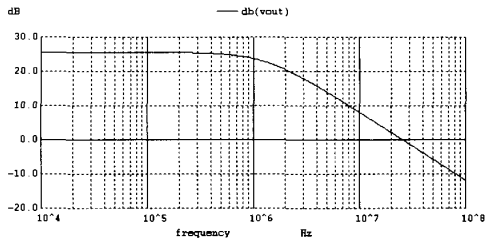


Figure 22.14 The AC response of the diff-amp in Fig. 22.12 when driving a 1 pF load capacitance.

The simulation results are seen in Fig. 22.14. ■

22.1.3 Common-Mode Rejection Ratio

An important aspect of the differential amplifier is its ability to reject a common signal applied to both inputs. Often, in analog systems, signals are transmitted differentially, and the ability of an amplifier to reject coupled noise into each line is very desirable. Consider the amplifier shown in Fig. 22.15. The bias current at the sources of M1/M2 has been replaced with its small-signal output resistance. The common-mode signal (the "noise" labeled in the figure) on both inputs is, ideally, rejected by the diff-amp. The diff-amp's output voltage doesn't vary. The noise causes the source potentials of M1/M2 to vary (the voltage across the tail current source).

If we apply an AC signal, v_c, to the gates of M1/M2, equivalent to saying that a common signal is applied to the input of the differential amplifier, we can calculate the common-mode gain. We begin by writing the AC small-signal, common-mode input voltage, v_c, as

$$v_c = v_{gs1,2} + 2i_d R_o \tag{22.23}$$

The MOSFETs M1 and M2 each source a current, i_d, through the output resistance of the current source, R_o (hence the factor of two in this equation). This equation may be rewritten as

$$v_c = i_d\left(\frac{1}{g_m} + 2R_o\right) \approx i_d 2R_o \tag{22.24}$$

The output voltage, because of the symmetry of the circuit, is given by

IPR2023-00698
Theta EX2019

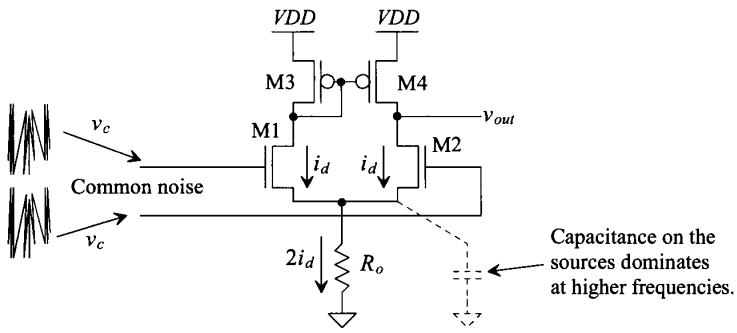


Figure 22.15 Calculating the CMRR of a diff-amp.

$$v_{out} = -i_d \cdot \frac{1}{g_{m3}} = -i_d \cdot \frac{1}{g_{m4}} \tag{22.25}$$

The common-mode gain is then

$$A_c = \frac{v_{out}}{v_c} = \frac{-1/g_{m3,4}}{2R_o} = -\frac{1}{2g_{m3,4}R_o} \tag{22.26}$$

Notice that by increasing the output resistance of the biasing current source, the common-mode gain drops. (This is good because we don't want any common signal to be amplified by the diff-amp.) The common-mode rejection ratio (*CMRR*), in dB, of a diff-amp with a current mirror load is given by

$$CMRR = 20 \cdot \log\left|\frac{A_d}{A_c}\right| = 20 \cdot \log[g_{m1,2}(r_{o2} \| r_{o4}) \cdot 2g_{m3,4}R_o] \tag{22.27}$$

The larger the *CMRR,* the better the performance of the diff-amp. For the diff-amp in Fig. 22.12, Eq. (22.27) evaluates to 86 dB. As the frequency of the input common-mode signal, v_c , increases, the *CMRR* goes down. This is the result of the impedance of the capacitance to ground on the sources of the diff-pair becoming much smaller than R_o.

Example 22.7

Simulate the *CMRR* of the diff-amp in Fig. 22.12. Show that the *CMRR* falls with increasing frequency.

As mentioned above, the (low-frequency) calculated *CMRR* is 86 dB. To simulate the *CMRR,* we perform AC simulations on two of the diff-amps in Fig. 22.12 so we can simultaneously determine A_d and A_c. The ratio of these gains is the *CMRR.* Figure 22.16 shows the simulation results. The common-mode gain, A_c , depends on the DC common-mode voltage. The results seen in Fig. 22.16 are with a common-mode (DC) voltage of 700 mV. Dropping the common-mode voltage down to 500 mV causes the low-frequency *CMRR* to drop down to 50 dB. The reason for this is that, with higher DC common-mode voltages, the voltage across the tail current source is higher, so its output resistance is larger. ■

IPR2023-00698
Theta EX2019

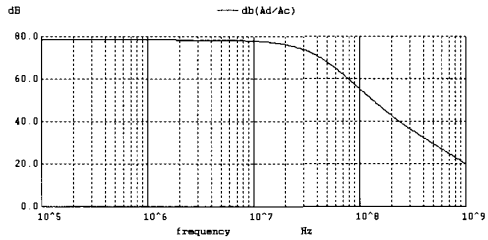


Figure 22.16 The simulated CMRR for the diff-amp in Fig. 22.12.

Input-Referred Offset from Finite CMRR

It's important we understand, in Eq. (22.26), why large R_o reduces A_c. Looking at Fig. 22.15, notice that as v_c varies so too will the voltage across the biasing current source (represented in this figure as R_o). As seen in Ch. 20, Fig. 20.4, variations in the voltage across a current source changes the output of the current source (unless its output resistance is infinite). This variation in current causes the current in M1 and M2 to change with variations in v_c. Changes in the drain current of M1/M2 (and thus M3/M4) causes the output voltage to change. Figure 22.17 shows how the output voltage changes as a function of the DC common-mode voltage. Understanding this circuit is *very important*. If we wanted to hold the diff-amp output voltage constant while varying the common mode voltage, we would have to apply a small voltage difference between the gates of M1 and M2 (*an offset voltage*). We'll revisit this important topic when we discuss op-amp *CMRR*.

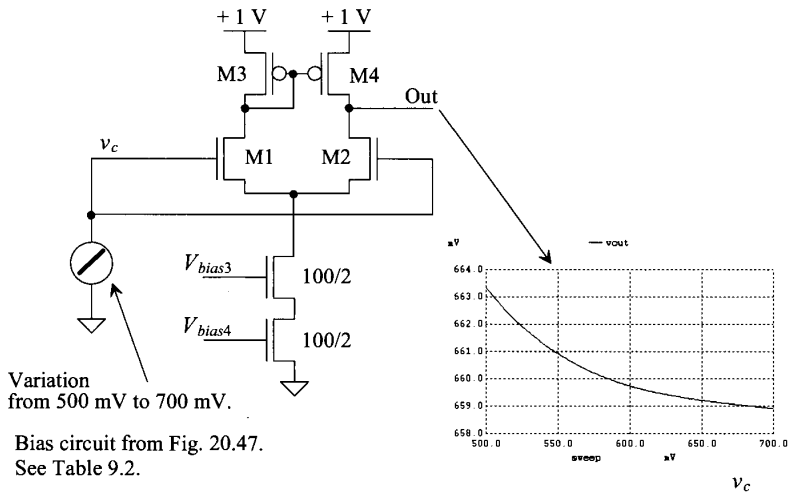


Figure 22.17 How variation in common-mode voltage affects the output voltage.

IPR2023-00698
Theta EX2019

22.1.4 Matching Considerations

The fact that performance depends on the matching of devices is another important aspect of differential amplifiers. Layout techniques can be used to minimize the first-order effects of mismatch due to oxide gradients and other process variations. Figure 22.18 shows how the diff-pair would be laid out in a *common-centroid* configuration (see Ch. 5). Common-centroid layout, as the name implies, constructs two devices symmetrically about a common center in the layout. This allows the two devices to cancel process gradients in both the x and y directions and exposes both devices to heat sources in an identical manner. Note how we've attempted to match the parasitic capacitances of the pair (on the gates and drains) and how we've minimized the capacitance at the sources (to keep the *CMRR* high at higher frequencies).

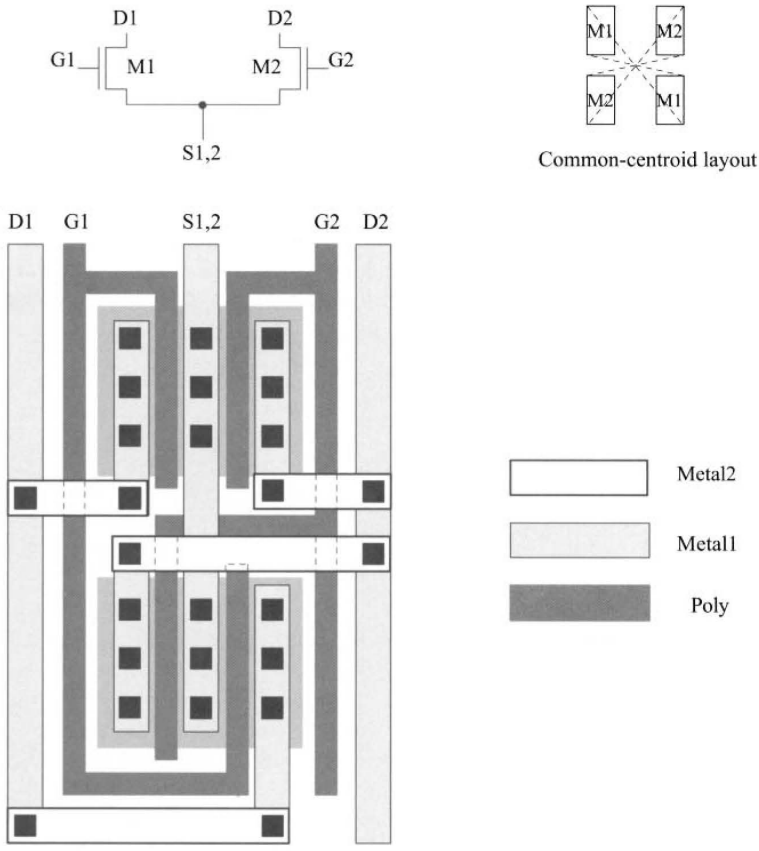


Figure 22.18 Common-centroid layout of a diff-pair.

IPR2023-00698
Theta EX2019

The input-referred offset voltage of a diff-pair, resulting from mismatches in threshold voltage, geometries, and load resistances, can be determined with the help of Fig. 22.19. When the diff-pair is offset-free, connecting the diff-amp's inputs together causes the drain currents and drain voltages of M1 and M2 to be equal ($V_o = 0$). When the load resistors or M1 and M2 are mismatched, V_o is not zero. (We have mismatch and thus V_o is offset from its ideal value.). To drive V_o to zero, we apply a voltage difference between the gates of M1 and M2 (an input-referred offset).

$$V_{OS} = V_{GS1} - V_{GS2} = V_{THN1} + \sqrt{\frac{2I_{D1}}{\beta_1}} - V_{THN2} - \sqrt{\frac{2I_{D2}}{\beta_2}} \qquad (22.28)$$

We can define the differences and averages of the threshold voltage, load resistance, and geometry as ΔV_{THN} (V_{THN1}–V_{THN2}), V_{THN} (average of V_{THN1} and V_{THN2}), ΔR_L (R_{L1}–R_{L2}), R_L (average of R_{L1} and R_{L2}), $\Delta(W/L)$ [(W_1/L_1)–(W_2/L_2)] and (W/L) [average of (W_1/L_1) and (W_2/L_2)]. Thus, we can make the appropriate substitutions into Eq. (22.28), requiring $I_{D1}R_{L1} = I_{D2}R_{L2}$, with the result:

$$V_{OS} = \Delta V_{THN} + \frac{V_{GS} - V_{THN}}{2} \cdot \left[\frac{-\Delta R_L}{R_L} - \frac{\Delta(W/L)}{(W/L)} \right] \qquad (22.29)$$

The threshold voltage mismatch must be reduced using layout techniques. Mismatches resulting from unequal geometry and differences in the load resistance can be reduced by designing with a small V_{GS} (V_{GS} close to V_{THN}). This should be compared with Eq. (20.8), which shows that using a small V_{GS} in a current mirror results in a large mirrored current difference.

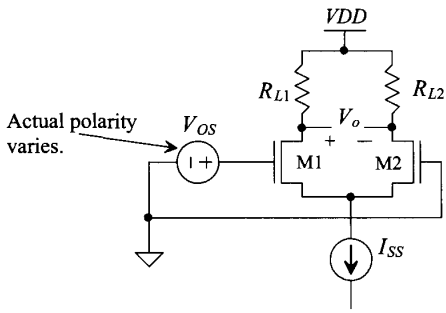


Figure 22.19 Determining diff-amp input offset voltage.

Input-Referred Offset with a Current Mirror Load

Suppose the diff-amp with current mirror load, Fig. 22.12, has an offset (the devices aren't perfectly matched). What this means is that the drain potential of M4 (M2) is not equal to the drain potential of M3 (M1 and most likely the drain currents in M1 and M2 aren't equal as well). The input-referred offset would then be the voltage difference on the gates of M1 and M2 needed to drive the output voltage to the correct value. Using Eq. (22.22), we can write the input-referred offset in terms of the difference-mode gain as

IPR2023-00698
Theta EX2019

$$|V_{OS}| = \frac{V_{o,ideal}}{|A_d|} = \frac{V_{o,ideal}}{g_m \cdot (r_{o2} || r_{o4})} \tag{22.30}$$

which indicates that to reduce the input-referred offset voltage we simply need to design the diff-amp with large gain. This is the same result as that of Eq. (22.29), if we pause and think about it. To reduce the offset, we said we design with a small V_{GS}. If we hold the diff-amp's bias current constant and reduce V_{GS} (by increasing the widths of M1 and M2), we are increasing g_m.

Finally, although not clearly indicated by Eq. (22.30) as it is in Eq. (22.29), a mismatch in the diff-pair's threshold voltage can only be reduced by layout techniques (e.g., using common-centroid layout). Using large devices (larger-width MOSFETs), reduces variation in device characteristics and leads to better matching.

22.1.5 Noise Performance

The diff-amp with noise sources is seen in Fig. 22.20. The noise from M6 feeds equally into M1 and M2 (see Fig. 21.44 for comments about noise contributions from M5) and thus, because of the current mirror action in M3 and M4, ideally doesn't affect the output voltage. Therefore, we will ignore noise from M6 in the following discussion.

The diff-amp's output noise power spectral density (PSD) is given by

$$V_{onoise}^2(f) = \left(I_{M1}^2(f) + I_{M2}^2(f) + I_{M3}^2(f) + I_{M4}^2(f)\right) \cdot (r_{o2} || r_{o4})^2 \tag{22.31}$$

The input-referred noise PSD is then

$$V_{inoise}^2(f) = \frac{V_{onoise}^2(f)}{A_d^2} \tag{22.32}$$

Substituting in the difference-mode gain of the diff-amp, Eq. (22.22), we get

$$V_{inoise}^2(f) = \frac{I_{M1}^2(f) + I_{M2}^2(f) + I_{M3}^2(f) + I_{M4}^2(f)}{g_m^2} \tag{22.33}$$

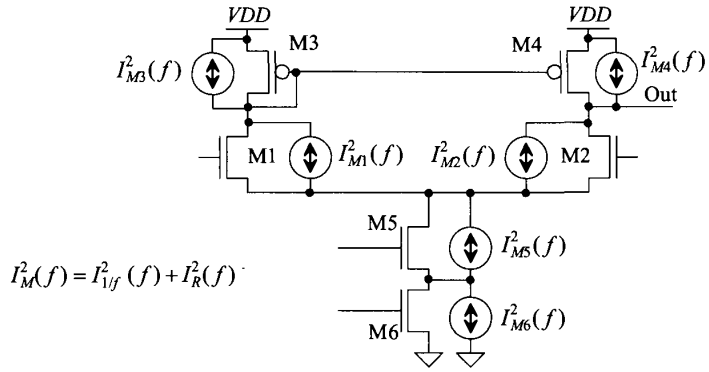


Figure 22.20 Noise model for the diff-amp.

IPR2023-00698
Theta EX2019

By maximizing the transconductance, g_m, of the diff-pair, we can reduce the input-referred noise. This statement alone is a little misleading. As we saw in Eqs. (9.63)–(9.65), increasing the drain current increases the noise PSD. So to reduce the input-referred noise, we would want to maximize the transconductance by increasing the diff-pair's width (make M1 and M2 wide), not by increasing I_{SS}.

22.1.6 Slew-Rate Limitations

Like the class A amplifiers studied in the last chapter, the basic diff-pair also exhibits slew-rate limitations. As seen in Fig. 22.21, when either M1 or M2 turns off, the total current available to charge a load capacitance is I_{SS}. The slew-rate is then (see also Eq. (21.38))

$$\text{slew rate} = \frac{dV_{out}}{dt} = \frac{I_{SS}}{C_L} \tag{22.34}$$

For high-speed, we need a large bias current. However, this causes excess power dissipation. Once again, to get high-speed (large current drive capability) with low quiescent power dissipation, let's look at a class AB configuration.

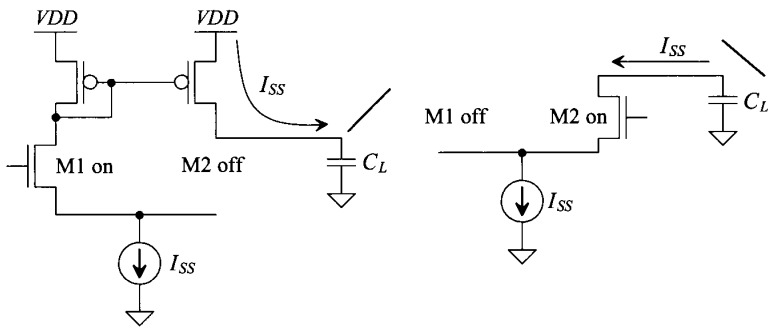


Figure 22.21 Slew-rate limitations in diff-amps.

22.2 The Source Cross-Coupled Pair

The source cross-coupled pair is shown in Fig. 22.22. This configuration is very useful in practical circuit design because it can eliminate slew-rate limitations. A diff-amp built with the source cross-coupled pair can operate in the class AB mode with significant output drive current. Assuming that both inputs to the pair are connected to a (the same) voltage within the pair's common-mode input range and that all NMOS are sized the same and that all PMOS are sized the same, a current I_{SS} flows in all the MOSFETs in the circuit. Note that M11, M21, M31, and M41 are simply behaving as biasing batteries. Their gate-source voltages are mirrored by M1–M4 to set the DC operating current. Using the parameters in Table 9.1 with the gates of M1 and M2 (v_{I1} and v_{I2}) at 3.5 V and I_{SS} of 20 μA, the gate of M3 is at a constant potential that we will label v_{I2}–V_{bias} (or $V_{bias} = V_{GS21} + V_{SG31} = 2.2$ V neglecting body effect).

IPR2023-00698
Theta EX2019

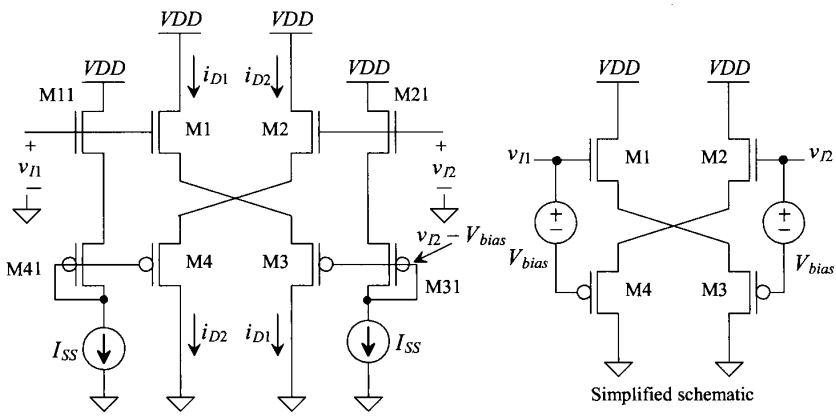


Figure 22.22 Source cross-coupled differential amplifier with NMOS input devices.

Operation of the Diff-Amp

An important characteristic of the source cross-coupled differential amplifier is that as v_{DI} is increased, i_{D1} continues to increase and i_{D2} shuts off, while the opposite is true when v_{DI} decreases. In other words, the amplifier is operating class AB where neither of the output currents is zero as long as their magnitudes remain less than I_{SS}. Looking at the simplified schematic in Fig. 22.22, if v_{I1} increases, then M1 turns on. Because the gate potential of M3 is constant, it turns on as well. In other words, an increase in v_{I1} causes the gate-source voltages of both M1 and M3 to increase (i_{D1} increases). At the same time, the gate potential of M4 increases. This cause both M4 and M2 to shut off (i_{D2} shuts off).

Example 22.8

Simulate the operation of the diff-amp in Fig. 22.22 using the parameters from Table 9.1.

The simulation results are seen in Fig. 22.23. Note that when both inputs are at 3.5 V, the output currents are equal. This figure should be compared to Fig. 22.3.

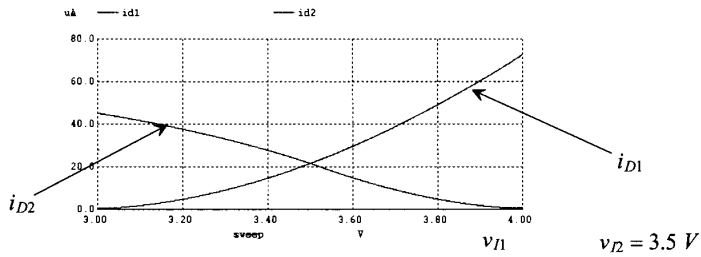


Figure 22.23 The output currents of the diff-amp in Fig. 22.22.

IPR2023-00698
Theta EX2019

Note how in Fig. 22.3, the output currents flatten out above a maximum input voltage. Here the current continues to increase. ■

After looking at the results of this example, we might wonder: why use such a complicated diff-amp if the output current is only three times more than what we get with the basic diff-amp discussed in the first section? We can always just scale up the current in a basic diff-amp to get more drive. The benefits of the source cross-coupled diff-amp are seen in special-purpose, *low-power* design. In low-power design, the gate source voltages are close to the threshold voltages (and so low power equates to low speed, see Eq. [9.36]). Consider the amplifier and simulation results seen in Fig. 22.24. Here we've increased the widths of the amplifying devices and reduced their bias current (by sizing up the lengths of the biasing current sources as seen in the figure). When the two inputs are the same voltage, the quiescent current that flows in the diff-amp is a few microamps. However, if one input gets a little larger than the other input, the drain-current increases dramatically. In a low-power design, this topology would be useful to quickly charge a capacitor while burning little quiescent power.

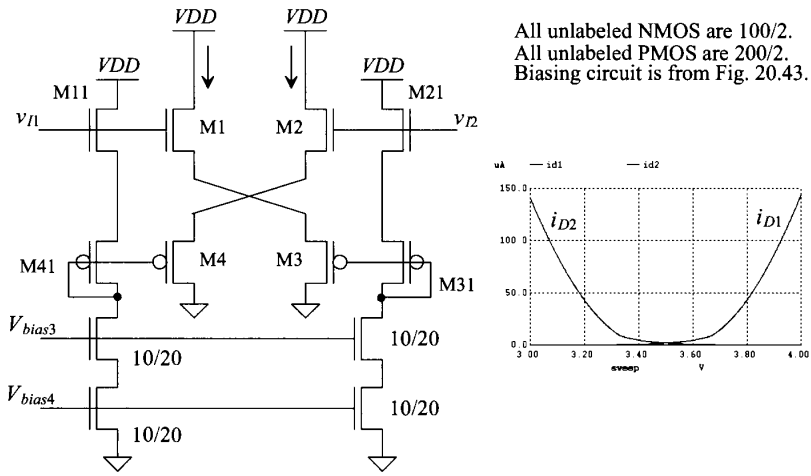


Figure 22.24 Lowering the power in the source cross-coupled diff-amp.

Input Signal Range

The drawback of the source cross-coupled diff-amp is the limited input range. The PMOS version of the source cross-coupled differential amplifier is shown in Fig. 22.25. Normally the version (PMOS or NMOS) of this differential amplifier depends on the input signal range. The NMOS version has the best positive input signal range, while the PMOS version has the best negative input signal range. For centering the input signal range the topology seen in Fig. 22.26 can be used. Again, both an NMOS V_{GS} and a PMOS V_{SG} are used to bias the cross-coupled diff-amp. Note that the biasing devices

IPR2023-00698
Theta EX2019

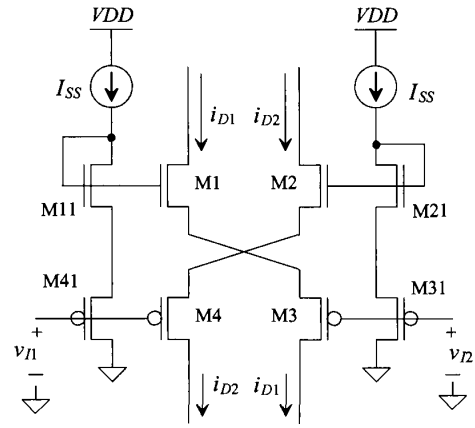


Figure 22.25 Source cross-coupled differential amplifier, PMOS.

(M11, M21, M31, and M41) experience body effect differently than M1–M4. This causes the currents in M1–M4 to be less than the current in the biasing devices. If possible, place M1–M4 and the biasing devices in their own wells to eliminate the body effect. Note that when very low currents are used, the constant current sources driving the gates of M1–M4 can cause slew-rate limitations. To eliminate this problem, capacitors can be placed between the gates of M1/M4 and M2/M3 to make the biasing appear more battery-like.

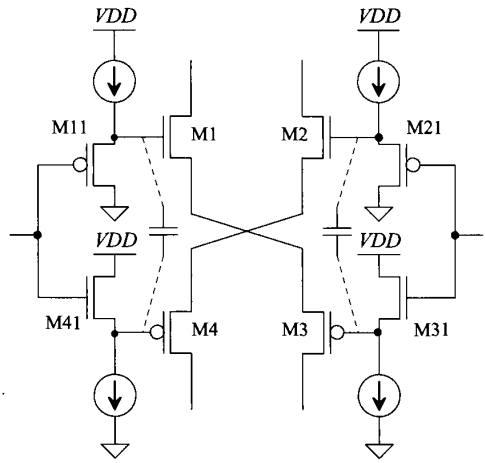


Figure 22.26 Biasing scheme used to shift input common-mode range. Note that the body effect affects the biasing, causing the current in M1–M4 to decrease.

IPR2023-00698
Theta EX2019

22.2.1 Current Source Load

The source cross-coupled differential amplifier with active loads is shown in Fig. 22.27. Note that we could have also added a pair of active loads in series with the drains of M1 and M4 to obtain a differential output amplifier (two complementary amplifier outputs). Superposition can be used to find the AC small-signal gain from the output back to each input. First, determine the small-signal gain, $\frac{v_o}{v_{i1}}$, with v_{i2} at AC ground. Neglecting the bulk effect, we can see that the voltage gain from the gate of M1 to the source of M1 is simply a common drain amplifier of the form

$$\frac{v_{s1}}{v_{i1}} = \frac{g_{m1}R_{Leq}}{1+g_{m1}R_{Leq}} \text{ V/V} \tag{22.35}$$

where R_{Leq} is simply the load seen by the source of M1. This can be found by using a test source (as seen in Fig. 22.28) to determine the effective impedance seen looking into the source of M3. This results in

$$R_{Leq} = \frac{v_t}{i_t} = \frac{1+\frac{1}{g_{m3}r_{o3}}}{g_{m3}+\frac{1}{r_{o3}}} \approx \frac{1}{g_{m3}} \tag{22.36}$$

Therefore, the voltage gain from v_{i1} to the source of M1 becomes

$$\frac{v_{s1}}{v_{i1}} = \frac{g_{m1}\frac{1}{g_{m3}}}{1+g_{m1}\frac{1}{g_{m3}}} \tag{22.37}$$

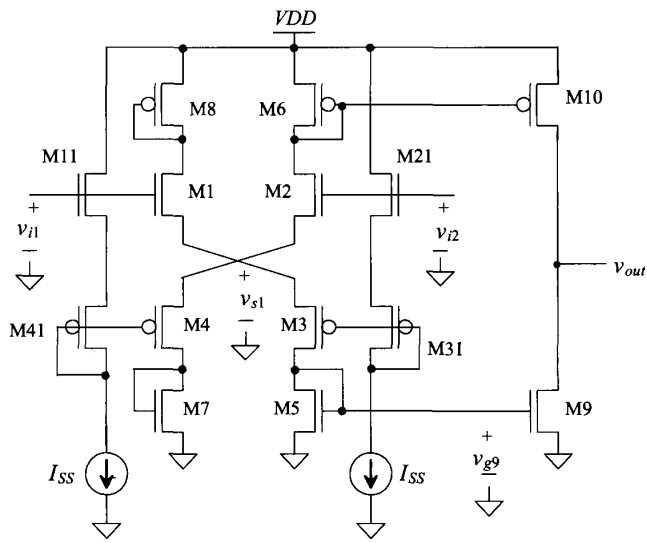


Figure 22.27 Source cross-coupled differential amplifier, with current source loads.

IPR2023-00698
Theta EX2019

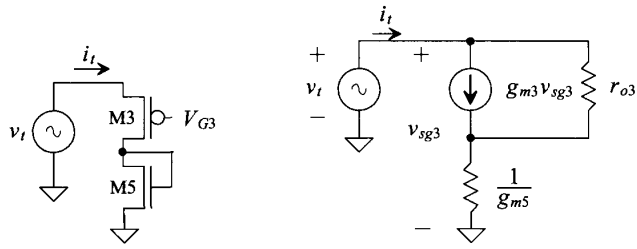


Figure 22.28 (a) Simplified schematic and (b) small-signal model.

The voltage gain from v_{s1} to v_{g9} is the same form as a common gate amplifier. This gain is

$$\frac{v_{g9}}{v_{s1}} = g_{m3}\frac{1}{g_{m5}} \tag{22.38}$$

And finally, the gain from the gate of M9 to the output is simply $-g_{m9}(r_{o9}||r_{o10})$, and

$$\frac{v_o}{v_{i1}} = -\frac{g_{m1}\frac{1}{g_{m5}}}{1+g_{m1}\frac{1}{gm3}}g_{m9}(r_{o9}||r_{o10}) \text{ for } v_{i2}=0 \tag{22.39}$$

Using the same type of analysis, we find that the expression from v_{i2} to the output is

$$\frac{v_o}{v_{i2}} = \frac{g_{m2}\frac{1}{g_{m6}}}{1+g_{m2}\frac{1}{gm4}}g_{m10}(r_{o9}||r_{o10}) \text{ for } v_{i1}=0 \tag{22.40}$$

Superposition can now be used to determine the total effect of each input on the output. Therefore, if M1 and M2 are matched, as are M3 and M4, and if M5 and M6 are designed so that their g_ms are equal as are M7 and M8 (respectively), then the differential gain can be expressed as

$$v_o = (v_{i2}-v_{i1})\frac{2\cdot g_{m1,2}\frac{1}{g_{m5,6}}}{1+g_{m1,2}\frac{1}{gm3,4}}g_{m9,10}(r_{o9}||r_{o10}) \Rightarrow \frac{v_o}{v_{di}} = \frac{2\cdot g_{m1,2}\frac{1}{g_{m5,6}}}{1+g_{m1,2}\frac{1}{gm3,4}}g_{m9,10}(r_{o9}||r_{o10}) \tag{22.41}$$

Input Signal Range

The positive-input, common-mode range (CMR) is determined in exactly the same manner as in the source coupled diff-amp of the first section. The positive CMR is given by

$$v_{CMMAX} \approx VDD - \sqrt{\frac{2I_{SS}}{\beta_6}} = VDD - V_{DS,sat} \tag{22.42}$$

while the negative CMR is given by

$$v_{CMMIN} = \overbrace{\sqrt{\frac{2I_{SS}}{\beta_1}} + V_{THN}}^{V_{GS1}} + \overbrace{\sqrt{\frac{2I_{SS}}{\beta_3}}}^{V_{DS3}} + \overbrace{\sqrt{\frac{2I_{SS}}{\beta_5}} + V_{THN}}^{V_{GS5}} \tag{22.43}$$

IPR2023-00698
Theta EX2019

An alternative current source (current mirror) load configuration is shown in Fig 22.29. The gain of this configuration is given by

$$A_d = \frac{v_{out}}{v_{i1} - v_{i2}} = 2 \cdot \frac{r_{o2} || r_{o6}}{\frac{1}{g_{m1}} + \frac{1}{g_{m3}}} \tag{22.44}$$

assuming $g_{m1} = g_{m2}$ and $g_{m3} = g_{m4}$.

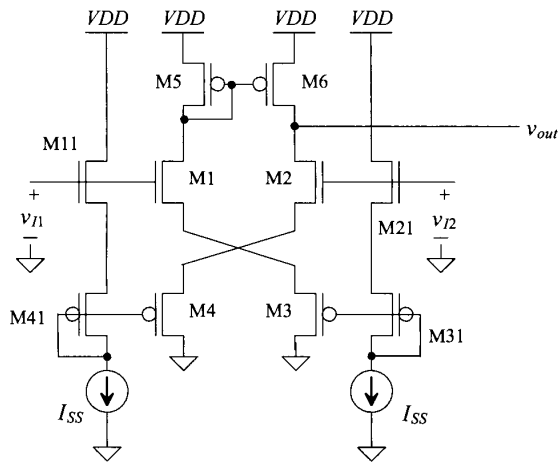


Figure 22.29 Alternative current source load configuration.

22.3 Cascode Loads (The Telescopic Diff-Amp)

The differential amplifier with current source load of Fig. 22.6 gave a voltage gain of $g_m(r_{o2} || r_{o4})$. For many applications, this gain may be too low. Using a cascode current source load in place of M3/M4, results in a gain of approximately $g_m r_{o2}$ (the output resistance of the load becomes much larger than the resistance looking into the drain of M2). This modest improvement in gain can be increased by also cascoding M1 and M2. Figure 22.30 shows the resulting circuit configuration (sometimes called a *telescopic diff-amp*). The MOSFET M6 is selected so that its V_{GS} keeps M1/M2 and MC1/MC2 in the saturation region. The gain of this configuration is given by

$$A_d = g_{m1} \cdot (R_{into\ DC2} || R_{into\ DC4}) \tag{22.45}$$

The resistance looking into the drain of MC2, assuming M2 and MC2 are the same size, is given by

$$R_{into\ DC2} \approx g_{m2} \cdot r_{o2}^2 \tag{22.46}$$

and the resistance looking into the drain of MC4, again assuming MC4 and M4 are the same size, is given by

IPR2023-00698
Theta EX2019

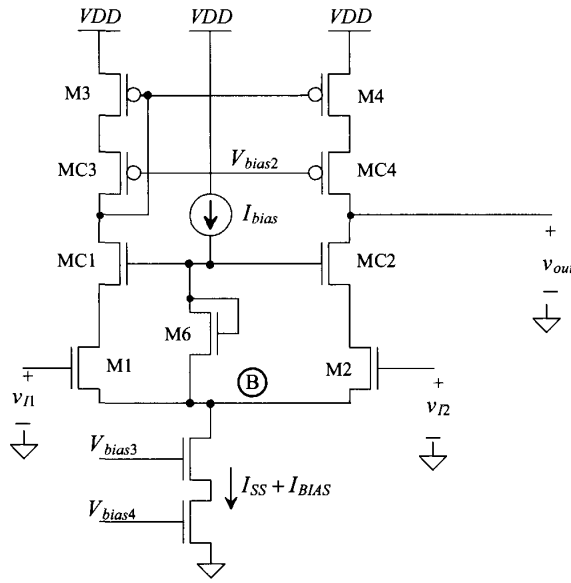


Figure 22.30 Cascode differential amplifier.

$$R_{into\ DC4} \approx g_{m4} \cdot r_{o4}^2 \tag{22.47}$$

The gain of the cascode differential amplifier may be written as

$$A_d = g_{m1}\left(g_{m2}r_{o2}^2 \| g_{m4}r_{o4}^2\right) \tag{22.48}$$

The main drawback of using the cascode differential amplifier is the reduction in positive common-mode range. In the negative direction, V_{CMMIN} is identical to the source-coupled diff-amp discussed in Sec. 22.1. The positive common-mode signal is limited by the additional circuitry required by the cascode loads. Notice that V_{CMMAX} in this case is limited by the amount of voltage necessary to keep M1, MC1, MC3, and M3 in their respective saturation regions.

An interesting issue concerning this circuit is the DC feedback through M6. Because the current through M6 is constant, V_{GS6} will be constant. Normally, an increase in the common-mode voltage on the gates of M1 and M2 causes their drain voltages to decrease. However, because the currents through M1 and M2 are constant, the common source node, marked as node B in Fig. 22.30, increases with the increasing input voltage. This causes the gate potential of MC1 and MC2 to increase. The result is that the drain voltages of M1 and M2 go up (since the gate-source voltages of MC1 and MC2 are constant). This feedback action attempts to keep M1 and M2 from going into nonsaturation. The maximum common-mode voltage, V_{CMMAX}, will be limited by the output voltage and by MC1/MC2 going into nonsaturation.

IPR2023-00698
Theta EX2019

Example 22.9

Determine the minimum and maximum input common-mode voltage for the diff-amp seen in Fig. 22.31. Note that the voltage across M6 is $V_{GS} + V_{DS,sat}$, where V_{GS} is the gate-source voltage of the other NMOS devices in this figure (see Fig. 20.32 and the associated discussion).

Parameters from Table 9.1
Unlabeled PMOS are 30/2.
Unlabeled NMOS are 10/2.
Bias circuit from Fig. 20.43.

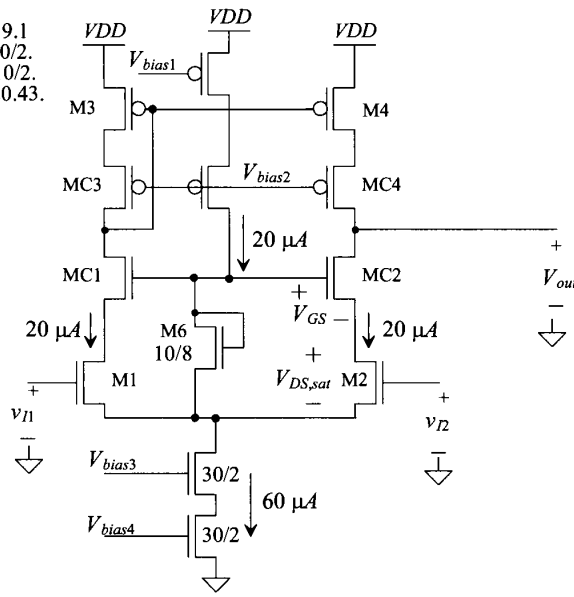


Figure 22.31 Cascode differential amplifier.

The minimum input common-mode voltage is given by

$$V_{CMMIN} = V_{GS1,2} + 2V_{DS,sat} = 1.05 + 0.5 = 1.55\ V$$

As mentioned a moment ago, MC1 and MC2 must be kept in the saturation region. If we assume that the drains of MC1 and MC2 are at $VDD - V_{SG}$ (= 3.85 V), then when MC1 and MC2 are on the verge of trioding (their drain-source voltages are $V_{DS,sat}$), the drains of M1 and M2 are at

$$V_{D1,2} = VDD - V_{SG} - V_{DS,sat} = 3.6\ V$$

We know for an NMOS device (M1 or M2) to be in saturation

$$V_D \geq V_G - V_{THN} \text{ and so } V_{D1,2} \geq V_{CMMAX} - V_{THN}$$

giving a $V_{CMMAX} = 4.4\ V$. In terms of an arbitrary output voltage, we can write

$$V_{CMMAX} = V_{out} - V_{DS,sat} + V_{THN}$$

so if V_{out} is 2.45 V, then $V_{CMMAX} = 3\ V$. ■

IPR2023-00698
Theta EX2019

22.4 Wide-Swing Differential Amplifiers

An increasingly important concern with the advent of low-voltage design is the need for improved common-mode range. A technique for extending the allowable input swing of a diff-amp is to use two complementary diff-amp stages in parallel, as shown in Fig. 22.32. To understand the operation of this circuit, consider the case when the DC component of the input signal, v_{in}, is such that both diff-amps are on (and the AC component is small compared to the DC). The current through the diff-pairs M1/M2 and M9/M10 is I, while the current through the summing MOSFETs M4 and M12 is $2I$. If M5 is the same size as M4 and if M7 is the same size as M12, then a current $2I$ flows in the output transistors. The small-signal voltage gain, assuming $g_{m1} = g_{m2}$ and $g_{m9} = g_{m10}$, is given by

$$A_v = (g_{m1} + g_{m9})[r_{o7}(2I) || r_{o5}(2I)] = \frac{(g_{m1} + g_{m9})}{\lambda_7 2I + \lambda_5 2I} = \frac{\sqrt{2\beta_1 I} + \sqrt{2\beta_9 I}}{2I(\lambda_7 + \lambda_5)} \qquad (22.49)$$

If the input is such that the p-channel diff-amp is on and the n-channel diff-amp is off, then a current I flows in the summing MOSFETs M4 and M12 and zero current flows in M1, M2, M3, and M6. The current that flows in M5 and M7 is now I. The small-signal voltage gain is

$$A_v = g_{m9}[r_{o7}(I) || r_{o5}(I)] = \frac{\sqrt{2\beta_9 I}}{I(\lambda_7 + \lambda_5)} \qquad (22.50)$$

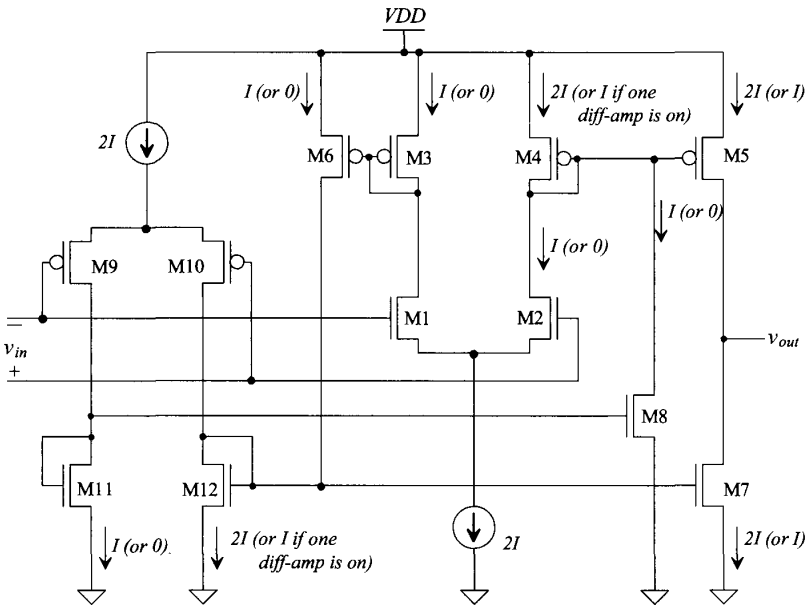


Figure 22.32 Two parallel differential amplifiers used to increase input swing.

IPR2023-00698
Theta EX2019

The small-signal gain when the p-channel diff-amp is off and the n-channel diff-amp on is given by

$$A_v = g_{m1}[r_{o7}(I) \| r_{o5}(I)] = \frac{\sqrt{2\beta_1 I}}{I(\lambda_7 + \lambda_5)} \tag{22.51}$$

It is desirable to transition smoothly from having a single diff-amp on to having both diff-amps on. If we require

$$\beta_1 = \beta_9 = \beta \Rightarrow G_m = \sqrt{2\beta I} \tag{22.52}$$

then Eqs. (22.49) through (22.51) can be rewritten as

$$A_v = G_m \cdot [r_{o7}(I) \| r_{o5}(I)] \tag{22.53}$$

For low distortion, a constant gain is important. Also, for a stable amplifier, it is important that the amplifier can be compensated to ensure stability. An amplifier whose gain depends on the input signal amplitude may be more challenging to compensate for.

22.4.1 Current Differential Amplifier

Another wide-swing, differential amplifier is the current differencing amplifier. The current diff-amp is shown schematically in Fig. 22.33. In the following discussion, let's assume that M1 through M4 are the same size. If both i_1 and i_2 are zero, then a current I_{SS} flows in all MOSFETs in the circuit. Now assume that i_1 is increased above zero, but less than I_{SS}. This causes the current in both M1 and M2 to increase. As a result, the drain current in M3 decreases to keep $i_{D2} + i_{D3} = 2I_{SS}$. The decrease in M3 drain current is mirrored in M4, forcing the current i_1 out of the diff-amp. A similar argument can be made for increasing i_2; that is, it causes the output of the differential amplifier to sink a current equal to i_2. The sizes of M1–M4 can be ratioed to give the diff-amp a gain or to scale the input currents.

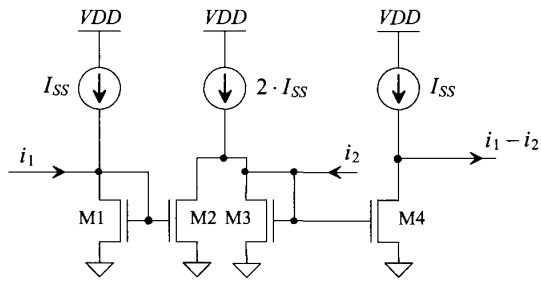


Figure 22.33 Current differential amplifier.

The input impedance of the current differential amplifier is simply the small-signal resistance of a diode-connected MOSFET, or

$$R_{in} = \frac{1}{g_m} \tag{22.54}$$

This configuration finds applications in both low-power and high-speed circuit design.

IPR2023-00698
Theta EX2019

22.4.2 Constant Transconductance Diff-Amp

A rail-to-rail differential amplifier made using n- and p-channel diff-amps is shown in Fig. 22.34. It is desirable, for good distortion and proper compensation, that the overall transconductance of the diff-amp remain constant independent of the region of operation, that is, operation with both n- and p-channel diff-amps on or only a single diff-amp on. A constant g_m is guaranteed over the input range if

$$g_m = g_{mn} + g_{mp} = \sqrt{2\beta_n I_n} + \sqrt{2\beta_p I_p} = \text{constant} \tag{22.55}$$

where g_{mn} and g_{mp} are the transconductances of the n- and p-channel diff-amps and g_m is the overall transconductance of the input stage. Since β_n and β_p are constant and can be made equal, Eq. (22.55) can be rewritten as (assuming both diff-amps are operating),

$$\sqrt{I_n} + \sqrt{I_p} = \text{constant} \tag{22.56}$$

This equation always holds if both differential amplifiers are on. The problem with nonconstant transconductance occurs if only one diff-amp is on. If, for example, the common-mode input (common input voltage on the + and – inputs of the differential amplifier) is large enough to shut the p-channel diff-amp off, then $I_p = 0$ and the transconductance of the overall input diff-amp changes.

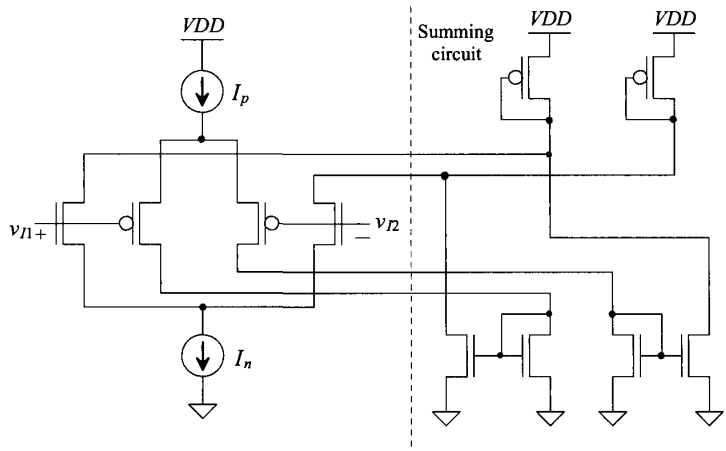


Figure 22.34 Rail-to-rail differential amplifier.

The solution to the problem of non-constant g_m begins by making

$$I_n = I_p = I_o \tag{22.57}$$

When both diff-amps are on (keeping in mind that we have already set the requirement that $\beta_n = \beta_p$), Eq. (22.56) reduces to

$$2\sqrt{I_o} = \text{constant} \tag{22.58}$$

IPR2023-00698
Theta EX2019

If we add a current of $3I_o$ to I_n (or to I_p) when the p-channel diff-pair (n-channel diff-pair) is off, the transconductance of the pair is constant, or

$$\overbrace{2\sqrt{I_o}}^{\text{both on}} = \overbrace{\sqrt{3I_o + I_n}}^{\text{n diff-amp on}} = \overbrace{\sqrt{3I_o + I_p}}^{\text{p diff-amp on}} \text{ if } I_o = I_p = I_n \qquad (22.59)$$

An example of a constant-g_m , rail-to-rail input stage is shown in Fig. 22.35. The summing circuit of Fig. 22.34 is not shown in this figure. MOSFETs M1–M4 make up the p- and n-channel diff-pairs, while MP1 and MN1 source the constant current I_o when both diff-pairs are on. (MP1 and MN1 represent the constant current sources shown in Fig. 22.28.) When both diff-pairs are on, the current out of each current diff-amp is approximately 0. (The drain current of M6 and M7 is approximately 0.) If the common-mode input voltage becomes large enough to shut the p-channel diff-pair off, then MOSFETs MS1 and MS2 are off as well. This causes the current in M5 to become I_o. The current in M6 mirrors the current in M5. Since M6 is three times larger than M5, the current in M6 becomes $3I_o$.

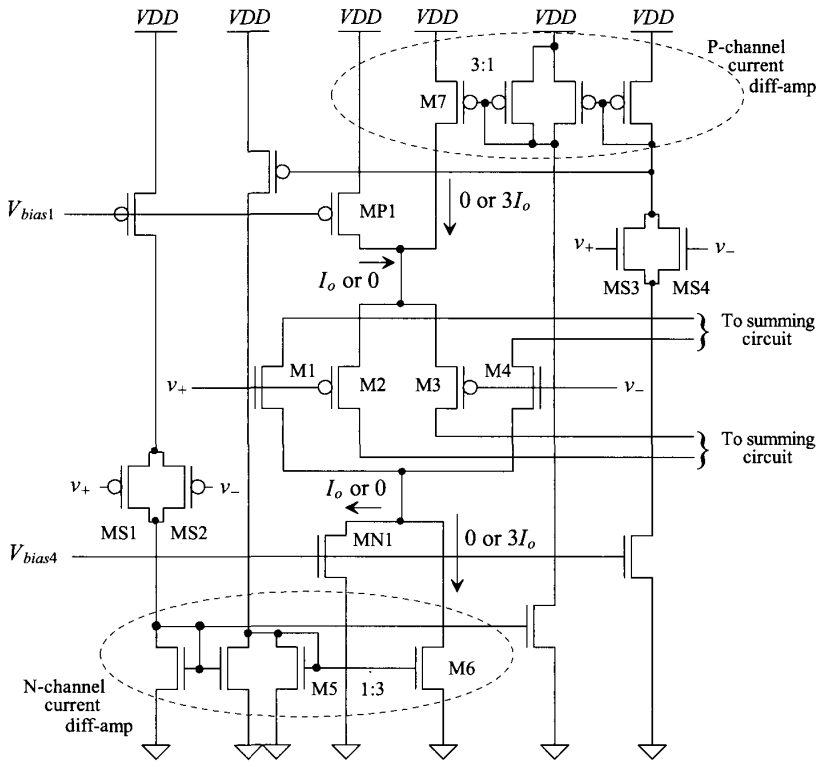


Figure 22.35 A wide-swing diff-amp with constant gm.

IPR2023-00698
Theta EX2019

Discussion

In general, the constant-g_m diff-amp is used in an op-amp to prevent overcompensation of the op-amp and to avoid distortion when the op-amp is used with large signals where the diff-amps, on the input of the op-amp, are turning on and off with variations in the input signal. Several practical problems exist with these uses of the constant-g_m diff-amp. Because the value of g_m may vary with shifts in the process, we may still overcompensate an op-amp design. Using the constant-g_m stage to avoid overcompensation is useful, but since we may overcompensate (or undercompensate) the op-amp with process variations, the added complexity and power draw may not justify the additional circuitry. Using the constant-g_m stage to avoid distortion is based on keeping the unity gain frequency of an op-amp, f_{un}, a constant independent of the common-mode input voltage. In practice, mismatches (threshold voltage and geometry) in the input diff-pair and changes in DC biasing conditions, resulting from the diff-amps turning off and on, can cause distortion. In some cases, this distortion can be worse than the distortion resulting from the nonconstant f_{un} (which we get with the nonconstant g_m diff-amp). The distortion resulting from mismatches can be modeled as an offset-voltage that is dependent on the input common-mode voltage. Since the DC currents sourced or sunk from a constant-g_m diff-amp are not constant, the DC operating point of the circuit summing the currents changes with the input signal's amplitude. The result is a change in the low-frequency gain and added distortion.

ADDITIONAL READING

[1] R. Jaeger and T. Blalock, *Microelectronic Circuit Design*, 3rd ed., McGraw-Hill Publishers, 2007. ISBN 978-0073309484.

[2] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, 2nd ed., Oxford University Press, 2002. ISBN 0-19-511644-5.

[3] A. L. Coban, P. E. Allen, and X. Shi, "Low-Voltage Analog IC Design in CMOS Technology," *IEEE Transactions on Circuits and Systems,* vol. 42, no. 11, November 1995.

[4] J. H. Botma, R. F. Wassenaar, and R. J. Wiegerink, "A Low-Voltage CMOS Op Amp with a Rail-to-Rail Constant-g_m Input Stage and a Class AB Rail-to-Rail Output Stage," *Proceedings of the 1993 IEEE ISCAS,* p. 1314.

[5] T. S. Fiez, H. C. Yang, J. J. Yang, C. Yu, and D. J. Allstot, "A Family of High-Swing CMOS Operational Amplifiers," *IEEE Journal of Solid State Circuits,* vol. 22, no. 6, pp. 1683–1687, December 1989.

[6] E. Seevinck and R. F. Wassenaar, "A Versatile CMOS Linear Transconductor/Square-Law Function Circuit," *IEEE Journal of Solid State Circuits,* vol. SC-22, no. 3, pp. 366–377, June 1987.

[7] M. Steyaert and W. Sansen, "A High-Dynamic-Range CMOS Op Amp with Low-Distortion Output Structure," *IEEE Journal of Solid State Circuits,* vol. SC-22, no. 6, pp. 1204–1207, December 1987.

[8] R. Castello and P. R. Gray, "A High-Performance Micropower Switched-Capacitor Filter," *IEEE Journal of Solid State Circuits,* vol. SC-20, no. 6, pp. 1122–1132, December, 1985.

IPR2023-00698
Theta EX2019

PROBLEMS

22.1 Determine the drain current of M1 as a function of the input voltage, $v_{I1} - v_{I2}$, for the diff-amp shown in Fig. 22.36. Neglect body effect. What does your derivation give for i_{D1} when v_{I1} is much larger than v_{I2}?

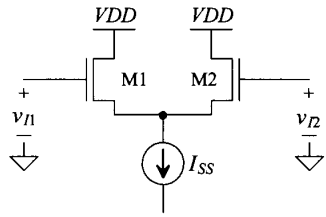


Figure 22.36 Diff-amp used in Problem 22.1

22.2 Repeat Ex. 22.1 if the widths of M1 and M2 are increased from 10 to 100. Determine the transconductance of the diff-amp. Write i_{d2} as a product of g_m (= g_{m1} = g_{m2}) and v_{I1} (with v_{I2} = AC ground), v_{I2} (with v_{I1} = AC ground), and $v_{I1} - v_{I2}$.

22.3 Determine the maximum and minimum common mode voltages for the PMOS version of the diff-amp seen in Fig. 22.4.

22.4 Determine the AC currents flowing in the circuit of Fig. 22.5 if the gate of M1 is grounded and the gate of M2 is an AC signal of 1 mV. Verify your answers with an AC SPICE simulation.

22.5 Determine the small-signal gain and the input common-mode-range (CMR) for the diff-amps shown in Fig. 22.37. Verify your answers with SPICE.

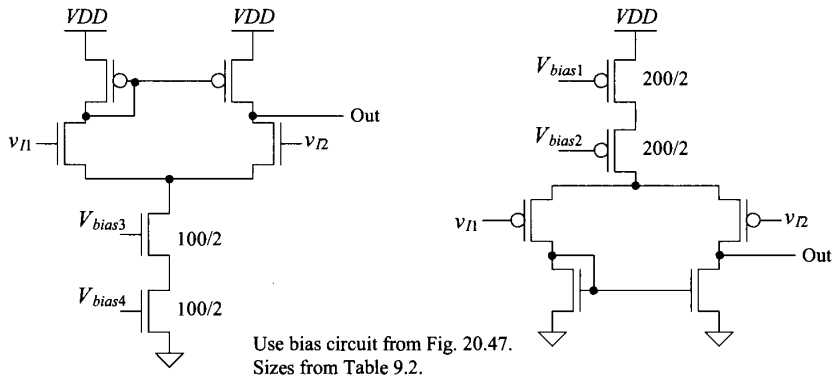


Figure 22.37 Diff-amps for Problem 22.5.

IPR2023-00698
Theta EX2019

22.6 Show that the capacitance on the sources of M1/M2 in Ex. 22.6 causes the *CMRR* to roll off quicker with increasing frequency. (Add a capacitance to the sources of M1/M2 in a simulation and show that *CMRR* decreases at a lower frequency.)

22.7 Estimate the slew-rate limitations in charging and discharging a 1 pF capacitor tied to the outputs of the diff-amps shown in Fig. 22.37. Verify with SPICE.

22.8 For the n-channel diff-pair shown in Fig. 22.38, show that the following relationships are valid if the body effect is included in the analysis of the transconductance.

$$i_{d1} = \frac{g_m}{2}\left[v_{i1}\left(2 - \frac{g_m}{g_m + g_{mb}}\right) - v_{i2} \cdot \frac{g_m}{g_m + g_{mb}}\right] - \frac{g_m \cdot g_{mb}}{g_m + g_{mb}} \cdot \frac{[v_{i1} + v_{i2}]}{2}$$

and

$$i_{d2} = \frac{g_m}{2}\left[v_{i2}\left(2 - \frac{g_m}{g_m + g_{mb}}\right) - v_{i1} \cdot \frac{g_m}{g_m + g_{mb}}\right] - \frac{g_m \cdot g_{mb}}{g_m + g_{mb}} \cdot \frac{[v_{i1} + v_{i2}]}{2}$$

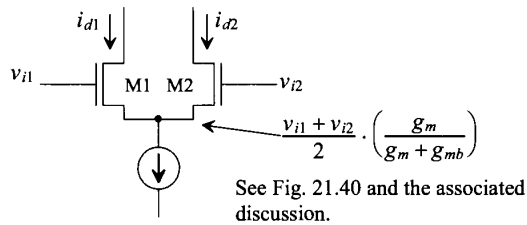


Figure 22.38 How body effect changes the AC behavior of the diff-amp.

22.9 The diff-amp configuration shown in Fig. 22.39 is useful in situations where a truly differential output signal is needed. Determine the following: (a) the transconductance of the diff-amp, (b) the AC small-signal drain currents of all MOSFETs in terms of the input voltages and g_{mn} (the transconductance of an n-channel MOSFET), and (c) the small-signal voltage gain, $(v_{O+} - v_{O-})/(v_{I+} - v_{I-})$. Verify your answers with SPICE.

22.10 Using the diff-amp topology in Fig. 22.24 and a current mirror load, show how quickly (or slowly) the diff-amp can drive a 1 pF load. Does this diff-amp exhibit slew-rate limitations? Verify your answers with SPICE.

22.11 Using the topology seen in Fig. 22.26 with the long-channel process (Table 9.1) without body effect (all MOSFET's bodies tied to their respective sources), show that the currents in M1–M4 match the biasing currents used in the source followers. Show how the currents become mismatched when the bodies of the NMOS are tied to ground and the bodies of PMOS are tied to *VDD*.

IPR2023-00698
Theta EX2019

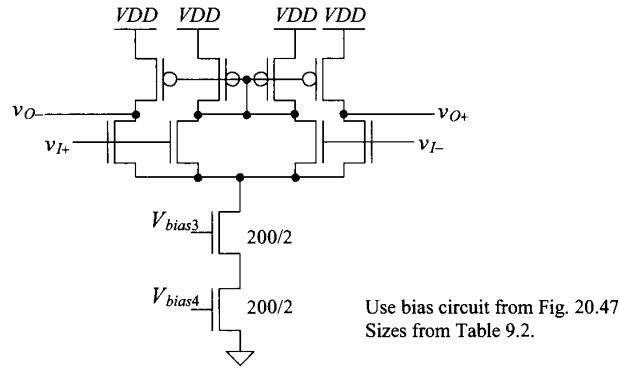


Figure 22.39 Fully differential diff-amp used in Problem 22.9.

22.12 Simulate the operation of the diff-amp seen in Fig. 22.31. Use a common-mode voltage of 2.5 V and an AC input voltage of 100 μV at 1 kHz. Compare the simulation results to your hand calculations.

22.13 Using SPICE with current sources for inputs, show the operation of the diff-amp in Fig. 22.33.

22.14 The circuit seen in Fig. 22.40 is an another example of a circuit that sums the currents from NMOS and PMOS diff-amps (so the input common-mode range can extend from beyond the power supply rails). Describe and simulate the operation of this circuit using the short-channel parameters in Table 9.2 and the bias circuit from Fig. 20.47.

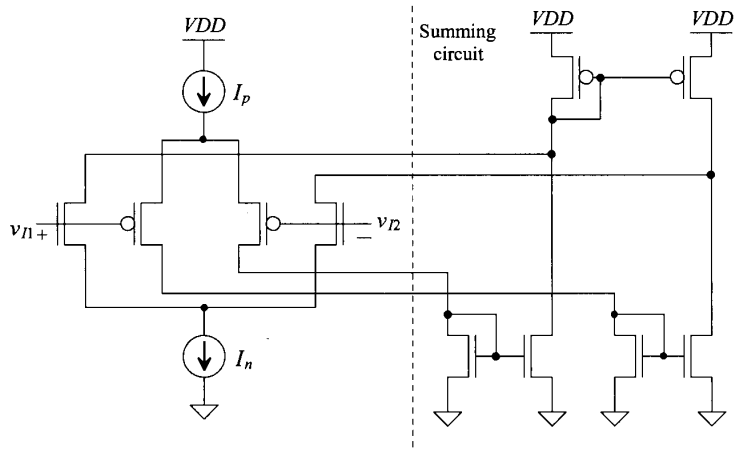


Figure 22.40 Summing circuit for Problem 22.14.

IPR2023-00698
Theta EX2019

22.15 Using symbolic analysis, determine the small-signal gain of the self-biased diff-amp seen in Fig. 22.41 (see also Fig. 18.17).

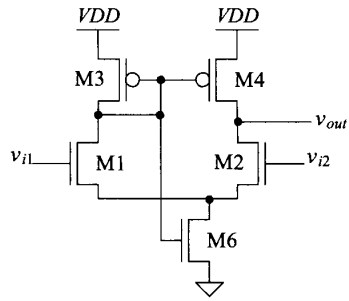


Figure 22.41 A self-biased NMOS diff-amp.

IPR2023-00698
Theta EX2019

Chapter

23

Voltage References

A general-use (ideal) voltage reference is a circuit used to generate a fixed voltage, V_{REF}, that is independent of the power supply voltage *VDD* (where $V_{REF} < VDD$), temperature, and process variations. In other words, the ideal reference voltage is independent of PVT. In some cases, we want to design a reference that varies with temperature. For example, if V_{REF} increases with temperature, Fig. 23.1a, we say that the reference voltage is *proportional to absolute temperature* or PTAT. If the reference voltage decreases with increasing temperature, Fig. 23.1b, the reference is said to be *complementary to absolute temperature* or CTAT. The PTAT and CTAT references can be used to design a voltage reference that changes very little with temperature called a *bandgap* reference. Unfortunately, the generation of PTAT and CTAT reference voltages requires using parasitic diodes. In a CMOS process, the electrical characteristics of the parasitic pn junctions are not monitored and controlled (like, say, the threshold voltage) during manufacturing. Therefore, if possible, the implementation of a reference voltage using MOSFET-resistor circuits is desirable.

This chapter is split into two sections. The first section covers the design of voltage references using MOSFETs and resistors, while the second section covers the design of voltage references using parasitic diodes. The first section also offers some overview material common to all voltage references.

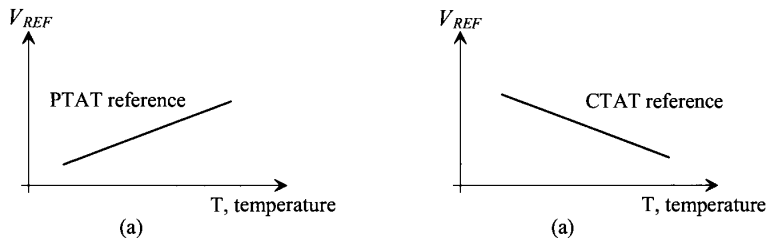


Figure 23.1 (a) PTAT and (b) CTAT voltage references.

IPR2023-00698
Theta EX2019

23.1 MOSFET-Resistor Voltage References

We can derive reference voltages from the power supplies using the resistor and the MOSFET, as seen in Fig. 23.2. The voltage divider formed with two resistors has the advantage of simplicity, temperature insensitivity (as was shown in Ch. 5), and process insensitivity; that is, changes in the sheet resistance have no effect on the voltage division. The main problem with this circuit is that in order to reduce the power dissipation (i.e., the current through the resistors), the resistors must be made large. Since large resistors require a large area on the die, this voltage divider may not be practical in many cases. One situation where we will use this simple voltage divider is in generating a voltage halfway between *VDD* and ground, *VDD*/2, (sometimes called the *common-mode* voltage of an analog circuit or system).

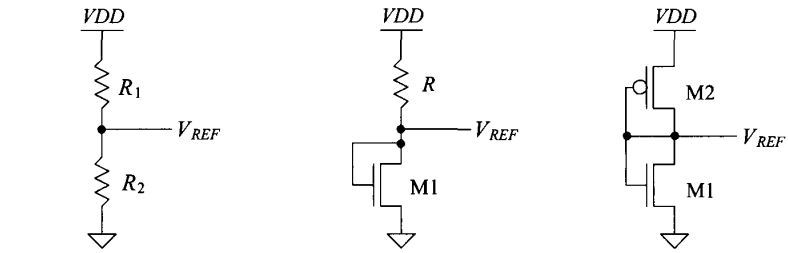


Figure 23.2 Voltage dividers implemented in CMOS.

The voltage divider formed between the resistor and the MOSFET can be recognized as the same circuit we used for a bias in the current mirror back in Ch. 20 (see Fig. 20.2 and Eqs. (20.26)–(20.31)). The final reference, a voltage divider between NMOS and PMOS devices, has the advantage that the layout can be small (see Fig. 20.13 and Eqs. (20.12)–(20.19)). In the following two subsections, we analyze the behavior of these last two voltage dividers.

23.1.1 The Resistor-MOSFET Divider

The reference voltage used in the resistor-MOSFET divider is equal to the V_{GS} of the MOSFET. We can write for this circuit

$$I_D = \frac{VDD - V_{REF}}{R} = \frac{\beta_1}{2}(V_{REF} - V_{THN})^2 \qquad (23.1)$$

or

$$V_{REF} = V_{THN} + \sqrt{\frac{2I_D}{\beta_1}} = V_{THN} + \sqrt{\frac{2(VDD - V_{REF})}{R \cdot \beta_1}} \qquad (23.2)$$

If V_{REF} is designed so that it is close to V_{THN}, then the reference voltage will be insensitive to changes in *VDD* and its temperature behavior will follow the threshold voltage (see

IPR2023-00698
Theta EX2019

Eqs. (9.43)–(9.48)). However, for the general case, the temperature coefficient of the resistor-MOSFET voltage divider is determined using

$$TCV_{REF} = \frac{1}{V_{REF}} \cdot \frac{\partial V_{REF}}{\partial T} \tag{23.3}$$

and if we assume $VDD >> V_{REF}$, then

$$TCV_{REF} = \frac{1}{V_{REF}}\left[V_{THN} \cdot TCV_{THN} - \frac{1}{2}\sqrt{\frac{2L_1}{W_1} \cdot \frac{VDD}{R \cdot KP(T)}} \cdot \left[\frac{1}{R}\frac{\partial R}{\partial T} - \frac{1.5}{T}\right]\right] \tag{23.4}$$

Example 23.1

Estimate the temperature performance of the resistor-MOSFET voltage references seen in Fig. 23.3. Assume that the temperature coefficient of the resistor is 2,000 *ppm/C*, the long-channel process is used where KP_n = 120 μA/V², and the nominal *VDD* is 5 V. Use simulations to verify the answers. Also show how the reference voltages change with *VDD*.

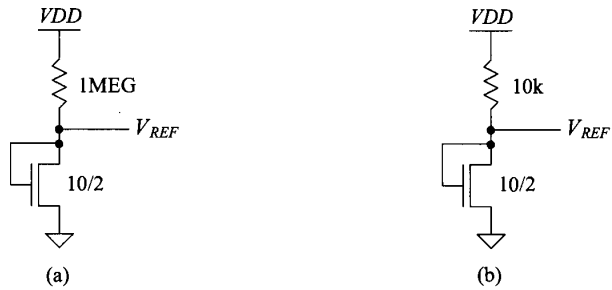


Figure 23.3 Resistor-MOSFET references used in Ex. 23.1

For the reference made using the 1MEG resistor, Fig. 23.3a, the current that flows in the circuit is

$$I = \frac{VDD - V_{REF}}{10^6} = \frac{KP_n}{2} \cdot \frac{W}{L}(V_{REF} - V_{THN})^2$$

which can be solved to determine I is around 4 μA and

$$V_{REF} = V_{GS} \approx 900\ mV$$

From Sec. 9.1.3 the rate the threshold voltage changes with temperature is

$$\frac{\partial V_{THN}}{\partial T} \approx -1\ mV/C^{\circ} \approx \frac{\partial V_{REF}}{\partial T}$$

The temperature coefficient of the reference voltage is

$$TCV_{REF} = \frac{1}{V_{REF}}\frac{\partial V_{REF}}{\partial T} = \frac{-0.001}{0.9} = -1,111\ ppm/C^{\circ} \tag{23.5}$$

and thus

$$V_{REF}(T) = V_{REF} \cdot (1 + TCV_{REF}(T - T_0)) = 0.9 \cdot (1 - 0.00111(T - 25))) \tag{23.6}$$

IPR2023-00698
Theta EX2019

where we assume that the threshold voltage was measured at 25 °C. The simulation results are seen in Fig. 23.4. Note that at higher temperatures the threshold voltage decreases and thus so does V_{REF}. Further note that a change in temperature of 25 °C corresponds to a change in the reference voltage of 31.25 mV (= 25 · 0.00125).

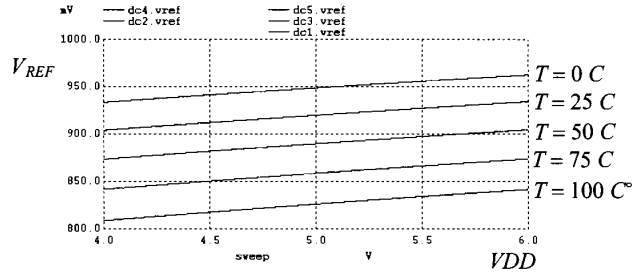


Figure 23.4 How the reference voltage changes with VDD and temperature for the reference in Fig. 23.3a.

To estimate the temperature behavior of the circuit in Fig. 23.3b, we must first determine the value of V_{REF}. Using Eq. (23.2), we can write

$$V_{REF} = V_{THN} + \sqrt{\frac{2(VDD - V_{REF})}{R \cdot \beta_1}} = 0.8 + \sqrt{\frac{2(5 - V_{REF})}{10k \cdot \frac{10}{2} \cdot 120\mu A/V^2}}$$

After a few iterations, we can determine $V_{REF} \approx 1.85\ V$. The temperature coefficient is calculated using Eq. (23.4)

$$TCV_{REF} = \frac{1}{1.85} \cdot \left[-1\ mV/C° - \frac{1}{2}\sqrt{\frac{2 \cdot 2}{10}} \cdot \frac{5}{10k \cdot 120\ \mu A/V^2} \cdot \left[0.002 - \frac{1.5}{300} \right] \right]$$

which evaluates to $TCV_{REF} = 500\ ppm/C$. The simulation results are seen in Fig. 23.5. Note how, when comparing Figs. 23.4 and 23.5, one reference circuit performs well with regard to temperature (23.3b), while the other reference circuit (23.3a) performs well with regard to *VDD* variations. ■

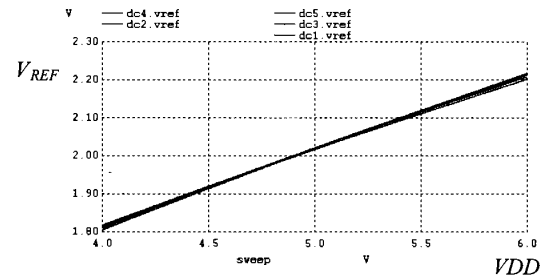


Figure 23.5 How the reference voltage changes with VDD and temperature for the reference in Fig. 23.3b.

IPR2023-00698
Theta EX2019

A modification to the basic resistor-MOSFET divider is shown in Fig. 23.6. The reference voltage in this circuit is given by

$$V_{REF} = V_{GS}\left(\frac{R_1}{R_2} + 1\right) \tag{23.7}$$

noting the ratio of the resistors (and so their temperature behavior or sheet resistance shift with process variations) doesn't affect the reference voltage. When V_{GS} is designed to be approximately V_{THN}, the resulting reference circuit is said to be a threshold-voltage multiplier reference circuit.

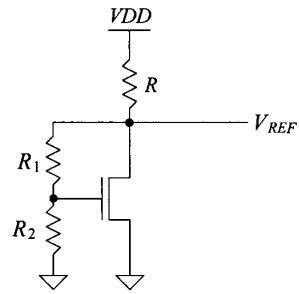


Figure 23.6 Modification of the resistor-MOSFET voltage divider.

23.1.2 The MOSFET-Only Voltage Divider

The MOSFET-only voltage divider shown in Fig. 23.2 generates a reference voltage that is equal to the voltage on the gates of the MOSFETs with respect to ground. Since $I_{D1} = I_{D2}$, we can write

$$\frac{\beta_1}{2}(V_{REF} - V_{THN})^2 = \frac{\beta_2}{2}(VDD - V_{REF} - V_{THP})^2 \tag{23.8}$$

or the reference voltage is given by

$$V_{REF} = \frac{VDD - V_{THP} + \sqrt{\frac{\beta_1}{\beta_2}} \cdot V_{THN}}{\sqrt{\frac{\beta_1}{\beta_2}} + 1} \tag{23.9}$$

or knowing the desired reference voltage and the power supply voltage

$$\frac{\beta_1}{\beta_2} = \left[\frac{VDD - V_{REF} - V_{THP}}{V_{REF} - V_{THN}}\right]^2 \tag{23.10}$$

The temperature dependence of the MOSFET-only voltage divider, assuming the temperature dependence of the ratio of the transconductance parameters, $\frac{\beta_1}{\beta_2}$, is negligible, is given by

$$TCV_{REF} = \frac{1}{V_{REF}} \cdot \frac{\partial V_{REF}}{\partial T} = \frac{1}{V_{REF}} \cdot \frac{1}{\sqrt{\frac{\beta_1}{\beta_2}} + 1} \cdot \left[\frac{\partial(-V_{THP})}{\partial T} + \sqrt{\frac{\beta_1}{\beta_2}} \frac{\partial V_{THN}}{\partial T}\right] \tag{23.11}$$

IPR2023-00698
Theta EX2019

From Ch. 9, we know for the long-channel devices

$$\frac{\partial V_{THN}}{\partial T} = -1\ mV/C^{\circ} \tag{23.12}$$

and

$$-\frac{\partial V_{THP}}{\partial T} = 1.4\ mV/C^{\circ} \tag{23.13}$$

To achieve $TCV_{REF} = 0$, requires

$$-\frac{\partial V_{THP}}{\partial T} = -\sqrt{\frac{\beta_1}{\beta_2}} \cdot \frac{\partial V_{THN}}{\partial T} \Rightarrow 1.4\ mV/C^{\circ} = \sqrt{\frac{\beta_1}{\beta_2}} \cdot 1\ mV/C^{\circ} \tag{23.14}$$

or

$$\sqrt{\frac{\beta_1}{\beta_2}} = 1.4 \tag{23.15}$$

Zero temperature coefficient, to a first order, can be met by satisfying this equation. However, this ratio is most often set by the desired V_{REF}.

23.1.3 Self-Biased Voltage References

We've already presented a voltage reference back in Ch. 20 using the beta-multiplier reference (BMR) (see Figs. 20.15, 20.19, and 20.22). Consider the BMR seen in Fig. 23.7. The added amplifier forces the drain/gates of M1 and M2 to the same potential. Because the gates and sources of M3/M4 are at the same potential, the same current is forced through each side of the reference. *This is an important common theme for all of the self-biased references discussed in this chapter.* We'll discuss this in greater detail in a moment. Before moving on to this topic, let's discuss why we connect the inverting and noninverting inputs to the added amplifier in the way seen in Fig. 23.7. (Why is M2 connected to the + amplifier input instead of the – input?) For the amplifier to be stable, we want the amount of signal fed back and subtracted from the input to be larger than the

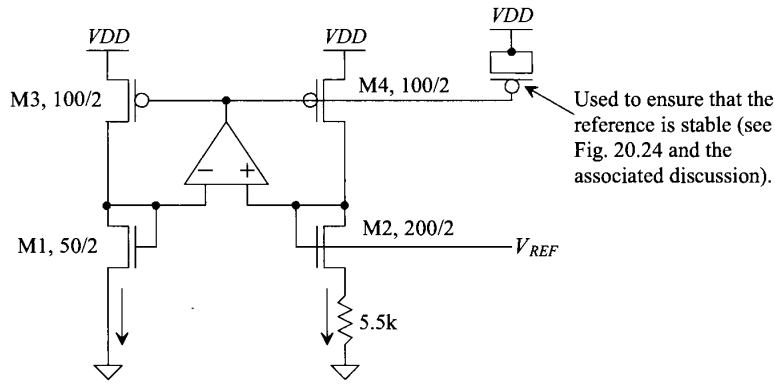


Figure 23.7 BMR from Figs. 20.19 and 20.22.

IPR2023-00698
Theta EX2019

amount of signal fed back and added to the input. Because M3 and M4 are *inverting* common-source amplifiers, we want the signal fed back to the + input of the amplifier to be larger than the signal fed back to the – input of the amplifier. Since the (AC) currents flowing in M3 and M4 are equal and the small-signal resistance of M2 + R is larger than the small-signal resistance of M1, we connect M2 to the + amplifier input.

Forcing the Same Current through Each Side of the Reference

The simplest method to force the same current through each side of a reference is to use a simple current mirror, Fig. 23.8a. As we saw in Fig. 20.15, the low-output resistance found in short-channel devices makes the resulting reference very sensitive to changes in *VDD*. By cascoding the current mirror, Fig. 23.8b, the currents can be made more equal and the sensitivity to *VDD* can be reduced. The problem with using a cascode current mirror is that the minimum allowable *VDD* increases. Consider the following.



Figure 23.8 Using a current mirror to force the same current through each side of a reference.

Example 23.2

Resimulate the BMR in Fig. 20.18 using a cascode current mirror for the PMOS devices as seen in Fig. 23.9.

Using the cascode current mirror, we expect the currents through each branch of the reference to be very nearly equal. As seen in the simulation results in Fig. 23.10, the currents *are* equal. However, they are not constant and independent of changes in *VDD*. Similarly, if we were to use V_{biasn} as a reference voltage (see Sec. 20.1.5), then V_{biasn} would show a large sensitivity to changes in *VDD*. The problem with cascoding only the PMOS devices is that while the currents through each branch are equal (but not constant), the variation with *VDD* is still present. The voltages at the drains of M1 and M2 change with *VDD*. To reduce this voltage variation, we might consider cascoding the NMOS devices as well (see Fig. 23.11). The PMOS devices are still used to force the same current through each side of the reference, while now the NMOS cascode stack is used to keep the voltages across M1 and M2 constant with changes in *VDD*. As the simulation results, Fig. 23.12, show, the currents do stabilize but at a voltage 20% higher than *VDD*. Clearly, cascoding devices won't be useful in a short-channel CMOS process (and so *we'll stick with using the added amplifier to both set the currents*

IPR2023-00698
Theta EX2019

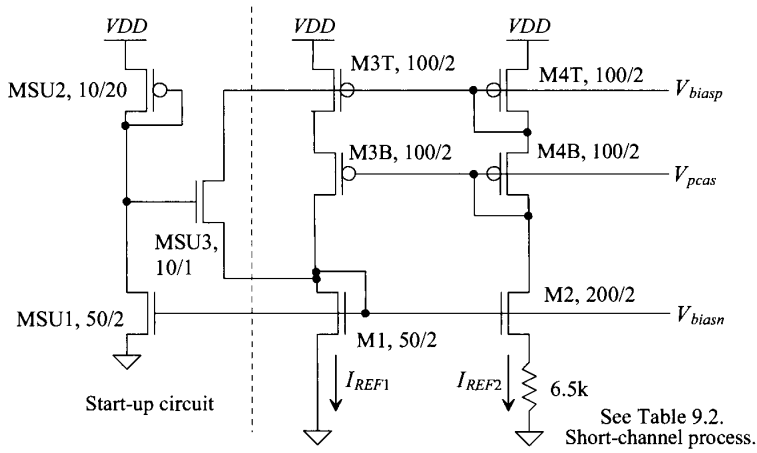


Figure 23.9 Cascoding the PMOS devices in the BMR circuit.

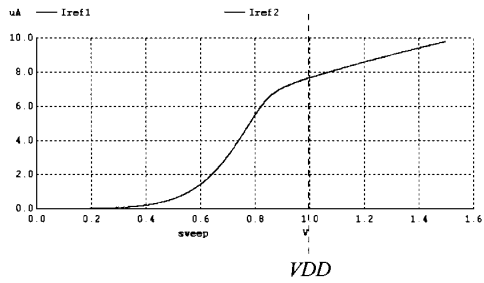


Figure 23.10 Simulating the behavior of the currents in the BMR of Fig. 23.9.

and hold the voltage across M1 and M2 constant in the remainder of the ***short-channel CMOS*** *designs presented in this chapter*). Note that cascoding devices can be useful when designing in long-channel CMOS processes. In these processes, the threshold voltage is a smaller percentage of the power supply voltage. For example, in our short-channel process, Table 9.2, this percentage is 28%. In the long-channel process, Table 9.1, it is 16%. Also note that using NMOS cascodes alone won't make the reference more tolerant to changes in *VDD*. The NMOS cascodes will keep the voltages across M1 and M2 constant (but not equal). Since the currents through each branch aren't equal (because the PMOS devices aren't cascoded), we still get significant sensitivity to *VDD*.

Finally, we need to, again (see Fig. 20.15 and the associated discussion), mention the importance of the start-up circuit. The start-up circuit is often overlooked and is often a cause of problems in practical designs. We include a start-up circuit in every self-biased reference to avoid the situation where zero

IPR2023-00698
Theta EX2019

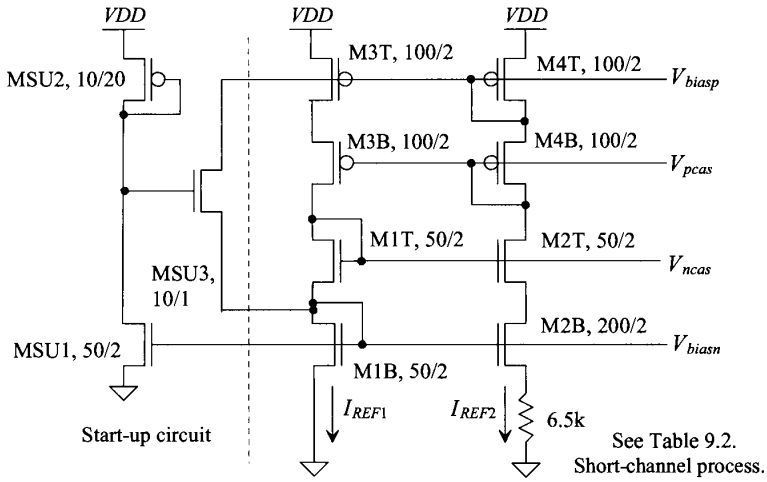


Figure 23.11 Cascoding both NMOS and PMOS devices in the BMR circuit.

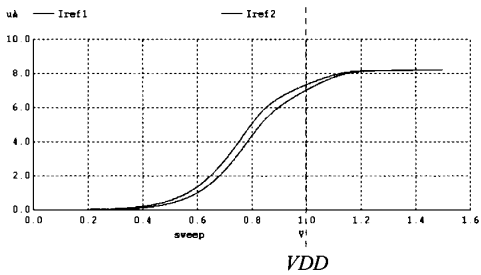


Figure 23.12 Simulating the behavior of the currents in the BMR of Fig. 23.11.

current flows in the reference. It's important to ensure that the start-up circuit doesn't affect normal operation or draw too much current from *VDD*. For the start-up circuit seen in Figs. 23.9 and 23.11, we use the V_{biasn} to set the current drawn by the bias circuit. In normal operation, MSU3 should be off (it should have a negative V_{GS}). ■

Example 23.3

Using the topology seen in Fig. 20.22, design a voltage reference with, ideally, zero temperature coefficient. Simulate the design to determine the reference's sensitivity to changes in temperature and *VDD*. Comment on the repeatability of the V_{REF} (with process variations from one process run to the next) and methods (and concerns) with trimming V_{REF}. Assume $TCR = 0.002$ (2,000 ppm/C).

IPR2023-00698
Theta EX2019

We begin by writing the long-channel equation, Eq. (20.40), for the resistance needed for zero temperature coefficient (ZTC)

$$R = \frac{2}{\frac{\partial V_{THN}}{\partial T} \cdot KP_n \cdot \frac{W}{L}} \left(1 - \frac{1}{\sqrt{K}}\right) \cdot \left(\frac{1}{R}\frac{\partial R}{\partial T} + \frac{1}{KP_n} \cdot \frac{\partial KP_n}{\partial T}\right) \quad (23.16)$$

While this equation won't directly apply for the design in a short-channel process, we can use it, with simulations, to help point us to the factors that influence the temperature behavior of the reference. Substituting in the appropriate numbers with $K = 4$, we get

$$R = \frac{1}{(-0.0006) \cdot KP_n \cdot \frac{W}{L}} \cdot \left((0.002) - \frac{1.5}{300}\right) \quad (23.17)$$

or

$$R = \frac{5}{KP_n \cdot \frac{W}{L}} \quad (23.18)$$

To move the reference towards a ZTC, let's set the resistor to nominally 5.5k (the same value we used before, see Fig. 20.22 and the associated discussion) and adjust the widths and lengths of the NMOS devices until the simulations show good temperature behavior (keeping in mind that we must keep the W/L of M2 K times the W/L of M1). The resulting circuit is seen in Fig. 23.13. Simulation results are seen in Fig. 23.14. (Note that the currents and gate-source voltages have nothing to do with the values listed in Table 9.2.) The reference voltage is nominally 500 mV (= *VDD*/2). The reference, according to simulations, moves 40 mV over a 100 °C temperature change (–400 μV/C or a TC of –800 ppm/C, not that great when compared to the references described later). Further, the reference is fairly insensitive to variations in *VDD* once *VDD* gets above 600 mV.

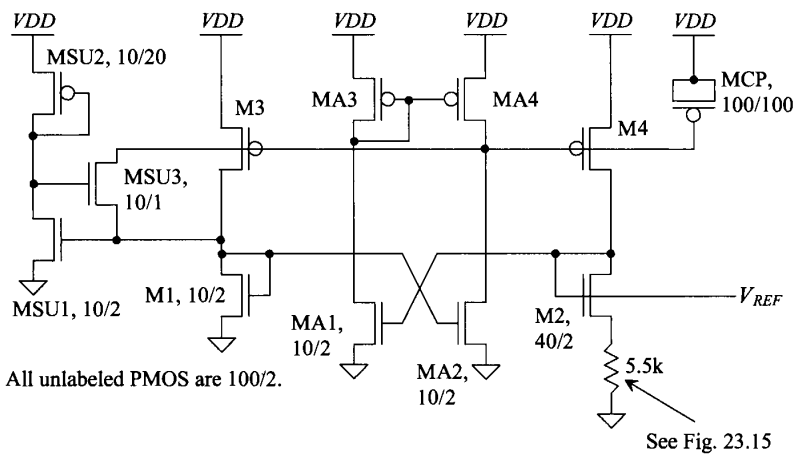


Figure 23.13 Voltage reference using the beta-multiplier.

IPR2023-00698
Theta EX2019

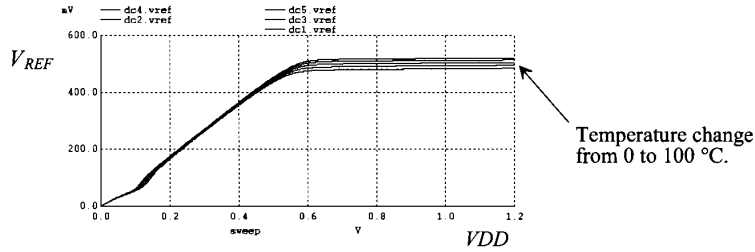


Figure 23.14 Simulating the operation of the reference in Fig. 23.13.

A couple of practical notes are needed at this point. While we've used simulations to zero in on a design with decent temperature behavior and a reference voltage at *VDD*/2, actually fabricating this reference would result in a V_{REF} different from the one seen in Fig. 23.14 (in most fabrication runs). The MOSFET characteristics and the sheet resistance vary with each process run.

To adjust the reference voltage to *VDD*/2, the resistor must be trimmed using fuses, Fig. 23.15. It can be shown that the temperature behavior doesn't vary significantly with small changes in V_{REF} when trimming. As seen in Eq. (20.38), lowering the resistor value causes V_{REF} to increase, while increasing the resistor value causes V_{REF} to decrease. The fuses in Fig. 23.15 short across the resistor until they are blown. To trim the resistor, we start blowing the resistors (electrically or with a laser). With each blown fuse we add (nominally) 200 Ω in series with the nominally 4k resistor. If the processes' sheet resistance increases by 20% (so that the resistors are now 4.8k and 240 Ω), then we only need to blow three fuses (neglecting the change in the MOSFET characteristics) to trim the resistor to 5.5k. If the sheet resistance decreases by 20%, then the resistors are 3.2k and 160 Ω. Fifteen fuses would need to be blown to trim the resistor to 5.5k.

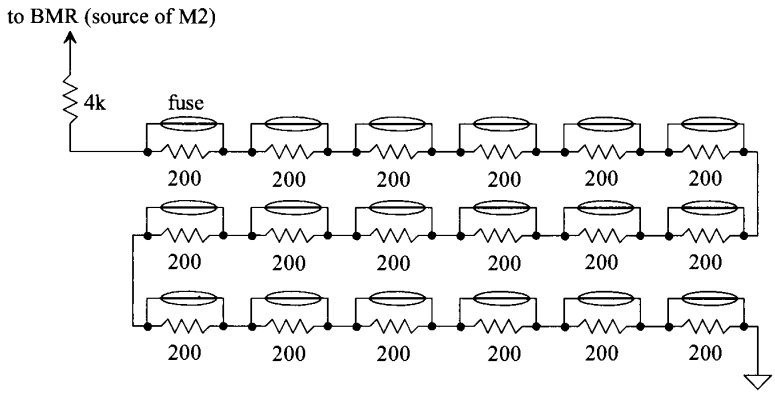


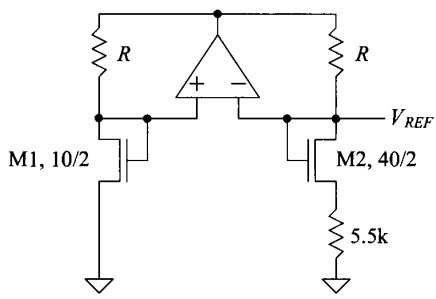
Figure 23.15 Trimming a resistor with fuses.

IPR2023-00698
Theta EX2019

Finally, it's important to remember that stability is of great importance when using feedback. As discussed in Sec. 20.1.4, the feedback loop is made stable by adding a capacitance on the output of the added amplifier. The bigger this capacitance, the more stable the reference. We stabilize the reference in Fig. 23.13 with the addition of MCP. ■

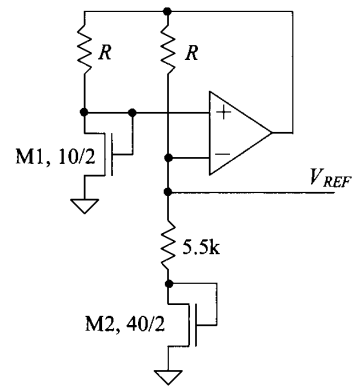
An Alternate Topology

Instead of using PMOS devices in the top of each branch of the reference, we might try holding the voltage across resistors constant to force the current through each branch of the reference to be the same. Figure 23.16 shows the idea drawn two different ways. In (a) the amplifier tries to hold both of its inputs at the same potential. This causes the voltage across each R to be the same and thus each branch of the reference has the same current. Note now we've tied M2 and the resistor (a resistance larger than M1 alone) to the – amplifier input. This is because the inverting PMOS common-source amplifiers (M3 and M4) are no longer in the feedback signal path; we need to ensure that the largest signal is fed back to the inverting amplifier terminal. Note also, in (b), that the series order of the resistor and MOSFET is not important. In fact, having both sources connected to ground eliminates the body-effect mismatch and may result in better temperature performance. We won't use these topologies in any of the MOSFET-resistor topologies in this chapter because the amplifier has to drive a resistive load (which, as we saw in Chs. 22 and 23, kills the amplifier's gain). A two-stage op-amp, discussed in the next chapter, can be used. However, then we have to be concerned with compensating the op-amp. Finally, note that a start-up circuit is still needed with these topologies. The start-up circuit would leak current into the node connected to the + amplifier input to start the reference circuit up (and then shut off after the reference turns on). Connecting the start-up circuit to the – amplifier input would move the amplifier output in the wrong direction and the reference would never start up (noting in Figs. 23.7 and 23.13 that we do connect the start-up circuit to the – input because M3 and M4 are inverting amplifiers).



(a) Holding the voltage across resistors to force the same current through each branch of the reference.

Note start-up circuit not shown (see text).



(b) Redrawing (a) and switching the places of the resistor and M2.

Figure 23.16 Other references useful in MOSFET-resistor circuits.

IPR2023-00698
Theta EX2019

23.2 Parasitic Diode-Based References

By using a junction diode, the variability encountered in the references of the last section that used a MOSFET's threshold voltage can be reduced. One practical implementation of a parasitic diode available in CMOS is seen in Fig. 23.17. A diode is formed between the p+ implant and the n-well. However, in a practical circuit the parasitic vertical PNP bipolar device formed by the p+, n-well, and p-substrate causes current to be injected into the substrate. Good guard rings are used to surround the p+ in both the n-well and the p-substrate to collect this current. This ensures that the injected carriers are collected and not causing substrate current to flow in other portions of the chip. Because the substrate is connected to ground, the diode's cathode (K) must be tied to ground. We might try to use an n+ implant directly in the substrate for diode formation. However, assuming the substrate (which forms the anode of the resulting diode) is grounded, we would have to apply a negative voltage to the n+ (K) to forward-bias the diode. One other practical parasitic diode available in CMOS is made using the lateral PNP device seen in Fig. 23.18. In this device, the p+ forms the emitter/collector of the resulting device (the n-well is the base). However, the vertical PNP (and the resulting substrate current) is still present. A significant portion of the lateral PNP's emitter current will flow into the substrate. In the following example, we'll assume vertical PNP (diodes) are used.

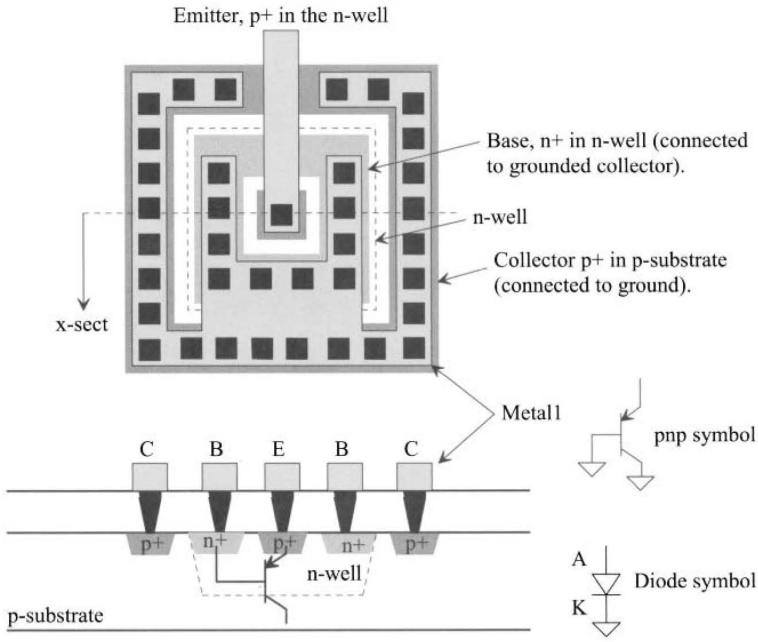


Figure 23.17 A vertical parasitic pnp bipolar junction transistor used as a diode.

IPR2023-00698
Theta EX2019

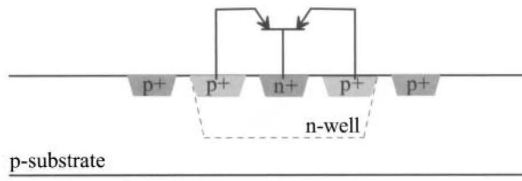


Figure 23.18 A lateral parasitic bipolar junction transistor.

Diode Behavior

It's important to realize that the parasitic diode's behavior must be thoroughly characterized in a particular CMOS process to determine its voltage-current (and temperature) characteristics. For the sake of illustrating design procedures, considerations, and examples in this chapter, we can develop a general diode model. The current through a forward-biased diode is given by

$$I_D = I_S \cdot e^{V_D/n \cdot V_T} \tag{23.19}$$

where V_D is the voltage across the diode, n is the emission coefficient (used to shape the diode's current-voltage curve to fit experimental data), V_T is the thermal voltage (kT/q or 26 mV at room temperature), and I_S is the diode's scale current. Here, we'll use a value of 10^{-18} A for the scale current and an n of 1. The SPICE statement corresponding to the schematic seen in Fig. 23.19 is

```
D1    1    0    PNPDIODE
D2    2    0    PNPDIODE    4
```

where the 4 indicates 4 diodes are connected in parallel. We may use, in schematics drawn in later in the chapter, the term K to indicate K diodes are connected in parallel. The model statement for the diode, using the parameters seen above, is

```
.MODEL    PNPDIODE    D    is=1e-18    n=1
```

Figure 23.20 shows how the voltage across the diode changes if the current through the diode is held constant and the temperature is changed. Using our model (n =1 and I_S =

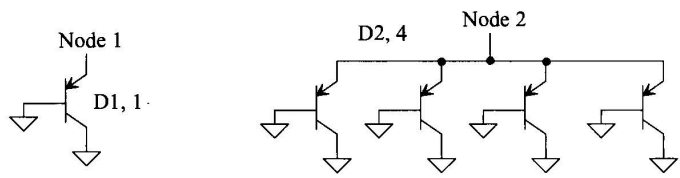


Figure 23.19 A schematic of diodes.

IPR2023-00698
Theta EX2019

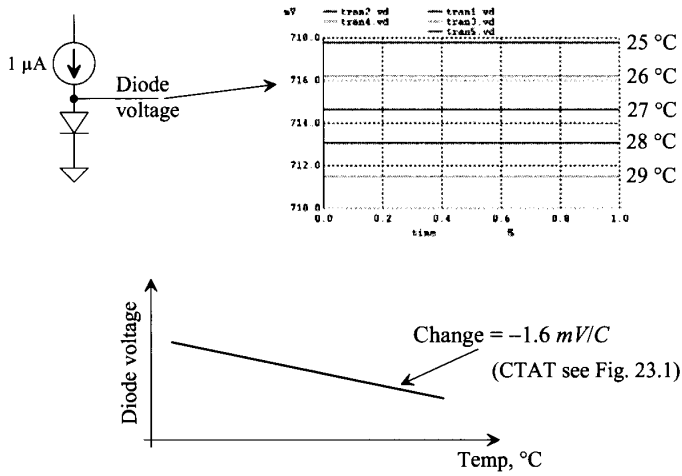


Figure 23.20 Change in diode voltage with temperature.

10^{-18}), we can estimate the change in the diode's voltage with temperature, from the simulation data, as

$$\frac{dV_D}{dT} = -1.6\ mV/C \tag{23.20}$$

noting that this number is very dependent on the DC current through the diode. (Also note this voltage change is complementary to absolute temperature, CTAT.) Increasing the bias current in Fig. 23.20 to 10 μA, causes the diode's voltage drop to increase (of course) and its voltage change with temperature to become −1.4 *mV/C*.

The Bandgap Energy of Silicon

The silicon bandgap energy as a function of temperature is given by

$$E_g(T) = 1.16 - (702 \times 10^{-6}) \cdot \frac{T^2}{T + 1108}\ (\text{eV}) \tag{23.21}$$

At room temperature, the bandgap of silicon is approximately 1.1 eV (where 1 electron-volt, eV, is 1.6×10^{-19} J). Looking at Eq. (23.21), notice that the bandgap energy decreases with increasing temperature. If we put a constant current through a diode and increase the temperature, the barrier height between the n and p sides of the diode decreases and thus so does the diode's voltage drop (and that is why the diode's voltage change with temperature in Fig. 23.20 is negative).

In the next two sections we'll talk about *BandGap References* (BGR). BGR's combine the CTAT behavior of the bandgap of silicon (the diode's forward voltage drop) with the PTAT behavior of the thermal voltage (the thermal voltage, kT/q , increases with increasing temperature see Fig. 23.1) to form a voltage reference (a BGR) that doesn't vary (much) with temperature.

IPR2023-00698
Theta EX2019

Lower Voltage Reference Design

While, at the time of this writing, power supply voltages are well above the forward turn-on voltage of the pn junction diode (parasitic PNP device), it is clear that the near future will bring *VDD* voltages below 0.7 V. To implement a reference in these lower supply voltage processes, consider the Schottky diode layout of Fig. 23.21. In this layout, we connect the metal (the anode) directly to the n-well (that is, without the p+) to form the Schottky diode. To connect the metal directly to the substrate, we need an opening in the oxide. To form this opening, we use the "active" layer without the select layer (without the implant). It's important that no select layer surrounds this Schottky contact.

The turn-on voltage of a Schottky contact is considerably lower than a standard pn junction. In Fig. 23.21 we draw the turn on voltage as 300 mV. In real silicon this voltage will depend highly on the doping of the n-well (which, as we saw in Ch. 6, scales up as process dimensions shrink). The design discussions and procedures developed in the next two sections can be applied to Schottky diode-based references. Again, though, thorough characterization of the parasitic diode is required prior to designing the reference. One concern, among others, is the repeatability of the diode's current-voltage characteristics from one process run to the next. We don't discuss Schottky diode-based references any further in this chapter, but, rather, leave it to the refereed literature.

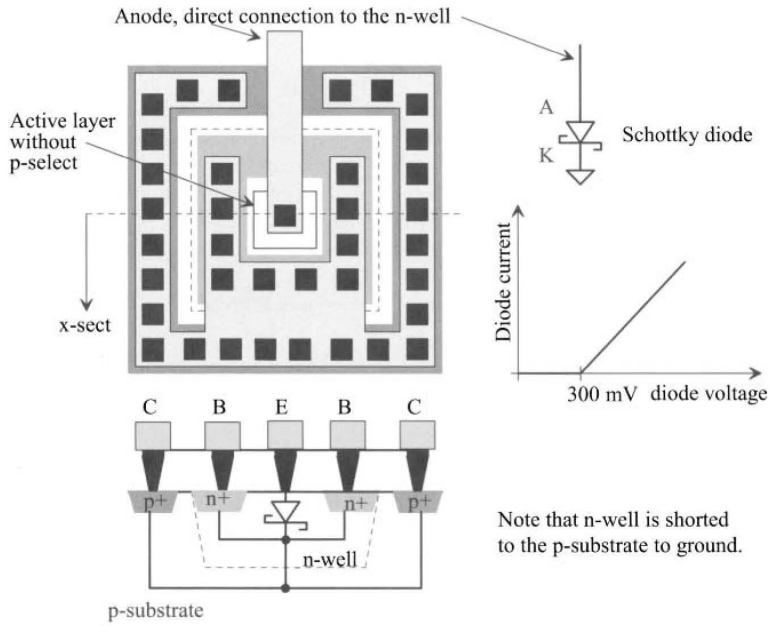


Figure 23.21 Connecting metal directly to n-well to make a Schottky diode.

IPR2023-00698
Theta EX2019

23.2.1 Long-Channel BGR Design

In this section we present some long-channel designs for bandgap references (BGR) using the cascode structure seen in Fig. 23.11. We will not show the (**required**) start-up circuit in the following schematics.

Diode-Referenced Self-Biasing (CTAT)

A diode-referenced, self-biasing circuit is seen in Fig. 23.22. The cascode structures force the same current through each branch of the reference. The voltage across the resistor now equals the forward voltage drop of the diode. The current in the circuit (the current flowing in each branch of the reference) is then

$$I_{REF} = \frac{V_D}{R} = I_S e^{V_D/nV_T} \tag{23.22}$$

Solving for a resistor to set the reference current level gives

$$R = \frac{nV_T}{I} \ln \frac{I}{I_S} \tag{23.23}$$

To determine the reference current's variation with temperature, we can write

$$\frac{\partial I_{REF}}{\partial T} = \frac{\partial}{\partial T}\left(\frac{V_D}{R}\right) = \frac{1}{R} \cdot \overbrace{\frac{\partial V_D}{\partial T}}^{\text{change in } V_D \text{ with } T} - \overbrace{\frac{V_D}{R^2}\frac{\partial R}{\partial T}}^{\text{TC of the resistor}} \tag{23.24}$$

Remembering that the overall current change with temperature is written as

$$I_{REF}(T) = I_{REF}(T_0) \cdot (1 + TCI_{REF}(T - T_0)) \tag{23.25}$$

Start-up circuit not shown.
NMOS are 10/2.
PMOS are 30/2.
Scale factor is 1 micron.

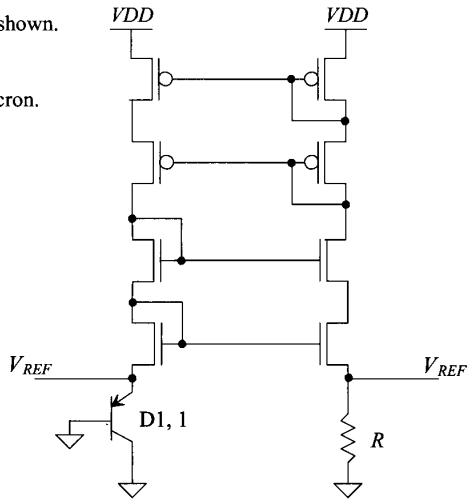


Figure 23.22 Diode-referenced, self-biasing circuit.

IPR2023-00698
Theta EX2019

where

$$TCI_{REF} = \frac{1}{I_{REF}} \cdot \frac{\partial I_{REF}}{\partial T} \tag{23.26}$$

we can, with the help of Eq. (23.24), write

$$\frac{1}{I_{REF}} \frac{\partial I_{REF}}{\partial T} = \frac{1}{V_D} \frac{\partial V_D}{\partial T} - \frac{1}{R} \frac{\partial R}{\partial T} \tag{23.27}$$

The behavior of the diode-referenced, self-biased circuit shows a large negative coefficient. If the TC of the resistor is 0.002 and the change in the diode voltage with T (Eq. (23.20)) is –0.0016 (or if the diode voltage, V_D, is nominally 0.7, the first term in Eq. (23.27) is –0.0023), then the overall TC of the current is, roughly, –0.004 or –4,000 ppm/C. If we were to use the voltage across the diode (or resistor) as a voltage reference, the fact that the current through the diode is decreasing with increasing temperature (causing the diode's voltage to drop) and the bandgap of silicon is decreasing with increasing temperature, causes V_D to drop (relatively quickly) with increasing temperature (CTAT). Consider the following example.

Example 23.4

Simulate the operation of the reference in Fig. 23.22 where the voltage across the diode is used as a reference voltage. Use a nominal reference current of 1 μA.

Using Eq. (23.23), we can set R to 700k. The simulation results are seen in Fig. 23.23. The reference voltage change is essentially set by the decrease in the diode voltage with increasing temperature, that is, –1.6 mV/C. At the risk of stating the obvious, this isn't a very good voltage reference. ■

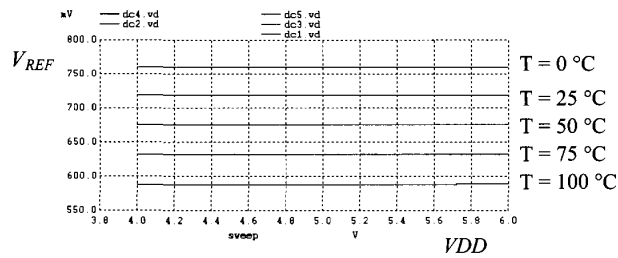


Figure 23.23 Temperature behavior of the diode-referenced circuit of Fig. 23.22. See Ex. 23.4.

Thermal Voltage-Referenced Self-Biasing (PTAT)

A thermal voltage, V_T, referenced, self-biasing circuit is seen in Fig. 23.24. In this configuration the voltage across D1 must equal the voltage across D2 and the resistor

$$V_{D1} = V_{D2} + I_{D2}R \tag{23.28}$$

Notice that D2 must be larger than D1 if the current flowing in the reference is to be nonzero. The larger D2 will drop a smaller voltage than D1 for the same current through each branch. As indicated in Eq. (23.28), the difference in the diode voltages is dropped

IPR2023-00698
Theta EX2019

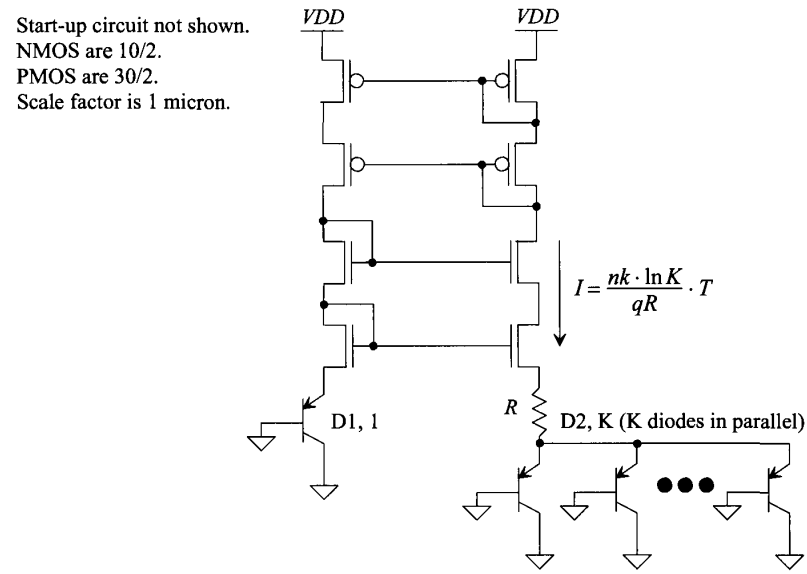


Figure 23.24 Thermal voltage-referenced, self-biasing circuit.

across R. Alternatively, we could size the diodes the same and increase the current flowing in the left branch of the reference by making the MOSFETs K times wider. The current in the left branch would be K times more than the current in the right branch of the reference.

We know

$$I_{D1} = I_S e^{V_{D1}/nV_T} \rightarrow V_{D1} = nV_T \cdot \ln \frac{I_{D1}}{I_S} \tag{23.29}$$

and

$$I_{D2} = K \cdot I_S e^{V_{D2}/nV_T} \rightarrow V_{D2} = nV_T \cdot \ln \frac{I_{D2}}{K \cdot I_S} \tag{23.30}$$

Knowing, because of the cascode structures, $I_{D1} = I_{D2} = I$, we can write

$$R = \frac{nV_T \cdot \ln K}{I} \text{ or } I = \frac{nk \cdot \ln K}{qR} \cdot T \tag{23.31}$$

Notice that the current is proportional to absolute temperature (PTAT). To get a voltage reference that is also PTAT, consider the reference seen in Fig. 23.25. Here the reference voltage is given by

$$V_{REF} = I \cdot L \cdot R = \frac{nk \cdot L \cdot \ln K}{q} \cdot T \tag{23.32}$$

Notice how the temperature behavior of the resistor falls out of the equation.

IPR2023-00698
Theta EX2019

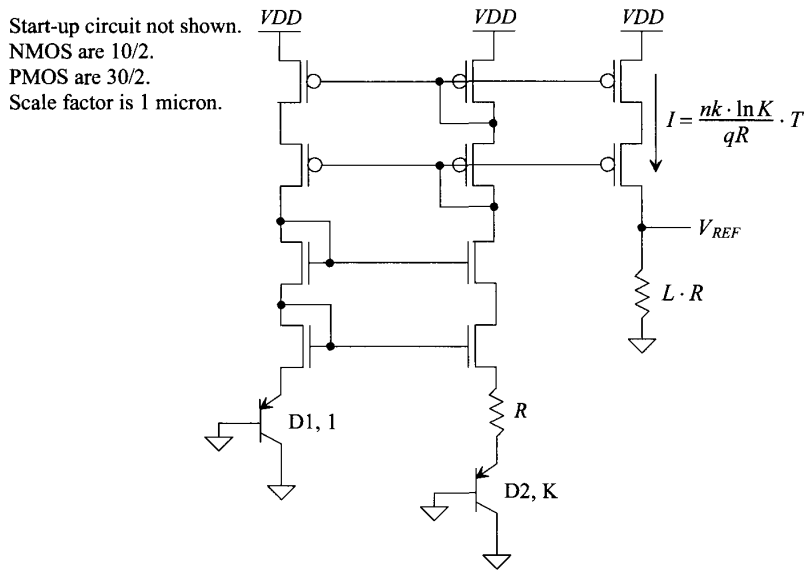


Figure 23.25 PTAT voltage reference based on the thermal voltage-referenced, self-biased circuit.

Example 23.5

Design a nominally 2.500 V voltage reference (at room temperature of 27 °C or 300 °K) PTAT voltage reference. Use the topology seen in Fig. 23.25 with a bias current of 1 μA. Simulate the operation of the design.

Let's set K (the number of diodes connected in parallel) to 8. We do this since it is easy to remember ln 8 = 2. Solving for R using Eq. (23.31) gives

$$R = \frac{1 \cdot 0.026 \cdot 2}{10^{-6}} = 52k$$

The voltage drop across the resistor is $1\ \mu A \cdot 52k = 52\ mV$. Using Eq. (23.32), we can solve for L as 48 (so $L \cdot R = 2.5\ M\Omega$). The simulation results are seen in Fig. 23.26. At room temperature, the reference voltage is roughly 2.5 V. Note how, as temperature increases, the reference voltage (PTAT) increases. Using Eq. (23.32), we can estimate the reference voltage's change with temperature as

$$\frac{\partial V_{REF}}{\partial T} = \frac{nk \cdot L \cdot \ln K}{q} \tag{23.33}$$

Using the numbers from this example with $k = 1.38 \times 10^{-23} J/K$ and $q = 1.6 \times 10^{-19} C$ gives

$$\frac{\partial V_{REF}}{\partial T} = \frac{1 \cdot 1.38 \cdot 10^{-23} \cdot 48 \cdot 2}{1.6 \times 10^{-19}} = 8.26\ mV/C$$

For every 25 °C increase in temperature, we expect V_{REF} to go up by 206 mV. ■

IPR2023-00698
Theta EX2019

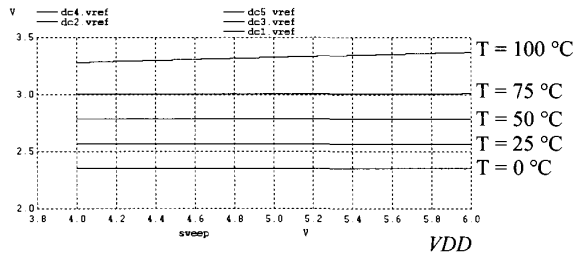


Figure 23.26 A PTAT voltage reference. See Ex. 23.5.

Bandgap Reference Design

The bandgap reference (BGR) is formed using CTAT and PTAT references. When designed properly, the TC of the BGR can be very small. Figure 23.27 shows a schematic of a BGR. The PTAT current generated from the circuit in Fig. 23.24 is driven into a resistive load (to give a PTAT voltage drop, see Eq. (23.32)) and a diode (to give a CTAT voltage drop, see Eq. (23.20)). The reference voltage is then the sum of the PTAT and CTAT voltage drops or

$$V_{REF} = V_{D3} + I \cdot L \cdot R = V_{D3} + L \cdot n \cdot \ln K \cdot V_T \qquad (23.34)$$

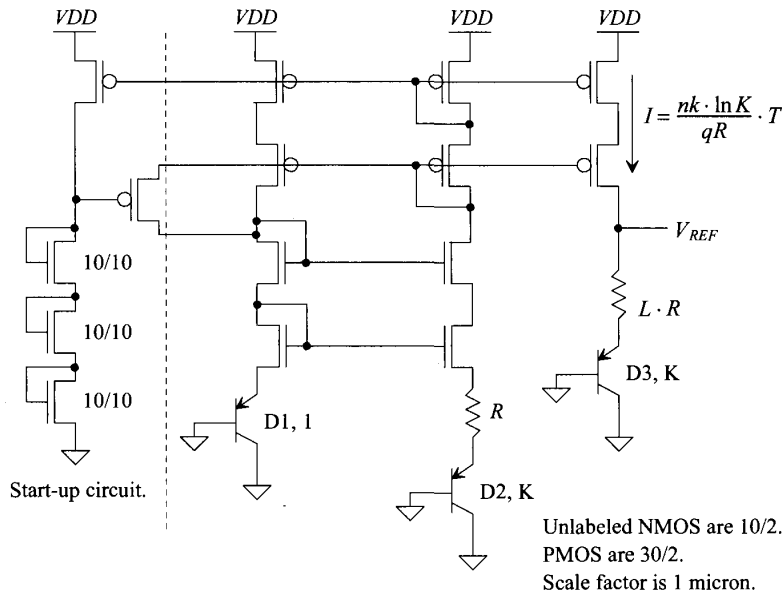


Figure 23.27 A bandgap reference circuit.

IPR2023-00698
Theta EX2019

The change in the reference voltage with temperature is

$$\frac{\partial V_{REF}}{\partial T} = \overbrace{\frac{\partial V_{D3}}{\partial T}}^{-1.6\ mV/C} + L \cdot n \cdot \ln K \cdot \overbrace{\frac{\partial V_T}{\partial T}}^{0.085\ mV/C} \tag{23.35}$$

where the change in the thermal voltage with temperature was taken from Eq. (9.44). To determine the value of L that causes the change in V_{REF} with temperature to go to zero, we set Eq. (23.35) to zero and solve

$$L = \frac{1.6}{n \cdot \ln K \cdot 0.085} \tag{23.36}$$

If $K = 8$ and $n = 1$, then $L = 9.41$. Using Eq. (23.34) with these values, gives a V_{REF} of 1.2V. Figure 23.28 shows the simulation results using these values and the schematic seen in Fig. 23.27. The start-up circuit causes the reference to turn on at a *VDD* of approximately 3V. The change in the reference voltage with temperature is roughly 1.5 mV per 75 °C or 20 μV/C. The TC of the reference, Eq. (23.3), is then 16.7 ppm/C. At higher temperatures the variation in V_{REF} (with *VDD*) is slightly poorer likely due to not keeping the currents in each branch equal. Note that if a precise voltage is needed, trimming the $L \cdot R$ resistor (see Fig. 23.15 and the associated discussion) is required.

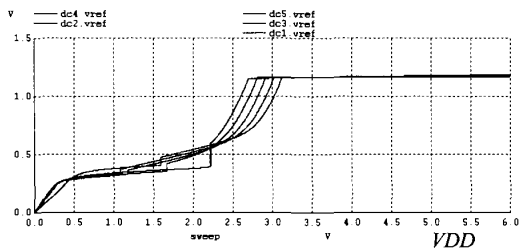


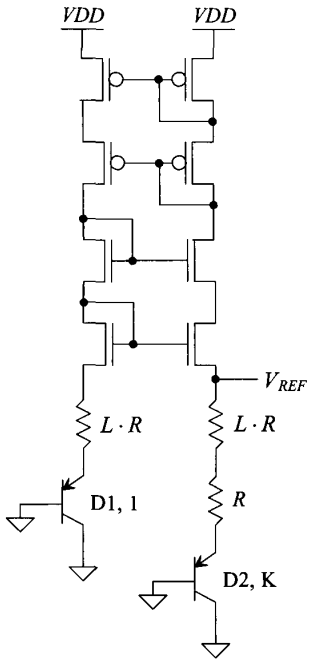
Figure 23.28 Simulating the BGR in Fig. 23.27 from 0 to 100 C.

Alternative BGR Topologies

We might wonder if there is a way to simplify the BGR of Fig. 23.27 to reduce power and layout area. Figure 23.29a shows one possibility. In this circuit we've combined the three branches in Fig. 23.27 into two branches. The same voltage, as in Fig. 23.25, is dropped across the single resistor R. The $L \cdot R$ resistors are a common element to both branches. The current is still PTAT, while the reference voltage at the top of the $L \cdot R$ resistor is still a bandgap voltage. The drawback of this topology is the fact that the minimum value of *VDD* for proper reference operation increases.

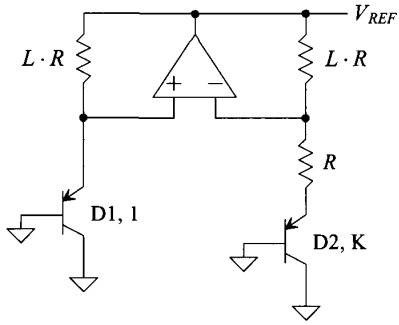
Figure 23.29b shows a topology based on the configuration, using resistors, seen in Fig. 23.16. As with all of the topologies in this chapter (again), the point of the added amplifier is to force the same current through each side of the reference. The benefit of this topology over the topology in (a) is that *VDD* can move lower before it affects V_{REF}. The drawback of this topology is that the amplifier must be capable of driving a resistive load (which we'll see in the next chapter requires two gain stages for the amplifier's overall gain to be reasonably high). In (c) PMOS devices are added to isolate the

IPR2023-00698
Theta EX2019

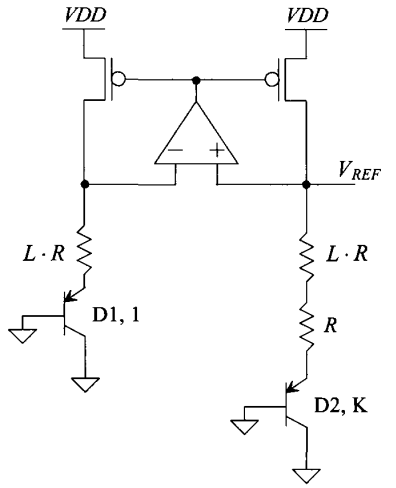


(a) Alternative BGR using cascodes.

Note: start-up circuits (required) not shown.



(b) Using an amplifier to force the same current through each branch of the reference, see Fig. 23.16.



(c) Using PMOS devices to isolate the amplifier from the resistive loading.

Figure 23.29 Alternative BGR topologies.

amplifier's output from the resistors. This makes using a diff-amp for the added amplifier possible. The currents are equal in (c) because the source-gate voltages of the PMOS devices are equal. Note how the branch containing D2 is a higher resistance (than the branch containing D1) and so it (the D2 branch) is always connected to an inverting point in the feedback loop. As seen in (c) the addition of the PMOS devices (which are inverting) means that we need to switch the inverting and noninverting amplifier inputs from (b). Finally, note that a start-up circuit is required for all three of these references.

IPR2023-00698
Theta EX2019

23.2.2 Short-Channel BGR Design

The bandgap reference voltage of 1.2 V developed in the last section is greater than *VDD* (= 1 V) in our short-channel process. To develop a BGR for short-channel processes, consider the schematic seen in Fig. 23.30. The diodes D1 and D2 together with the resistor, R, form a PTAT current generator, as seen in Fig. 23.24. To provide a CTAT current to sum with the PTAT current, consider the addition of the $L \cdot R$ resistors, as seen in the schematic. As temperature increases, the diode voltage decreases, causing the current through the $L \cdot R$ resistors to decrease (CTAT). We know that the current due to the PTAT portion of the circuit, see Eq. (23.31), is

$$I_{PTAT} = \frac{nV_T \cdot \ln K}{R} \tag{23.37}$$

The current through the added resistors (the CTAT portion of the circuit) is

$$I_{CTAT} = \frac{V_{D1}}{L \cdot R} \tag{23.38}$$

The total current is driven through the $N \cdot R$ to generate the reference voltage

$$V_{REF} = nV_T \cdot N \cdot \ln K + \frac{N}{L} \cdot V_{D1} \tag{23.39}$$

The temperature behavior of the BGR is

$$\frac{\partial V_{REF}}{\partial T} = n \cdot N \cdot \ln K \cdot \overbrace{\frac{\partial V_T}{\partial T}}^{0.085\ mV/C} + \frac{N}{L} \cdot \overbrace{\frac{\partial V_{D1}}{\partial T}}^{-1.6\ mV/C} \tag{23.40}$$

For zero TC, we get an L of

$$L = \frac{1.6}{n \cdot \ln K \cdot 0.085} \tag{23.41}$$

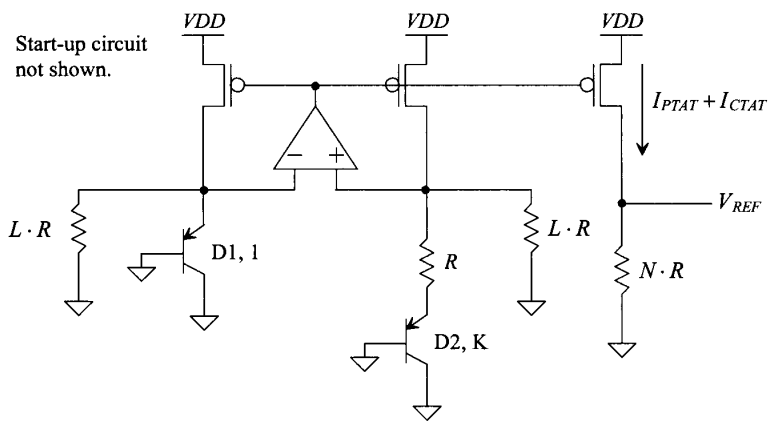


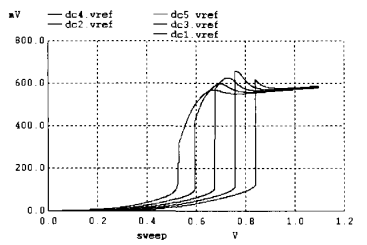
Figure 23.30 Lower voltage BGR.

IPR2023-00698
Theta EX2019

Using a K of 8 results in, again, an L of 9.41. To get a particular reference voltage, we use Eq. (23.39) to determine N. For example, if we want a reference voltage of 500 mV (half of VDD), we get an N of (using a K of 8)

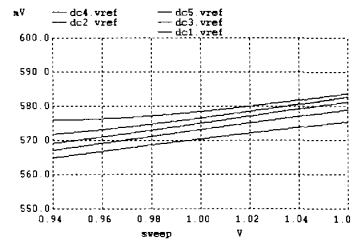
$$N = \frac{V_{REF}}{nV_T \cdot \ln K + \frac{V_{D1}}{L}} = \frac{0.5}{0.052 + \frac{0.7}{9.41}} = 3.91 \qquad (23.42)$$

Figure 23.31 shows some simulation results using these numbers (the schematic of the full design is seen in Fig. 23.32). In (a) the reference turns on at a VDD of approximately 900 mV. The temperature behavior of the reference is seen in (b). Notice that the reference voltage is higher than what we designed for. In all practical situations, the $N \cdot R$ resistor (the 205k resistor in Fig. 23.32) would need trimming (see Fig. 23.15) to set the reference voltage to a precise value.



VDD

(a) The BGR turns on at 900 mV.



VDD

(b) Close-up temperature behavior.

Figure 23.31 Simulating the behavior of the reference in Fig. 23.30.

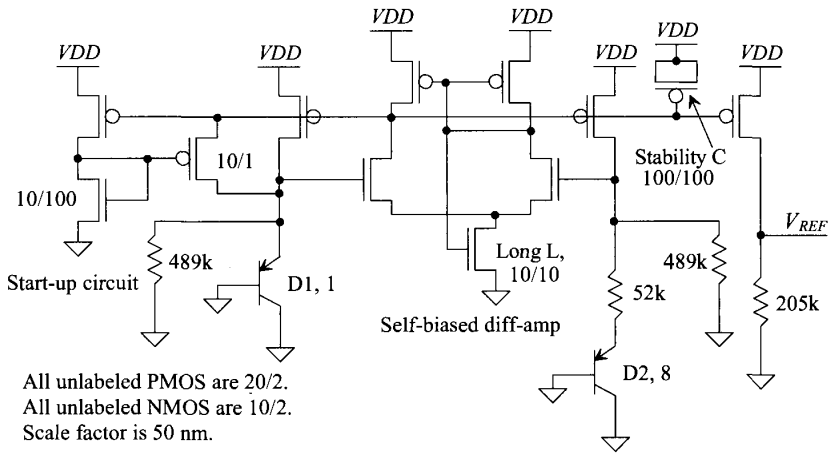


Figure 23.32 Lower voltage BGR used to generate the simulation data seen in Fig. 23.31.

IPR2023-00698
Theta EX2019

The Added Amplifier

In Fig. 23.32 we used a self-biased amplifier (see Fig. 18.17) to hold the voltage across D1 to the same value as the voltage across the 52k resistor and D2. This amplifier is critical for good BGR performance. For example, notice how the reference voltage in Fig. 23.31b varies with changes in *VDD*. By using a higher-gain amplifier, we can reduce this variation. While we'll discuss op-amp design in detail in the next chapter, one concern is the added amplifier's input common-mode voltage. Because the forward voltage drop of the diodes is approximately 0.7 V (a significant percentage of *VDD*), the performance (gain) of the amplifier may drop if any of the MOSFETs in the diff-amp triode or shut off. To lower the input common-mode voltage, consider the topology of Fig. 23.33. Here we set

$$R_1 + R_2 = L \cdot R \tag{23.43}$$

The amplifier sets the voltages across each R_2 to the same value and so the current flowing in R_1 and R_2 is the same as in Fig. 23.30, provided Eq. (23.43) is satisfied. The big benefit of using the topology is that the DC input common-mode voltage of the added amplifier can be reduced.

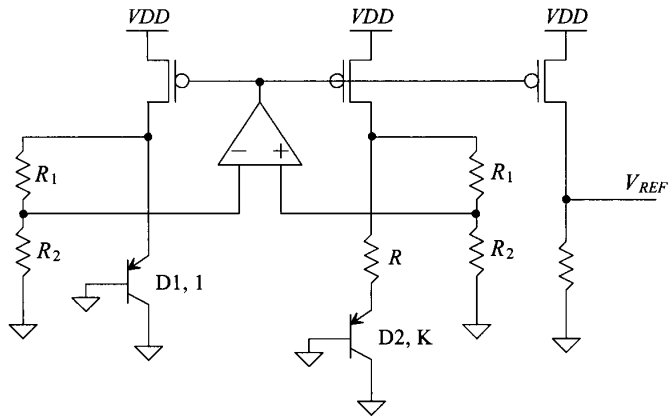


Figure 23.33 Lowering the input common-mode range of the added amplifier.

Lower Voltage Operation

Fundamentally, the limitation on how low *VDD* can be in the BGR reference of Fig. 23.30 (or 23.33) is the forward voltage drop of the diodes. As the dimensions of CMOS technology continue to scale downwards, so does *VDD*. While the MOSFET-resistor references, such as the beta-multiplier reference (BMR), discussed in the first section of this chapter may provide a solution to many reference circuit needs at lower *VDD*s, it is unlikely that they will replace the BGR in all applications. Even though both the BMR and BGR require trimming to set V_{REF}, the temperature performance of the BGR is, in general, better than that of the BMR. What may happen is that the pn junction used in the BGR is replaced by a junction with a lower built-in potential (and thus lower forward turn-on voltage) like the Schottky diode (see Fig. 23.21).

IPR2023-00698
Theta EX2019

ADDITIONAL READING

[1] K. N. Leung and P. K. T. Mok, "A Sub-1-V 15-ppm/C CMOS Bandgap Voltage Reference Without Requiring Low Threshold Voltage Device," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 4, April 2002, pp. 526–530. Develops the scheme seen in Fig. 23.33 for lowering the input common-mode range of the added amplifier.

[2] H. Banba, H. Shiga, A. Umezawa, T. Miyaba, T. Tanzawa, S. Atsumi, and K. Sakui, "A CMOS Bandgap Reference Circuit with Sub-1-V Operation," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 5, May 1999, pp. 670–674. Presents the design of the lower voltage BGR seen in Fig. 23.30.

[3] G. Tzanateas, C. A. T. Salama, and Y. P. Tsividis, "A CMOS Bandgap Voltage Reference," *IEEE Journal of Solid-State Circuits*, vol. SC-13, no. 3, June 1979, pp. 655–657. Good paper discussing BGR reference design.

[4] E. Vittoz and J. Fellrath, "CMOS Analog Integrated Circuits Based on Weak Inversion Operation," *IEEE Journal of Solid-State Circuits*, vol. SC-12, no. 3, June 1977, pp. 224–231. Covers the BMR design operating in the weak inversion region.

[5] K. E. Kuijk, "A Precision Reference Voltage Source," *IEEE Journal of Solid-State Circuits*, vol. SC-8, no. 3, June 1973, pp. 222–226. Development of the BGR seen in Fig. 23.29b.

PROBLEMS

23.1 Using the MOSFET-only reference seen in Fig. 23.2, design a nominally 500 mV reference in the short-channel CMOS process. Using simulations, characterize the sensitivity of the reference voltage to changes in *VDD* and temperature.

23.2 Use the long-channel CMOS process and the topology seen in Fig. 23.6 to design a voltage reference of $3V_{THN}$. Simulate your design and show the *VDD* sensitivity and temperature behavior of the reference. Do TCs of R_1 and R_2 affect the performance of the reference? Why or why not?

23.3 Suppose it was desired, in Fig. 23.7 (see also Fig. 20.22), to make M1 and M2 the same size. However, to increase the gate source voltage of M1, relative to M2, the width of M3 is increased by *K*. How do the equations governing the operation of the BMR change? How does the current flowing in the BMR change?

23.4 Verify that if the PMOS devices in Fig. 23.11 are not cascoded (that is, they have only NMOS cascodes), the currents in each branch will not be equal and there will be significant sensitivity to *VDD* (a sensitivity similar to what is seen in Fig. 23.10).

23.5 In a CMOS process, several of the layers including poly, n+, p+, and n-well can be used for resistor formation. Each of these layers has a different temperature coefficient (TC). For the BMR that generated Fig. 23.14, use simulations to determine the optimum resistor TC.

23.6 Derive the equations that govern the operation of the reference in Fig. 23.16b.

IPR2023-00698
Theta EX2019

23.7 Show why n+ directly in the p-substrate cannot be used as a diode in a CMOS process.

23.8 Generate a diode model that produces a forward voltage drop of 700 mV when driven with 1 μA and has a change with temperature, dV_D/dT, near room temperature, of −2 mV/C. Use simulations to verify your model meets the requirements.

23.9 Generate a SPICE model for the Schottky diode seen in Fig. 23.21. Assume that the series resistance of the diode is 1 kΩ.

23.10 Estimate, using hand calculations, the minimum allowable *VDD* for the reference of Ex. 23.4. What are the PMOS and NMOS gate-source and drain-source voltages when the reference current is 1 μA? Note that the parameters in Table 9.1 have nothing to do with the operating conditions in this question. Verify your answers using SPICE.

23.11 Show that K forward-biased diodes in parallel behave like a single diode with a scale current of $K \cdot I_S$, as assumed in Eq. (23.30).

23.12 Suggest a reference design that would output a voltage of $n \cdot V_T$.

23.13 Determine whether the performance of the BGR of Fig. 23.32 can be enhanced by using the topology of Fig. 23.33. Use simulations to verify your answer.

IPR2023-00698
Theta EX2019

Chapter

24

Operational Amplifiers I

The operational amplifier (op-amp) is a fundamental building block in analog integrated circuit design. A block diagram of the two-stage op-amp with output buffer is shown in Fig. 24.1. The first stage of an op-amp is a differential amplifier. This is followed by another gain stage, such as a common source stage, and finally by an output buffer. If the op-amp is intended to drive a small purely capacitive load, which is the case in many switched capacitor or data conversion applications, the output buffer is not used. If the op-amp is used to drive a resistive load or a large capacitive load (or a combination of both), the output buffer is used.

Design of the op-amp consists of determining the specifications, selecting device sizes and biasing conditions, compensating the op-amp for stability, simulating and characterizing the op-amp A_{OL} (open-loop gain), CMR (common-mode range on the input), CMRR (common-mode rejection ratio), PSRR (power supply rejection ratio), output voltage range, current sourcing/sinking capability, and power dissipation.

We'll start this chapter off with a very simple two-stage op-amp (without an output buffer). By pointing out the weaknesses with this two-stage topology, we'll set the stage for developing practical op-amps in the rest of the chapter.

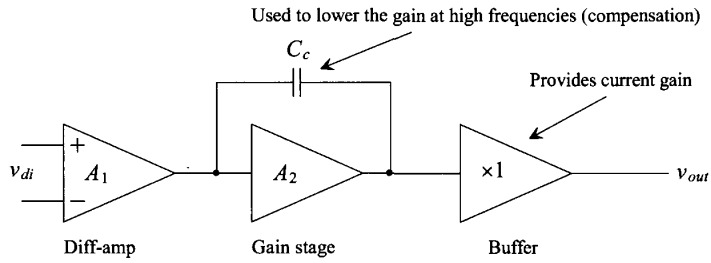


Figure 24.1 Block diagram of two-stage op-amp with output buffer.

IPR2023-00698
Theta EX2019

24.1 The Two-Stage Op-Amp

Figure 24.2 shows the basic two stage op-amp made using an NMOS diff-amp and a PMOS common-source amplifier (M7). As seen in Fig. 22.8 M7 is biased to have the same current as M3 and M4 (10 μA from Table 9.2). Note also the addition of the compensating network consisting of a compensation capacitor, C_c, (*Miller compensation*) and a zero-nulling resistor R_z (see Figs. 21.25 and 21.33 along with the associated discussion). Because the op-amp doesn't have an output buffer, it is limited to driving capacitive loads and very large resistances (comparable to the output resistance of a MOSFET, that is, megaohms).

Low-Frequency, Open Loop Gain, A_{OLDC}

The low-frequency, open loop gain of the op-amp is calculated as the product of each stage gain, that is,

$$A_{OLDC} = A_1 \cdot A_2 = \overbrace{g_{mn} \cdot (r_{on} || r_{op})}^{A_1 = \text{diff-amp's gain}} \cdot \overbrace{g_{mp} \cdot r_{op}}^{A_2,\ \text{M7's gain}} \qquad (24.1)$$

where the output resistance of the cascode current source load, M8, is assumed to be much larger than the M7 output resistance, r_{op}. Using the values from Table 9.2, we get an A_{OLDC} of 832 *V/V*.

Input Common-Mode Range

The minimum input common-mode voltage is given by Eq. (22.11) or 450 mV. The maximum input common-mode voltage is given by Eq. (22.12) or 930 mV. This means, for proper operation of our two-stage op-amp, the input voltages (v_p and v_m) should fall within the range of 450 to 930 mV. If they go outside this range, the op-amp gain drops, and it is likely that the circuit employing the op-amp will not function properly. Figure 24.3 shows a SPICE DC sweep where the *inverting* input (v_m) is held at 500 mV and the *noninverting* input (v_p) is swept from 495 to 505 mV. The slope of this transfer curve is the DC open-loop gain of the op-amp, A_{OLDC}.

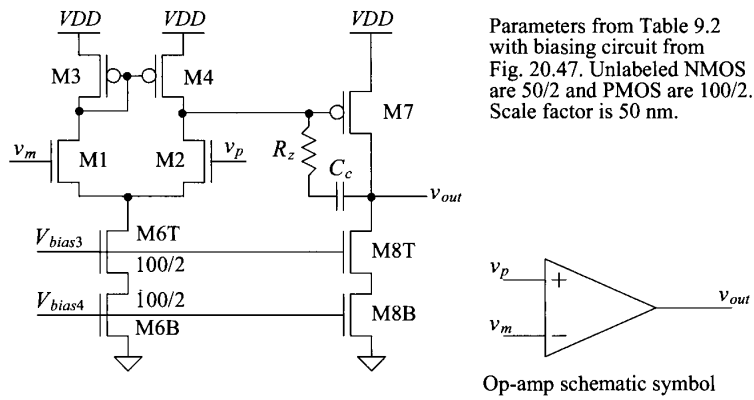


Figure 24.2 Basic two-stage op-amp.

IPR2023-00698
Theta EX2019

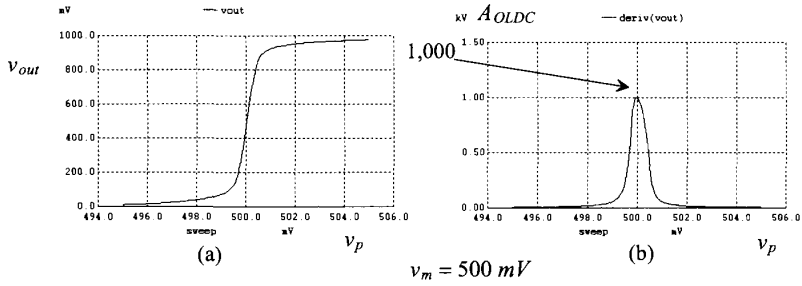


Figure 24.3 (a) DC transfer curves for the op-amp in Fig. 24.2 and (b) its gain (the derivation of (a)).

Power Dissipation

To determine the power dissipated by the op-amp, we sum the currents supplied by the constant current sources and multiply the result by *VDD*. For the op-amp in Fig. 24.2, the current through M6 is 20 μA and the current through M8 is 10 μA. The total power dissipation is then 30 μW ($VDD = 1\ V$).

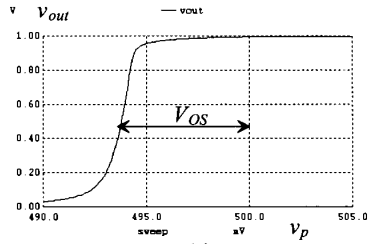
Output Swing and Current Sourcing/Sinking Capability

The maximum output swing (for the op-amp in Fig. 24.2) is limited by M7 going into the triode region. If we must keep at least 100 mV across M7, then the maximum output voltage is 900 mV. When M8 goes into triode or roughly 100 mV (see Fig. 20.48), the minimum output voltage is set. As seen in Fig. 24.3a, the high-gain (or large slope) region falls between v_{out} of 100 and 900 mV. Note that the maximum amount of current that this op-amp can sink is limited by the constant current sink M8 or 10 μA. The op-amp can source considerably more than 10 μA by pulling the gate of M7 downwards. Because this topology can source a considerable amount of current, it is useful as a voltage regulator (because the op-amp is always only sourcing current in a voltage regulator application).

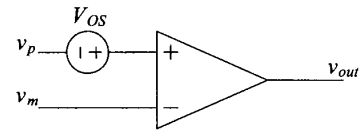
Offsets

We've talked in great length (earlier in the book) about random offsets and how to design and layout circuits to minimize their effects. Another type of offset (that is not random) is termed a *systematic offset*. When we sized the MOSFETs in Fig. 24.2, for example, we made sure that M7 was sized to source 10 μA and M8 was sized to sink 10 μA. What would happen if we sized M7 to source 100 μA of current instead (changed its size from 100/2 to 1000/2)? Because M8 is a constant bias of 10 μA, M7 would move into the triode region until it was sourcing the 10 μA of current that M8 wants to sink. The output voltage would be very close to *VDD* with M7 in the triode region. Effectively we would get a shift or offset in the transfer curves seen in Fig. 24.3a (see Fig. 24.4a). To model this output voltage shift, we can refer it back to the op-amp input as an input-referred offset voltage, Fig. 24.4b. Note that unlike a random offset, which may be positive or negative in value, a systematic offset will always be a known polarity. Finally, note that while we chose to model the offset in series with the noninverting input terminal, in Fig. 24.4, we could just as easily have placed it in series with the inverting op-amp terminal (with a change in polarity).

IPR2023-00698
Theta EX2019



(a)



(b) How the offset is modeled

$v_m = 500\ mV$

Figure 24.4 Showing how increasing M7's width to 1000 in Fig. 24.2 causes an input-referred (systematic) offset voltage.

Compensating the Op-Amp

An important step in op-amp design is the design of the compensation network. The op-amp takes the difference between the inverting and noninverting input terminal voltages and, ideally, multiplies the result by a very large number (the gain). A block diagram representation of an op-amp is seen in Fig. 24.5. The open-loop gain as a function of frequency is labeled $A_{OL}(f)$.

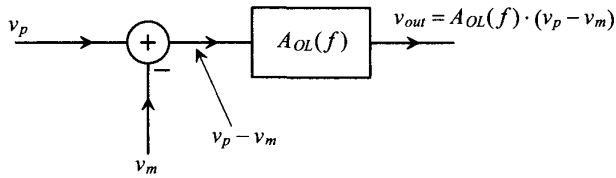


Figure 24.5 Block level diagram of an op-amp.

In all practical situations, the op-amp is used with feedback, Fig. 24.6. While our op-amp in Fig. 24.2 cannot drive a resistive load (unless it is megaohms, as a resistor on the op-amp output will kill the gain of the second stage), we'll still use this figure (Fig. 24.6) to illustrate the concept of feedback and compensation. We can write

$$v_{out} = A_{OL}(f) \cdot (v_{in} - v_f) \tag{24.2}$$

and

$$v_f = v_{out} \cdot \frac{R_2}{R_1 + R_2} \tag{24.3}$$

The amount of the output that is fed back is often called the feedback factor β or

$$\beta = \frac{R_2}{R_1 + R_2} \tag{24.4}$$

Substituting Eqs. (24.3) and (24.4) into Eq. (24.2) and solving for the closed loop gain gives

IPR2023-00698
Theta EX2019

$$A_{CL}(f) = \frac{v_{out}}{v_{in}} = \frac{A_{OL}(f)}{1 + \beta \cdot A_{OL}(f)} \tag{24.5}$$

If we take $A_{OL}(f) \to \infty$, then the closed loop gain of this non-inverting topology is

$$A_{CL}(f) \to \frac{1}{\beta} = 1 + \frac{R_1}{R_2} \tag{24.6}$$

We have several important points that we need to discuss. To begin, notice that in Eq. (24.5) if

$$\beta \cdot A_{OL}(f) = -1 \tag{24.7}$$

or more precisely

$$|\beta \cdot A_{OL}(f)| = 1 \text{ and } \angle\beta \cdot A_{OL}(f) = \pm 180° \tag{24.8}$$

the closed loop gain blows up (the feedback amplifier becomes unstable). The worst case situation (the largest value of β) occurs when all of the output is fed back to the op-amp input (assuming no transformer, amplifier, etc. in the feedback path). The voltage follower, Fig. 24.7, is an example of this situation. *To determine the stability of an op-amp, we'll look at the open loop gain when the feedback factor is one, that is,*

$$|A_{OL}(f)| = 1 \text{ and } \angle A_{OL}(f) = 180° \tag{24.9}$$

Notice that the larger the closed loop gain, the smaller the value of β (the less output signal we feed back) and the more likely the op-amp circuit, with feedback, will be stable. *This is important.* While feedback helps to desensitize an amplifier's gain to variations in an op-amp's A_{OL}, the drawback is stability. In a high-performance op-amp that will never operate in a unity-follower configuration, we can get an enhancement in speed by reducing the amount of compensation (the amount of the reduction guided by Eq. (24.8) with the actual value of β used). While this discussion has focused on the non-inverting topology (voltage or series-shunt amplifiers, $\beta = v_f/v_{out}$) a similar discussion can be given for the inverting amplifier (shunt-shunt where $\beta = i_f/v_{out}$) topology, see Sec. 30.2.

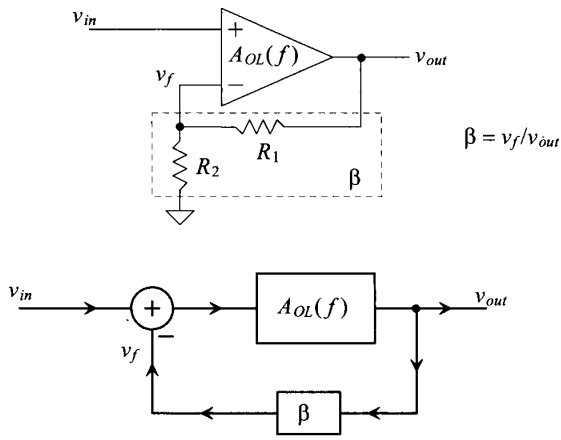


Figure 24.6 An example of feedback in an op-amp, see Sec. 30.2 for additional discussion.

IPR2023-00698
Theta EX2019

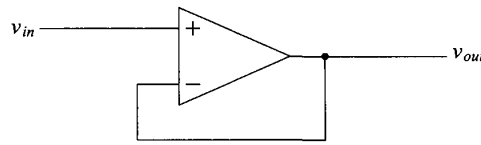


Figure 24.7 Voltage follower configuration, an example of a closed-loop amplifier with unity feedback factor.

To estimate the open-loop frequency response of the op-amp, let's use the generic model seen in Fig. 21.25. Figure 24.8 shows the location of nodes 1 and 2 on our two-stage op-amp. We've included a load capacitance in this figure. With the help of Table 9.2 we can write

$$R_1 = r_{on} || r_{op} = 111\ k\Omega$$

$$R_2 = r_{op} || R_{ocasn} \approx r_{op} = 333\ k\Omega$$

$$g_{m1} = g_{mn} = 150\ \mu A/V \text{ (diff-amp)}$$

$$g_{m2} = g_{mp} = 150\ \mu A/V \text{ (common-source)}$$

$$C_1 = C_{dg4} + C_{gd2} + C_{gs7} = 13.6\,fF$$

$$C_2 = C_L + C_{gd8} \approx C_L + 1.56\,fF$$

Let's calculate the open-loop response with a load and compensation capacitance of 100 fF, that is, $C_L = C_c = 100\,fF$. The pole associated with node 1, from Eq. (21.65), is

$$f_1 \approx \frac{1}{2\pi g_{m2} R_1 R_2 C_c} = 287\ kHz$$

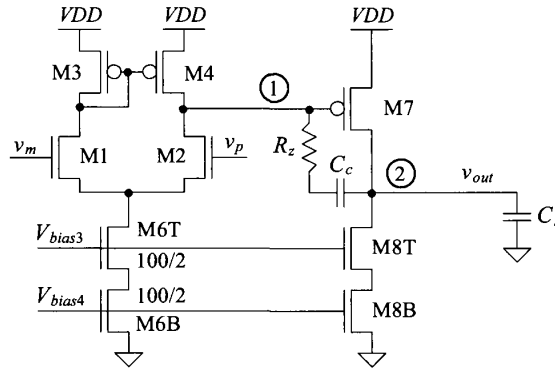


Figure 24.8 Calculating the frequency response of the op-amp.

IPR2023-00698
Theta EX2019

The location of the pole associated with the output node (node 2) is, from Eq. (21.66),

$$f_2 = \frac{g_{m2}C_c}{2\pi(C_cC_1 + C_1C_2 + C_cC_2)} = 210\ MHz$$

The location of the zero, f_z in Eq. (21.63), is at 240 MHz (very near the second pole).

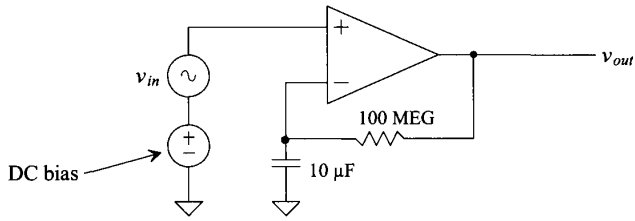


Figure 24.9 Circuit configuration used to simulate open-loop frequency response.

To simulate the open-loop response of the op-amp, we can use the configuration seen in Fig. 24.9. The feedback resistor and capacitor form a time constant so large that for all intents and purposes none of the AC output voltage is fed back to the inverting input. However, the DC bias level is fed back so that the op-amp biases up correctly (all MOSFETs are operating in the saturation region). With a DC bias voltage of 500 mV, Fig. 24.10 shows the open loop responses of the op-amp in Fig. 24.8 with 100 fF compensation capacitance and load capacitance. The simulated values are close to the

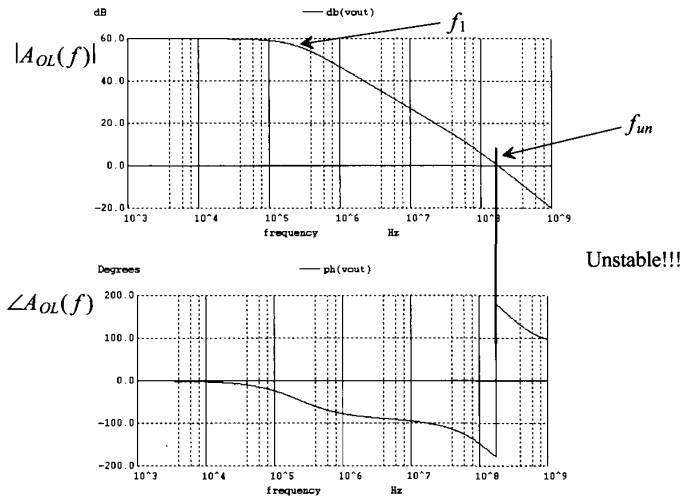


Figure 24.10 The open-loop frequency response of the op-amp in Fig. 24.8 with a 100 fF compensation capacitance and load capacitance.

IPR2023-00698
Theta EX2019

hand-calculated values given above. However, from the criteria given in Eq. (24.9), the op-amp will clearly be unstable. *We want the open loop gain of the op-amp to be much less than one when the phase shift is 180°.* Looking at Fig. 24.10, we effectively want to shift the lower frequency pole (f_1) downwards in frequency. At the same time, we want to move the higher frequency pole, f_2 , higher in frequency (we want to *split the poles*). Remember this (pole-splitting) was the point of adding the compensation capacitor to an amplifier back in Ch. 21. Let's use Eq. (21.72) to select C_c. Looking at Fig. 24.10, we see that the phase shift is −100° at 10 MHz, so let's set the unity gain frequency to this value (in an attempt to make sure that the open loop gain is well under one when the phase shift is 180°).

$$f_{un} = \frac{g_{m1}}{2\pi C_c} = \frac{150\ \mu A/V}{2\pi \cdot C_c} = 10\ MHz \rightarrow C_c = 2.4\ pF \qquad (24.10)$$

The simulation results are seen Fig. 24.11. The unity-gain frequency, f_{un} , is close to 10 MHz. However, while the lower frequency moved downwards and the higher frequency pole moved upwards, the zero moved down to unity-gain frequency, see Eq. (21.63). This causes the open-loop gain to hover around unity instead of continuing to decrease (this is bad). The stability and step response of the op-amp will be degraded because of the zero. To illustrate this, consider the simulation results seen in Fig. 24.12. Notice that, in a transient simulation, we have to wait some time to let the bias circuit start up. Here, in this simulation, we've waited 500 ns before applying the input signal. Next, notice that our step input signal has a small amplitude change (5 mV). If we apply a larger step, the small-signal analysis used to derive pole-splitting is no longer valid (and we'll see the slew-rate limitations). Clearly the output of the op-amp doesn't behave well with the zero present. To eliminate the zero, we add the zero-nulling resistor, as seen in Fig. 21.33b. Figure 24.11 is regenerated in Fig. 24.13 with the addition of a zero-nulling resistor ($1/g_{mn}$ = 6.5k). Figure 24.14 shows the well-behaved step response.

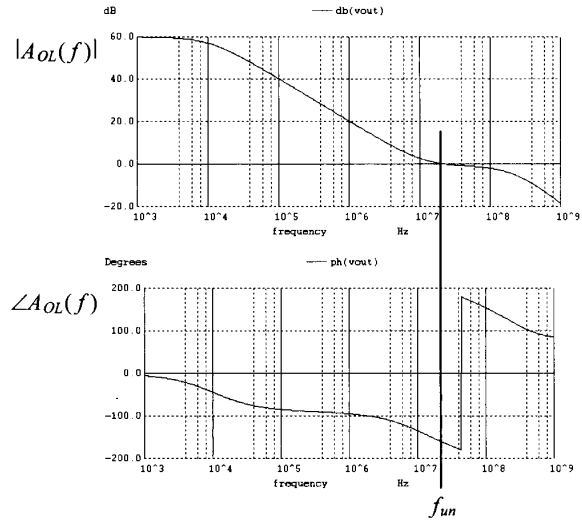


Figure 24.11 Increasing the compensation capacitor's value to 2.4 pF.

IPR2023-00698
Theta EX2019

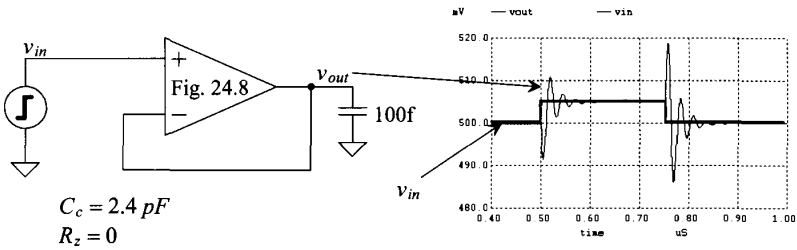


Figure 24.12 The poor step response of the op-amp with the zero present.

Gain and Phase Margins

In the previous discussions we assumed a load of 100 fF. If the load capacitance varies, the stability can (will) be affected. Further, with temperature, process, and power supply variations, the stability of the op-amp can change. In order to specify "how stable" the op-amp is at a given set of operating conditions, the parameters gain margin (GM) and phase margin (PM) are used. To determine an op-amp's PM, we look at the phase shift when the open-loop gain is unity. The amount of phase shift away from 180° is the PM of the op-amp. As seen in Fig. 24.13, the phase shift when the gain is unity is –90°. Taking the difference between this value and 180° gives a PM of 90°. To calculate the GM, we look at the difference between the open-loop gain and unity when the phase of the op-amp is ±180°. For the op-amp response in Fig. 24.13, the GM is approximately 25 dB. Note that as seen in Fig. 24.14, it is nice to have a PM of 90° because the step response of the op-amp has a first-order response (like an RC circuit).

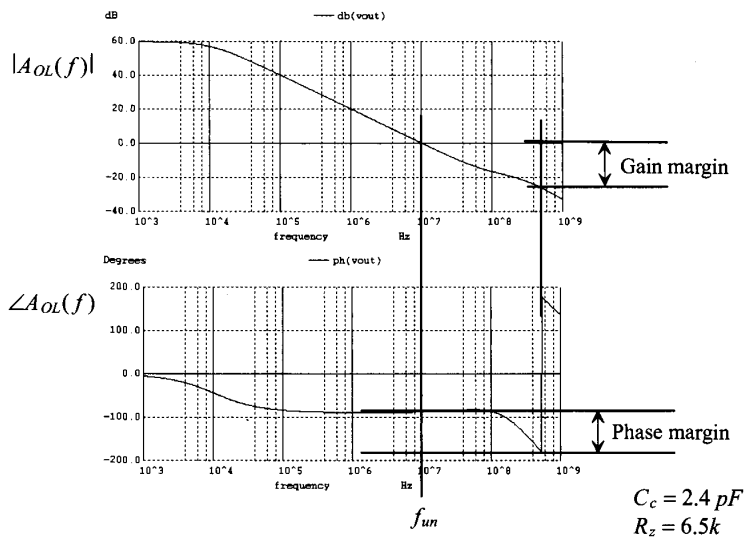


Figure 24.13 Adding a zero nulling resistor to the op-amp in Fig. 24.8.

IPR2023-00698
Theta EX2019

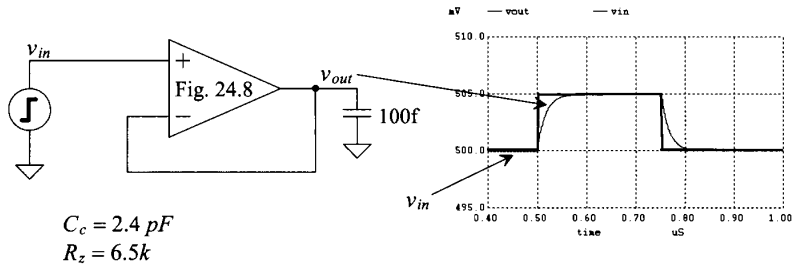


Figure 24.14 Good step response of the op-amp with the zero absent.

Removing the Zero

As the previous discussion illustrated, the RHP zero can raise issues concerning stability and settling time (this is important). As seen in Eq. (21.73) and the associated discussion, R_z can be added to eliminate ($R_z = 1/g_{m1}$) or move the zero into the LHP ($R_z > 1/g_{m1}$). When the zero is moved to the LHP, the phase response of the zero adds to the overall phase response, increasing the PM (this is called *lead compensation*). The practical problem with using R_z is that setting its value to a precise number (say $1/g_{m1}$) is challenging with shifts in process, temperature, or voltage. One solution to this problem is to replace the resistor with a MOSFET operating in the triode region, Fig. 24.15. Mz behaves like a resistor with a value of $1/g_{m1}$. Ideally, the source-gate voltage of M7 is the same as the V_{SG} of MP1. It then follows that the source-gate potential of MP2 equals the source-gate potential of Mz. The channel resistance of Mz is then, from Eq. (9.16), set to $1/g_{m1}$. The practical issues with this method are the wasted power dissipated by the additional circuitry and the fact that if the output swing becomes large (especially at higher frequencies where C_c has a small impedance), Mz can move out of the triode region, which can affect the large-signal behavior of the op-amp.

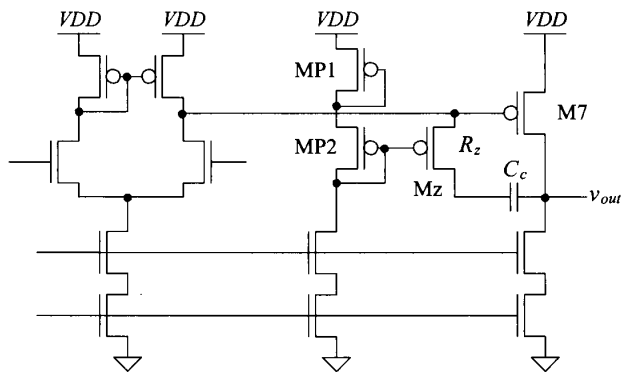


Figure 24.15 Making the zero-nulling resistor process independent.

IPR2023-00698
Theta EX2019

The other method to remove the RHP zero seen in Fig. 21.33a is to add an amplifier with a gain of +1 in series with the compensation capacitor. Figure 24.16 shows the idea. A source follower allows the output signal to feed back through the compensation capacitor (so that the pole-splitting effect is still present). However, the root cause of the RHP zero, namely, C_c shorting the input of the second stage (the output of the diff-amp) to the output of the second stage (the output of the op-amp) at higher frequencies is removed. The concerns with this topology are, again, power dissipation and, perhaps more importantly, large signals. If the overall output voltage swings too low, it causes the source-follower to shut off.

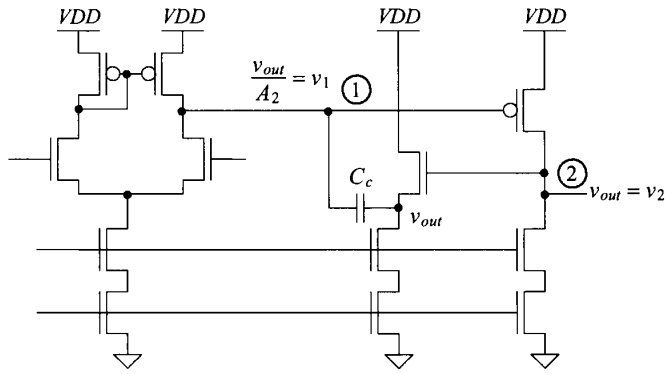


Figure 24.16 Using an amplifier to eliminate forward signal feedthrough via the compensation capacitor.

Compensation for High-Speed Operation

Notice, in Fig. 24.16 (or in any other of the topologies we've discussed up to this point), that the current fed back through C_c is

$$i_{Cc} = \frac{v_{out} - \frac{v_{out}}{A_2}}{1/j\omega C_c} \tag{24.11}$$

If the second-stage gain, A_2 ($g_{m2}R_2 = g_{m7}r_{o7}$) is reasonably large, then this equation can be approximated using

$$i_{Cc} \approx \frac{v_{out}}{1/j\omega C_c} \tag{24.12}$$

If we can feed back this current *indirectly* to the output of the diff-amp, we can still compensate the op-amp (and have pole-splitting). Further, if we do it correctly, we avoid connecting the compensation capacitor directly to the output of the diff-amp and thus avoid the RHP zero. Towards this goal, consider the modified op-amp schematic seen in Fig. 24.17. The added MOSFETs form a common-gate amplifier. (Note that here we are assuming $v_p << v_{out}$, so the source-follower action of MCG is negligible. MCG is connected to v_p to set its gate at the DC voltage of the input.) The current i_{Cc} is fed back through the common-gate MOSFET, MCG, to node 1 (the output of the diff-amp).

IPR2023-00698
Theta EX2019

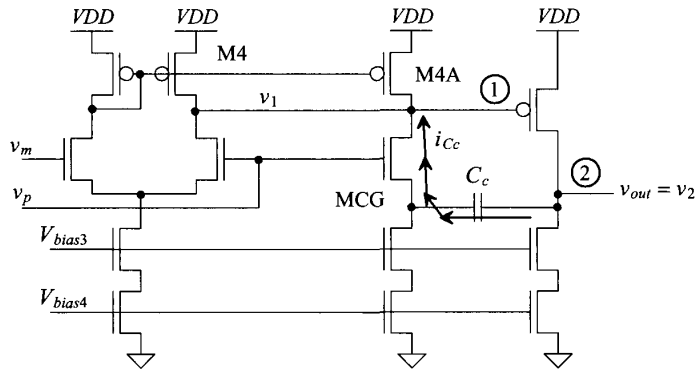


Figure 24.17 Feeding back a current indirectly to avoid the RHP zero.

To determine the frequency response of this amplifier, consider the model seen in Fig. 24.18 (modified from Fig. 21.25). Summing the currents at node 1 gives

$$-g_{m1}v_s + \frac{v_1}{R_1 \| \frac{1}{j\omega C_1}} - \overbrace{\frac{v_{out}}{1/j\omega C_c + 1/g_{mcg}}}^{i_{Cc}} = 0 \qquad (24.13)$$

where the resistance looking into the source of MCG is $1/g_{mcg}$. Solving for v_1 gives

$$v_1 = \frac{g_{m1}R_1}{1+j\omega R_1 C_1} \cdot \left(\frac{j\omega \frac{C_c}{g_{m1}}}{1+j\omega \frac{C_c}{g_{mcg}}} \cdot v_{out} + v_s \right) \qquad (24.14)$$

For the output node (node 2), we can write (noting that now C_2 includes both C_c and C_L)

$$v_{out} = -g_{m2}v_1 \cdot \left(\frac{R_2}{1+j\omega R_2 C_2} \right) \qquad (24.15)$$

Plugging Eq. (24.14) into this equation gives

$$v_{out} = -g_{m1}R_1 g_{m2}R_2 \cdot \frac{\frac{j\omega \frac{C_c}{g_{m1}}}{1+j\omega \frac{C_c}{g_{mcg}}} \cdot v_{out} + v_s}{(1+j\omega R_1 C_1)(1+j\omega R_2 C_2)} \qquad (24.16)$$

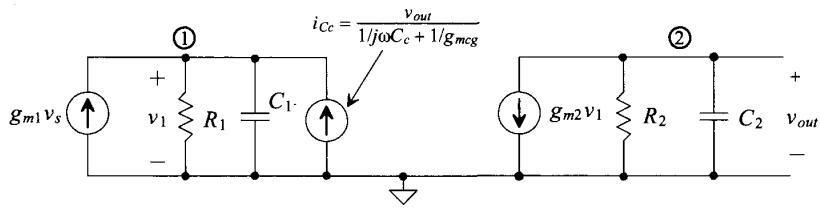


Figure 24.18 Model used to estimate bandwidth when indirect feedback current is used.

IPR2023-00698
Theta EX2019

Letting

$$K = \frac{-g_{m1}R_1 g_{m2}R_2}{(1+j\omega R_1 C_1)(1+j\omega R_2 C_2)} \tag{24.17}$$

we can rewrite Eq. (24.16) as

$$v_{out} = \frac{j\omega\frac{C_c}{g_{m1}}}{1+j\omega\frac{C_c}{g_{mcg}}} \cdot K \cdot v_{out} + K \cdot v_s \tag{24.18}$$

and, assuming $|K \cdot \frac{C_c}{g_{m1}}| >> \frac{C_c}{g_{mcg}}$, thus

$$\frac{v_{out}}{v_s} \approx \frac{K \cdot \left(1+j\omega\frac{C_c}{g_{mcg}}\right)}{1-j\omega\frac{C_c K}{g_{m1}}} \tag{24.19}$$

Resubstituting in K gives

$$\frac{v_{out}}{v_s} = \frac{-g_{m1}R_1 g_{m2}R_2\left(1+j\omega\frac{C_c}{g_{mcg}}\right)}{(1+j\omega R_1 C_1)(1+j\omega R_2 C_2)+j\omega\frac{C_c}{g_{m1}} \cdot g_{m1}R_1 g_{m2}R_2} \tag{24.20}$$

or, with $s = j\omega$, we can write

$$\frac{v_{out}}{v_s} = \frac{-g_{m1}R_1 g_{m2}R_2\left(1+s\frac{C_c}{g_{mcg}}\right)}{s^2 \cdot (R_1 C_1 R_2 C_2)+s \cdot (R_1 C_1 + R_2 C_2 + R_1 g_{m2} R_2 C_c)+1} \tag{24.21}$$

Notice there is a LHP zero at

$$f_z = \frac{g_{mcg}}{2\pi C_c} \tag{24.22}$$

Since this zero is in the LHP, it will add to the phase response and enhance the speed of the op-amp. Intuitively, we can think that at high speeds the phase shift through C_c will cause the output signal to feed back and add to the signal at node 1. This *positive feedback* enhances the speed of the op-amp. To determine the location of the second pole, let's assume $R_1 g_{m2} R_2 C_c >> R_1 C_1$ or $R_2 C_2$ or $R_1 C_1 R_2 C_2$ so

$$s_{1,2} \approx \frac{-R_1 g_{m2} R_2 C_c \pm R_1 g_{m2} R_2 C_c}{2(R_1 C_1 R_2 C_2)} \tag{24.23}$$

Our approximation won't tell us the location of the lower frequency pole (it's given by Eq. (21.65)). The location of the second pole is

$$f_2 \approx \frac{g_{m2}C_c}{2\pi \cdot C_1 C_2} \approx \frac{g_{m2}C_c}{2\pi \cdot C_1(C_L + C_c)} \tag{24.24}$$

This result should be compared to Eq. (21.66). The location of the second pole is at a considerably higher frequency using this technique. *The result is that we can set the unity gain frequency to a higher value and still have a stable op-amp.* Further, the load capacitance (which is included in C_2) can be considerably larger for a given PM or GM. Again, Eq. (21.72) can be used to set the unity-gain frequency, f_{un}, assuming $f_2 \approx f_z$

$$f_{un} = \frac{g_{m1}}{2\pi C_c} \ (\approx f_z \ \text{if} \ g_{m1} \approx g_{mcg}) \tag{24.25}$$

IPR2023-00698
Theta EX2019

Using the *indirect compensation* method seen in Fig. 24.17 in the op-amp of Fig. 24.8 with C_c = 240 *fF* (f_{un} = 100 MHz) gives the results seen in Fig. 24.19. The small-signal step response is seen in Fig. 24.20. This response should be compared to Figs. (24.12) and (24.14).

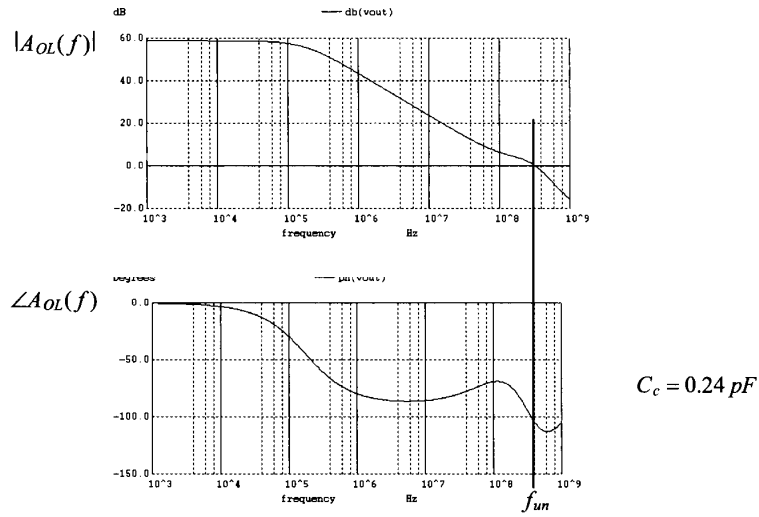


Figure 24.19 Simulating the op-amp of Fig. 24.8 using the indirect compensation scheme seen in Fig. 24.17 with a compensation capacitor of 240 fF.

The indirect feedback of the current through the compensation capacitor results in faster op-amp circuits and less layout area (the compensation capacitor generally dominates the layout area of an op-amp). However, the added circuit in Fig. 24.17 dissipates more power. To eliminate the additional power dissipation, consider the op-amp topology seen in Fig. 24.21. Here we've used the fact that a 100/2 PMOS device can be laid out as two 100/1 PMOS in series (see also Fig. 20.35). The current through the compensation capacitor is fed back through M4B to the output of the diff-amp. Note how we've also split M7 into two devices. It's important, for small offsets, to try to match both the drain-source and the gate-source voltages of the MOSFETs used in the op-amp. The step-response of the op-amp is seen in Fig. 24.22.

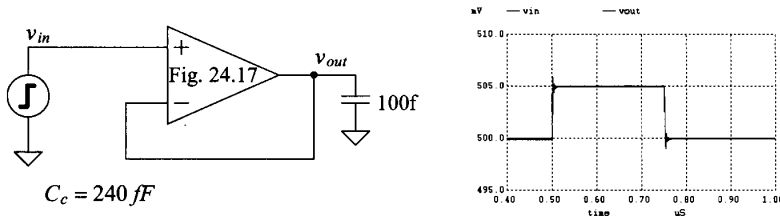


Figure 24.20 The step response of the op-amp in Fig. 24.17 with frequency response seen in Fig. 24.19.

IPR2023-00698
Theta EX2019

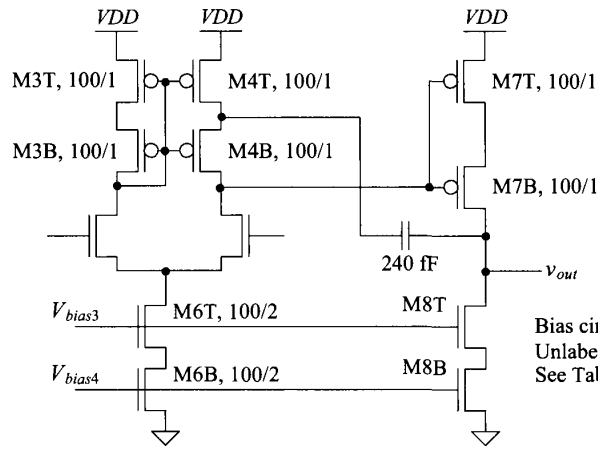


Figure 24.21 Implementing indirect feedback compensation without additional power dissipation in a two-stage op-amp.

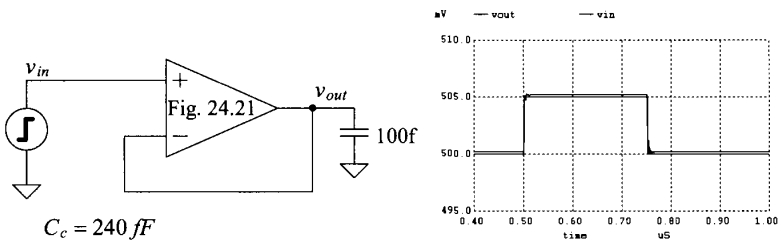


Figure 24.22 The step response of the op-amp in Fig. 24.21 driving 100 fF.

Slew-Rate Limitations

When we talked about step response, we limited the amplitude of the signals to small steps (5 mV) to avoid slew-rate limitations. Let's take the op-amp from Fig. 24.8 and put it in the configuration seen in Fig. 24.14, driving a 1 pF load capacitance. Further, let's increase the input pulse amplitude to (close to) the maximum allowable range, that is, from 500 mV to 900 mV (limited by the op-amp's input common-mode range). The simulated results are seen in Fig. 24.23. Referring to Fig. 24.8, notice that when the op-amp's noninverting input terminal (v_p) is driven high, M2 turns on and pulls the gate of M7 down. Since there isn't a current source in series with M7, it can quickly charge the 1 pF load capacitance. The only large-signal limitation is the bias current of the diff-amp, I_{SS} (= 20 μA here), charging C_c (= 2.4 pF here). This limits the output rate of change since any variation in v_{out} causes a displacement current through C_c, which must be sourced/sunk by the diff-amp. This slew-rate limitation is calculated as 20 μA/2.4 pF or 8.3 mV/ns. The change in the output voltage is 400 mV, so we would require roughly 50 ns to change the voltage across C_c.

IPR2023-00698
Theta EX2019

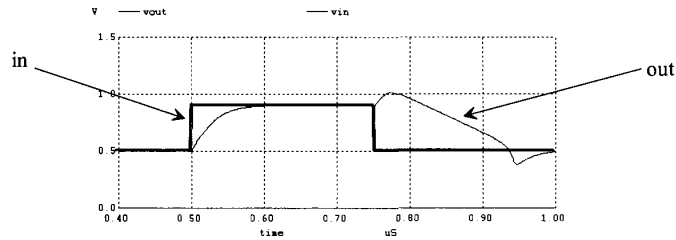


Figure 24.23 The large-signal (500 mV to 900 mV) performance of the op-amp in Fig. 24.8 with $C_c = 2.4$ pF and $R_z = 6.5$ k driving a 1pF load. The low-to-high settling time is roughly 100 ns, while the high-to-low settling time, which is slew-rate limited, is roughly 500 ns.

The more interesting behavior occurs when the input signal drops from 900 mV to 500 mV. This causes M2 to shut off, allowing all of the current from M4 to charge the gate of M7. The result is a positive pulse on the gate of M7 that feeds directly to the output through the compensation capacitor (this is bad, and it is another example of why it is desirable to use the indirect compensation method). This pulse accounts for the (unwanted) positive movement in the output voltage seen in Fig. 24.23 just after the input pulse transitions negative. Now the load capacitor and the compensation capacitor must both be discharged through the constant current source M8. The output slew-rate is estimated by how fast this constant current (here 10 μA) can discharge 3.4 pF (the sum of the compensation capacitor and the load capacitor). This rate is calculated as (roughly) 3 mV/ns. For the output to transition 400 mV requires 133 ns (this time will be longer because of the positive movement in the output voltage just after the input switches). Note that we might think that simply by sizing up the current conducting in M8 we can enhance the op-amp's slew-rate. This is true to the point where the diff-amp's current charging the compensation capacitor becomes the limiting factor.

Finally, note how the output voltage shoots past the final (desired) voltage level of 500 mV. This is because the gate of M7 is pulled to *VDD* to shut it off when slewing is taking place (both M4 and M7 are shut off because of this). When the inputs of the op-amp move to the same value (500 mV here), the gate of M7 must be pulled downwards to the quiescent value. The time it takes for this to happen results in the undershoot in the op-amp's output voltage seen in Fig. 24.23.

The same test setup used to generate the data in Fig. 24.23 is also used to generate the data in Fig. 24.24 except that here the op-amp in Fig. 24.21 is used. Note the compressed time scale. The settling time for the low-to-high transition is now 10 ns, while the high-to-low transition time is 60 ns. We would expect much faster settling because the compensation capacitor is now ten times smaller than the value used in the op-amp of Fig. 24.8. Again, the settling time is limited by how fast M8 can discharge the load capacitance and the compensation capacitance. This can be estimated as 10 μA/1.24 pF or 8 mV/ns. It takes 50 ns for the output to transition 400 mV. The size of M8 can be increased to meet a specific load capacitance and settling time requirement.

Because of the improved speed performance, smaller layout area, and (as we'll see) better power supply noise rejection, we'll use indirect compensation in place of Miller compensation in the remaining two-stage op-amps that we discuss in this book.

IPR2023-00698
Theta EX2019

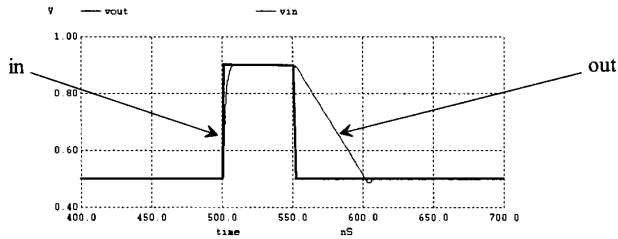


Figure 24.24 The large-signal (500 mV to 900 mV) performance of the op-amp in Fig. 24.21 with C_c = 0.24 pF driving a 1 pF load. The low-to-high settling time is roughly 10 ns, while the high-to-low settling time, which is slew-rate limited, is roughly 60 ns. Note the different time scale when compared to Fig. 24.23.

Common-Mode Rejection Ratio (CMRR)

The *CMRR* of an op-amp is calculated in the same way as the diff-amp in Sec. 22.1.3. The common-mode gain of the diff-amp is A_c. The common-mode gain of the op-amp is $A_c A_2$. The differential gain of the op-amp is $A_{OL}(f) = A_d \cdot A_2$ (where $A_d = A_1$). The *CMRR* of an op-amp in dB is given by

$$CMRR = 20 \cdot \log \left| \frac{A_{OL}(f)}{A_c \cdot A_2} \right| = 20 \cdot \log \left| \frac{A_d}{A_c} \right| \qquad (24.26)$$

which shows that the op-amp *CMRR* is determined by the differential stage. Simulating the *CMRR* of an op-amp can be accomplished with the circuits seen in Fig. 24.25. For the op-amps discussed so far in this chapter, the *CMRR* is approximately 50 dB (see Ex. 22.7, where the common-mode voltage is 500 mV). We can think of A_c as the open-loop gain of the op-amp for common-mode signals. If the applied differential voltage is zero, and we change the common-mode voltage by ΔV_c, then the output voltage will change by ΔV_o

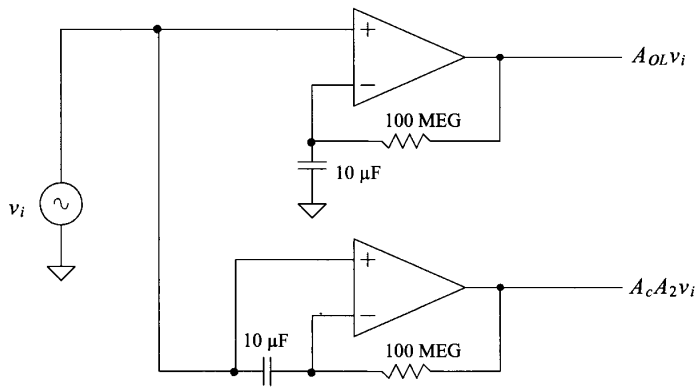


Figure 24.25 Circuit configuration used to simulate CMRR.

IPR2023-00698
Theta EX2019

$= A_{cm} \cdot \Delta V_c$ (see Fig. 22.17). To compensate for the change in the output voltage, a nonzero input differential voltage develops on the input of the op-amp (an offset voltage that is a function of the common-mode voltage). This offset voltage can be estimated by

$$\Delta V_{OS} = \frac{\Delta V_o}{A_{OL}} = \frac{\Delta V_c \cdot A_{cm}}{A_{OL}} = \frac{\Delta V_c}{CMRR} \tag{24.27}$$

Knowing that 50 dB = 316, a change in the common-mode level on the inputs of the op-amp by 500 mV results in an input-referred offset voltage change of 500 mV/316 or 1.6 mV (if the gain of the diff-amp is 10, then its output voltage will change by 16 mV). This may not seem like such a big deal. However, notice that at higher frequencies, in Fig. 22.16, the *CMRR* falls off, indicating that the common-mode gain is increasing. This leads to distortion and forces the use of op-amp topologies in which common-mode voltage doesn't vary (the inverting op-amp topology, see Fig. 24.26).

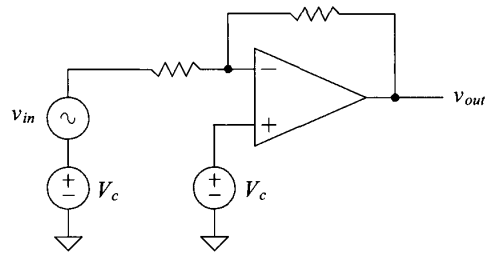


Figure 24.26 An inverting op-amp topology. The common mode voltage is held constant.

Power Supply Rejection Ratio (PSRR)

The power supply rejection ratio (PSRR) is a term used to describe how well an amplifier rejects noise or changes on the *VDD* and or ground power buses. This parameter can be extremely important in precision analog design. Consider the test setup shown in Fig. 24.27. The positive PSRR is defined by

$$PSRR^+ = \frac{A_{OL}(f)}{v_{out}/v^+} \tag{24.28}$$

while the negative PSRR is defined by

$$PSRR^- = \frac{A_{OL}(f)}{v_{out}/v^-} \tag{24.29}$$

Ideally, v_{out} doesn't vary with changes in *VDD* and ground (and so the *PSRR* is infinite).

To understand how variations in *VDD* or ground can feed to the output of the amplifier, consider the op-amp in Fig. 24.8. Variations in *VDD* feed through to the gate of M3. Because of the circuit's symmetry, this causes the drain of M4, and thus the gate of M7, to move around. However, the source of M7 is moving the same amount, causing the v_{sg} of M7 to remain constant. Again, because of the symmetry, the output voltage moves

IPR2023-00698
Theta EX2019

at the same rate. Thus, noise on *VDD* feeds directly to the output of the op-amp. This is indicated in Fig. 24.27c, where v_{out} / v^+ is one. At higher frequencies (in the kHz range), the gate and drain of M7 get shorted together through the compensation capacitor. This, again, causes all of the noise from *VDD* to feed directly to the output of the amplifier (see netlists at cmosedu.com). The indirect compensation scheme used in the op-amp of Fig. 24.21 can be designed for larger f_{un} and, thus, has better *PSRR* at higher frequencies. Using the common-gate stage, Fig. 24.17, can result in an even higher *PSRR* by isolating the compensation capacitor from the power supplies. For this reason op-amp's that use a cascode-load diff-amp, Fig. 24.29, are very useful in practical design.

Noise on ground ideally won't affect the operation of the op-amp (in either of the topologies of Fig. 24.8 or 24.21). The noise appears as a voltage variation across the current sources. In reality, at low frequencies, some of the ground noise, v^-, appears at the output (this is especially true in nanometer processes that have low r_o). However, at higher frequencies, because of the compensation capacitance connected to the output of the op-amp, ground noise contributions to the output signal decrease.

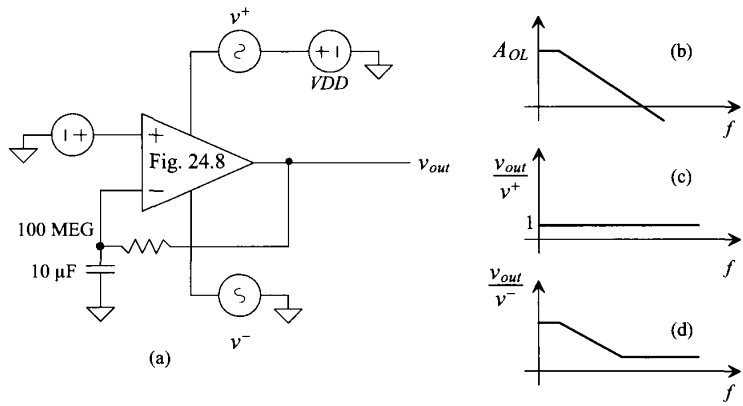


Figure 24.27 (a) Test setup to determine PSRR, (b) open-loop gain, (c) gain from AC signal on VDD to output, and (d) gain from AC signal on ground to output.

Increasing the Input Common-Mode Voltage Range

It may be useful in some situations to have an input common-mode voltage range that extends close to the power supply rails. For our op-amp in Fig. 24.21, the allowable input common-mode voltage range is from (roughly) 450 mV to 900 mV (basically half of *VDD*). Towards the goal of extending the input common-mode voltage range, consider the addition of a PMOS diff-amp to Fig. 24.21, as seen in Fig. 24.28. The circuitry on the right side of the schematic is the op-amp in Fig. 24.21 drawn a little differently. Here, we've drawn the current source of the diff-amp as two parallel current sources. In Fig. 24.21 the current source biasing the NMOS diff-amp used MOSFETs with 100/2 sizes. Here we supply the same current but use two 50/2-sized current sources. Also, we've bumped up the size of the PMOS devices (M3, M4, and M7) by two. This is to accommodate the extra current supplied by the added PMOS diff-amp (the left side of the schematic). We could keep the PMOS devices the same size as seen in Fig. 24.21 and the

IPR2023-00698
Theta EX2019

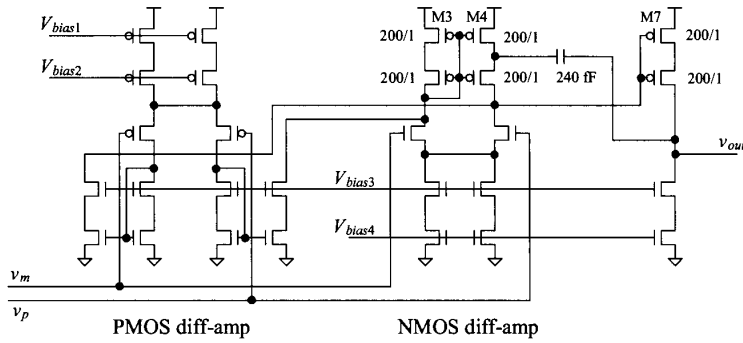


Figure 24.28 Two-stage op-amp of Fig. 24.21 with rail-to-rail input range.

circuit would still work fine (the V_{SG} voltages of the PMOS would simply be a little larger when both diff-amps are on). Two other notes: notice that the DC currents flowing in M3, M4, and M7 change depending on the input common-mode voltage. When the input common-mode voltage is large, the PMOS diff-pair is off; when it's small, the NMOS diff-pair is off. In the middle, both are conducting current. This means that the transconductance of the diff-amp will vary from g_{mn} to $g_{mn} + g_{mp}$ to g_{mp}. When both diff-amps are on (assuming $g_{mn} = g_{mp}$), the transconductance is twice as high as when a single diff-amp is on. This may require doubling the compensation capacitor. Finally, the change in the DC biasing conditions result in an offset that is a function of the input common-mode voltage.

Estimating Bandwidth in Op-Amp Circuits

When we finish designing an op-amp, its frequency response can be written using a single dominant-pole response (see Fig. 21.26 on page 680 with $A_{DC} = A_{OLDC}$) as

$$A_{OL}(f) = \frac{A_{OLDC}}{1 + j \cdot \frac{f}{f_{3dB}}} \tag{24.30}$$

For the op-amp response in Fig. 24.19, we can write

$$A_{OL}(f) = \frac{1{,}000}{1 + j\frac{f}{10\,kHz}} \tag{24.31}$$

To estimate the bandwidth of a closed-loop op-amp circuit where $f >> f_{3dB}$, we can write Eq. (24.30) as

$$|A_{OL}(f)| \approx \frac{A_{OLDC}}{\frac{f}{f_{3dB}}} = \frac{A_{OLDC} \cdot f_{3dB}}{f} = \frac{f_{un}}{f} \tag{24.32}$$

noting the unity-gain frequency is calculated using

$$f_{un} = A_{OLDC} \cdot f_{3dB} \tag{24.33}$$

IPR2023-00698
Theta EX2019

From this equation, we should see why the unity-gain frequency, f_{un}, is called the *gain-bandwidth product.* Note that for every increase in frequency by 10 (decade) above f_{3dB}, we get a decrease in A_{OL} by 10 (–20 dB). Alternatively, for every increase in frequency by 2 (octave) above f_{3dB}, we get a decrease in A_{OL} by 2 (– 6 dB).

The closed-loop bandwidth (which we'll call f_{3dBCL}) of the op-amp circuit cannot be larger than the bandwidth of the op-amp. Therefore, to estimate the bandwidth of a closed-loop op-amp circuit, assuming the op-amp is the limiting factor, we can write

$$A_{CL} \cdot f_{3dBCL} = f_{un} = \text{gain-bandwidth product} \qquad (24.34)$$

If the op-amp circuit has a closed-loop gain of one (a follower configuration as seen in Fig. (24.7)), the bandwidth is f_{un}. The op-amp in Fig. 24.21 used in a follower configuration would have a bandwidth of (roughly) 100 MHz. If this op-amp were used in a closed-loop configuration with a gain of 10, then the bandwidth of the op-amp circuit would be estimated as 10 MHz.

24.2 An Op-Amp with Output Buffer

In the previous section, the low-frequency open loop gain of the op-amps in Figs. 24.8 or 24.21 was determined by a product of the diff-amp's gain, A_1, and the common-source's gain, A_2 or

$$A_{OLDC} = A_1 \cdot A_2 = \overbrace{g_{mn}(r_{o2} \| r_{o4})}^{A_1} \cdot \overbrace{g_{mp} r_{o7}}^{A_2} \qquad (24.35)$$

which, from the simulations (see Fig. 24.19 for example) was 1,000. The small-signal resistance on the output of the op-amp, r_{o7}, is (from Table 9.2) 333 kΩ. If we were to connect a 10 kΩ resistor to the op-amp output, then A_{OLDC} would drop to 33. As seen in Fig. 24.1, we can add a buffer to the output of second stage to isolate it from a load resistance (or a large capacitance). We could use a source-follower (that has a voltage gain of 1), as seen in Figs. 21.46 or 21.48. However, this addition limits the output swing of the op-amp (which may be OK in some situations). Next we could try using a push-pull topology, as seen in Fig. 21.49. However, the voltage gain of the push-pull amplifier is >1 (see Eq. (21.116)) so if we were to include it in our basic op-amp of Fig. 24.21, we would have a three-stage op-amp (meaning all three stages have voltage gains greater than one). A three-stage op-amp can be challenging to compensate over production corners (and temperature). Also, as seen in Eq. (21.116), a small load resistor (kΩ) would still kill the push-pull amplifier's gain. To keep the gain high, let's try to use a first-stage topology with a large gain (increase A_1) while using a push-pull amplifier for the second stage. The gain of the second stage can still drop but, since the gain of the first stage is large, we still end up with a reasonable overall gain.

Towards the goal of increasing the first-stage gain, consider the op-amp schematic seen in Fig. 24.29. For the first stage, we've used the cascode diff-amp originally seen in Fig. 22.30. The gain of this diff-amp (now A_1) is 500 (from Tables 9.2, and 20.1 and Eq. (22.48)). The second stage of the op-amp is the topology seen in Fig. 21.50. Its gain depends on the load resistance it drives. If no load is connected to its output, then

$$A_2 = -10 \cdot (g_{mn} + g_{mp}) \cdot \frac{(r_{op} \| r_{on})}{10} = -31.6 \; V/V \qquad (24.36)$$

IPR2023-00698
Theta EX2019

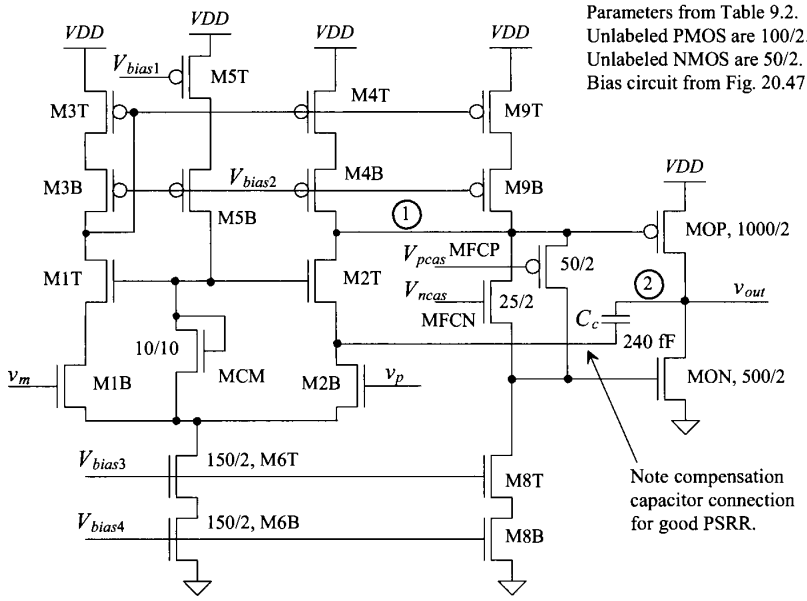


Figure 24.29 A CMOS op-amp with output buffer.

giving an overall low-frequency gain, A_{OLDC}, of 15,800 (= 84 dB). Note that the widths of MON and MOP are ten times wider than the values specified in Table 9.2 and so, as seen in Fig. 21.50, they conduct 100 μA of current. In Eq. (24.36) we multiplied the g_ms in Table 9.2 by 10 ($g_{mon} = 10 \cdot \beta_n(V_{GS} - V_{THN}) = 10 \cdot g_{mn,Table\ 9.2}$) and divided the output resistances given in Table 9.2 by 10 $\left(r_{omon} = \frac{1}{\lambda 10 \cdot 10\mu A} = \frac{r_{on}}{10}\right)$.

Compensating the Op-Amp

We increased the gain of the first stage by increasing R_1 (the resistance at node 1 is now set by cascoded current sources). As seen in Eq. (21.65), this pushes the pole lower in frequency. As just described in Eq. (24.33), this decrease in f_1 (= f_{3dB}) and increase in low-frequency gain *doesn't have any effect* on the gain-bandwidth product (= f_{un}) of the op-amp. What this indicates, to a certain extent, is that we can still compensate the op-amp with a 240 fF capacitor as we did in the last section. To feed the indirect compensation current to node 1 we connect the compensation capacitor to the source of M2T, as seen in Fig. 24.29. This ensures good PSRR.

The open-loop response of the op-amp in Fig. 24.29 (without a load) is seen in Fig. 24.30. The phase-margin is roughly 70 degrees. The step response of the op-amp driving a 1k resistor and a 10 pF capacitor is seen in Fig. 24.31. The input signal is a pulse from 100 mV to 900 mV. Since the gain of the amplifier is −1 and the common-mode voltage is 500 mV, the output swings from 900 mV to 100 mV.

IPR2023-00698
Theta EX2019

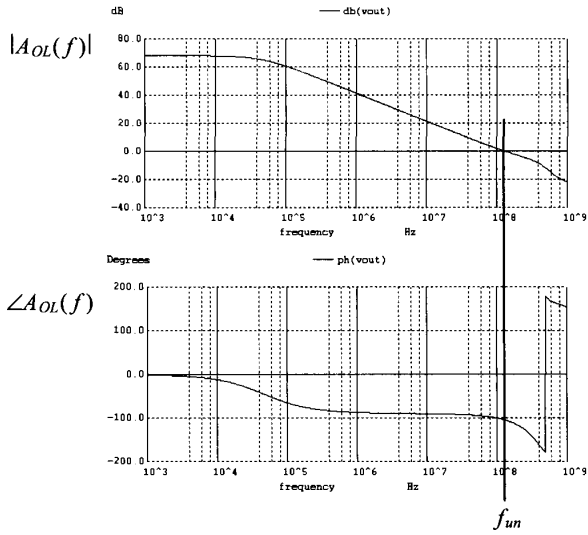


Figure 24.30 Open-loop response of the op-amp in Fig. 24.29.

This is a good point to remember one of the fundamentals we presented in Ch. 9: that is, for high-speed design, we must use minimum length devices (see Eq. (9.59)). If we reduce all of the length of 2 (100 nm) devices in Fig. 24.29 to length of 1 (50 nm) devices, we can increase the speed of the op-amp. Figure 24.32 shows how the step response in Fig. 24.31 changes when we make this reduction in length. The big problem with using minimum channel lengths is the reduction in gain. Close inspection of Fig. 24.32 shows that the output is never quite pulled up to the correct voltage (it is either slightly below 900 mV or 100 mV). Simulating the AC open-loop response results in a gain of 44 dB (= 159). An output voltage of 900 mV would result in a difference on the inputs of the op-amp of 5.66 mV (0.9/159 = $v_p - v_m$).

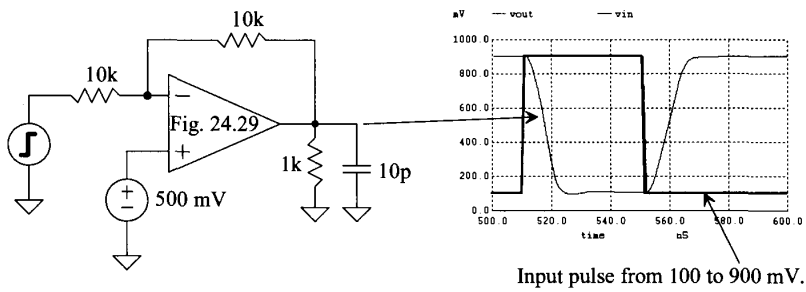


Figure 24.31 Step response of the op-amp in Fig. 24.29 driving 1k and 10 pF.

IPR2023-00698
Theta EX2019

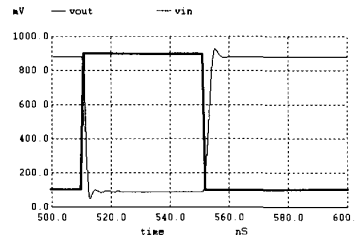


Figure 24.32 Step response of the amplifier in Fig. 24.29 when the lengths of the devices are reduced from 2 to 1 (from 100 nm to 50 nm). The load is still, as seen in Fig. 24.31, 1k and 10 pF.

24.3 The Operational Transconductance Amplifier (OTA)

The operational transconductance amplifier (OTA) can be defined as an amplifier where all nodes are low impedance except the input and output nodes. A simple example of an OTA is the diff-amp with current mirror load seen in Fig. 22.6. *An OTA without buffer can only drive capacitive loads.* A resistive load (unless the resistor is very large) will kill the gain of the OTA.

A sample OTA is shown in Fig. 24.33. Note that the basic op-amp in Fig. 24.2 is not considered an OTA because the drain of M4 is a high-impedance node and not the input or output of the amplifier. As seen in Fig. 24.33, except for the input and output nodes, all nodes have either a gate-drain-connected device or a source connected to them. The terminology "1:K" indicates that M4 and M5 can be sized K times wider (K > 1) than the other MOSFETs in the circuit. Assuming that $\beta_1 = \beta_2$, $\beta_{31} = \beta_{41}$, we observe that the current i_{d31} or i_{d41} is given by

$$-i_{d31} = i_{d41} = \frac{g_{mn}}{2}(v_p - v_m) = i_d \tag{24.37}$$

Furthermore, if $\beta_4 = K \cdot \beta_{41} = K \cdot \beta_{31} = K \cdot \beta_3$ and $K \cdot \beta_{51} = \beta_5$, then $i_{d4} = -i_{d5} = K \cdot i_{d41} = -K \cdot i_{d31}$. If the impedance of the capacitor is large compared to $r_{o4} \| r_{o5}$, then the output voltage of the OTA is given by

$$v_{out} = 2Ki_d(r_{o4} \| r_{o5}) \tag{24.38}$$

and the voltage gain is given by

$$A_v = \frac{v_{out}}{v_p - v_m} = K \cdot g_m \cdot (r_{o4} \| r_{o5}) \tag{24.39}$$

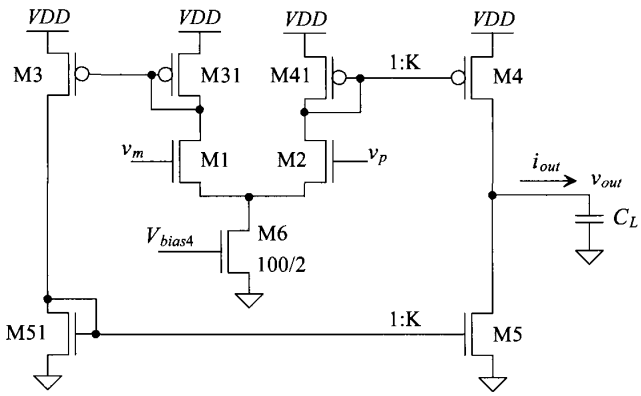
noting that the noninverting input of the OTA is the gate of M2. While the derivation is interesting, as the name implies, we are interested in the transconductance of OTA. If the impedance of the capacitor is small compared to the OTA output resistance (at higher frequencies), we can write

$$i_{out} = i_{d4} - i_{d5} = 2Ki_d \tag{24.40}$$

The transconductance of the OTA is given by

$$g_{mOTA} = \frac{i_{out}}{v_p - v_m} = K \cdot g_m \tag{24.41}$$

IPR2023-00698
Theta EX2019



Unless otherwise indicated, parameters from Table 9.2 with biasing circuit in Fig. 20.47.

Figure 24.33 Example of an operational transconductance amplifier (OTA).

Unity-Gain Frequency, f_{un}

To simulate the unity-gain frequency, f_{un}, (where the open loop gain is one) for the OTA, we can use the DC stability scheme seen in Fig. 24.34. We can write (for higher frequencies) the output voltage as

$$v_{out} = i_{out} \cdot \frac{1}{j\omega \cdot C_L} = g_{mn} v_{in} \cdot \frac{1}{j\omega \cdot C_L} \tag{24.42}$$

or

$$\frac{v_{out}}{v_{in}} = \frac{g_{mn}}{j\omega C_L} \rightarrow \left|\frac{v_{out}}{v_{in}}\right| = \frac{g_{mn}}{2\pi f \cdot C_L} \tag{24.43}$$

As before, we define the unity-gain frequency as the point where the open-loop gain is unity

$$\frac{g_{mn}}{2\pi f_{un} \cdot C_L} = 1 \rightarrow f_{un} = \frac{g_{mn}}{2\pi C_L} \tag{24.44}$$

For the OTA in Fig. 24.33 driving a 1 pF load capacitance, we can estimate the unity-gain frequency (with $K = 1$) as

$$f_{un} = \frac{150\ \mu A/V}{2\pi \cdot 1\,pF} = 24\ MHz$$

The simulation results are seen in Fig. 24.34. The location of the pole is estimated using

$$f_{3dB} = \frac{1}{2\pi(r_{o4} \| r_{o5})C_L} \tag{24.45}$$

For the OTA in Fig. 24.33, f_{3dB} = 1.4 MHz using the values in Table 9.2. The low-frequency gain is $g_m(r_{o4} \| r_{o5})$ = 16.65 V/V (= 24.4 dB).

IPR2023-00698
Theta EX2019

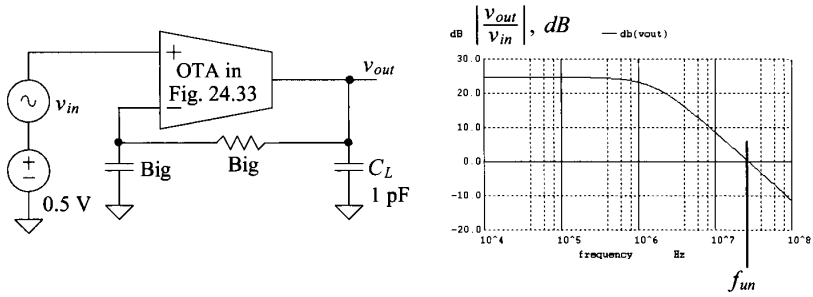


Figure 24.34 Simulating the open-loop response of the OTA in Fig. 24.33.

Note that the maximum value of i_{out} is KI_{SS} (the drain current of M6 multiplied by the increase in widths, K, in the current mirrors M41/M4 and M51/M5). This is important to understand because using an OTA for the first stage of an op-amp results in the same slew-rate limitations of I_{SS} driving C_c that we had when using a diff-amp.

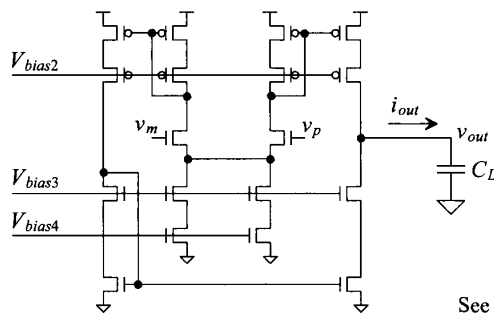
Increasing the OTA Output Resistance

The ideal OTA has infinite output resistance. All of i_{out} flows in the external capacitive load and none flows in the OTA's own output resistance (which is $r_{o4}||r_{o5}$ for the OTA in Fig. 24.33). Towards increasing the OTA output resistance, consider the configuration seen in Fig. 24.35. This topology is based on the one seen in Fig. 24.33 except that now we've cascoded the current mirrors. Note how we've drawn the diff-amp's tail current source as four MOSFETs instead of two with twice the width (this is the way we would lay it out too). The low-frequency gain of the OTA is now increased to

$$A_v = g_{mn} \cdot (R_{ocasn}||R_{ocasp}) \tag{24.46}$$

The 3-dB frequency is now

$$f_{3dB} = \frac{1}{2\pi(R_{ocasn}||R_{ocasp}) \cdot C_L} \tag{24.47}$$



See Table 9.2 and Fig. 20.47.

Figure 24.35 A cascode OTA circuit (higher output resistance).

IPR2023-00698
Theta EX2019

As discussed before (Sec. 24.2 in the discussion concerning compensating the op-amp), when we get a decrease in the 3-dB frequency and an increase in the low-frequency gain, the effects cancel and the gain-bandwidth product (f_{un}) remains constant. The unity-gain frequency is still given by Eq. (24.44). For the OTA in Fig. 24.35, we can estimate the low-frequency gain, using Tables 9.2 and 20.1, as

$$A_v = (150) \cdot (16.6||4.2) = 500\ (54\ dB)$$

and the 3-dB frequency, driving a 1 pF load, is

$$f_{3dB} = \frac{1}{2\pi(16.6||4.2 \times 10^6)(1\ pF)} = 47\ kHz$$

Simulation results are seen in Fig. 24.36. The PM is roughly 85°.

An Important Note

Note, in Eq. (24.44), that increasing the load capacitance, decreases the unity-gain frequency, making the OTA more stable. Unlike the two-stage op-amps discussed in the first two sections of this chapter, where the op-amp can become unstable with a large load capacitance, the OTA only becomes more stable with large load capacitance. Also, the PM approaches 90° as the load capacitance increases. The step response of the OTA has a first-order shape (just as in an RC circuit). In many on-chip applications, the load is purely capacitive and so the OTA works great as an "op-amp" in a closed-loop configuration (thus we'll often call OTAs op-amps). Note that adding a second stage to the basic OTA forms a (two-stage) op-amp like those discussed earlier. We can use the methods already presented to compensate the resulting structure (but, again, driving a load capacitance that's too large with these two-stage structures can result in instability).

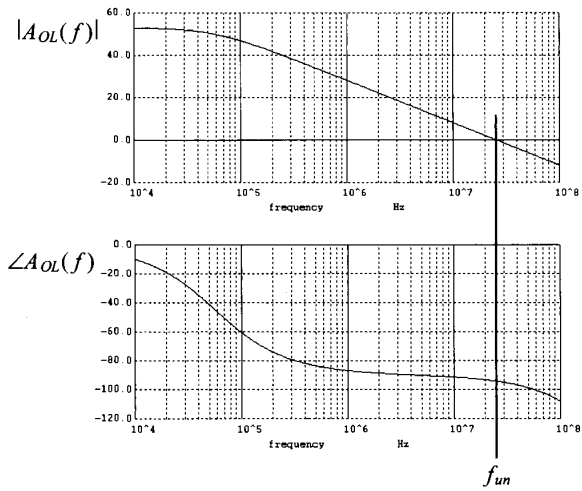


Figure 24.36 The open-loop gain and phase response for the OTA in Fig. 24.35.

IPR2023-00698
Theta EX2019

OTA with an Output Buffer (An Op-Amp)

If we use the OTA in Fig. 24.35 (or Fig. 24.33) to drive a 1 pF load, we're limited to charging/discharging the capacitor at a rate of

$$\frac{I_{SS}}{C_L} = \frac{dv_{out}}{dt} = \frac{20\ \mu A}{1\ pF} = \frac{20\ mV}{ns} \tag{24.48}$$

For the OTA's output to change from ground to *VDD* (1 V) requires (just due to slewing) 50 ns. This length of time (the slew-rate limited portion of the settling time) may be fine for many applications. However, for high-speed design, we want faster settling times. Looking at Eq. (24.48), we see that we can decrease the settling time by increasing I_{SS}. However, as discussed in Ch. 9, this causes the gain of the circuit to decrease. To avoid the reduction in gain, we can make a corresponding increase in the size of the devices to keep the overdrive voltage constant. But this is doing nothing more than effectively reducing the size of the load capacitor (the load capacitance becomes comparable in size to the parasitic transistor capacitances).

To speed up the circuit, we may add an output stage, Fig. 24.37. The added stage is a common-source amplifier (with a negative gain indicating **we swap the inverting and noninverting terminals** of the op-amp or else we waste a lot of time). While M7 can turn on and charge a load capacitor quickly, the current source, M8, still limits the rate of load capacitance discharge. For this reason, we've bumped up the current in the output stage to 100 μA. The open-loop simulations driving a 1-pF load are seen in Fig. 24.38. The PM is 45° (too low if the op-amp will be used in a unity follower configuration). Figure 24.39 shows the step response of the op-amp.

We should compare this figure to the response we saw in Fig. 24.31. The op-amp in Fig. 24.31 is driving a heavier load but still has cleaner settling behavior. Further, a comparison of the current pulled from *VDD* reveals, under the same loading and operating conditions, that the op-amp in Fig. 24.37 pulls more current (see Fig. 24.39). The main difference in these two op-amps is the output buffer used. Let's consider adding a class AB buffer to the OTA in Fig. 24.35 (like the one used in the op-amp of Fig. 24.29). Since we are only driving a 1 pF load, we'll leave the output devices, MON and

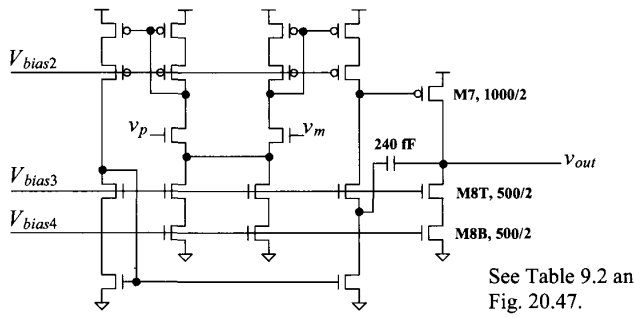


Figure 24.37 A cascode OTA circuit with common-source output buffer (an op-amp). Note the inverting and noninverting terminals are swapped from Fig. 24.35.

IPR2023-00698
Theta EX2019

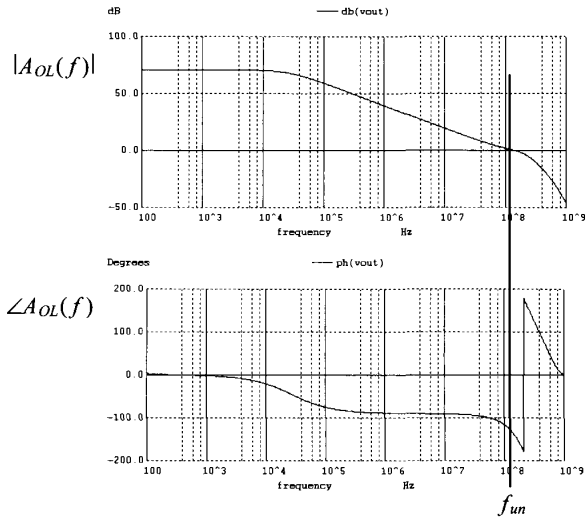
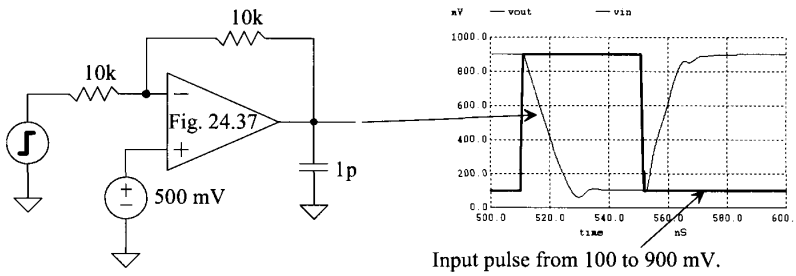
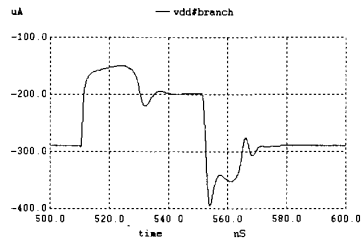


Figure 24.38 Open-loop response of the op-amp in Fig. 24.37 driving a 1-pF load.





Current from VDD. Includes the bias circuit's (Fig. 20.47) current (140 uA).

Figure 24.39 Step response of the op-amp in Fig. 24.37 driving a 1 pF.

IPR2023-00698
Theta EX2019

MOP, at the same size as the other MOSFETs in the circuit (PMOS are 100/2 and NMOS are 50/2). The resulting op-amp is seen in Fig. 24.40. The step response (in the same configuration used to generate Fig. 24.39) is seen in Fig. 24.41. Notice that the current pulled from *VDD* by the op-amp in Fig. 24.40 is 100 μA less than the current pulled by the op-amp in Fig. 24.37. This lower current is the result of using a class AB output stage.

The performance of the op-amps in Figs. 24.37 and 24.40, with regard to high-speed operation and settling time, won't be quite as good as the topology in Fig. 24.29 (assuming the same biasing conditions, device sizes, and overdrive voltages). To understand why, look at the path, in Fig. 24.40, that the noninverting op-amp input, v_p, takes to get to the op-amp output, v_{out} (that is, M1 - M31 - M3 - M51 - M5 - MON). The delay through this path is longer than the delay, in Fig. 24.29, through M2 - MOP. Further, for high speed, M5 of the op-amp in Fig. 24.40 must turn on (above its quiescent current level) to pull the gate of MON down. For the op-amp in Fig. 24.29, the fact that both the gate of M9T and MOP are driven (in opposite directions) by the diff-amp allows us to use a simple pull-down current (M8) without losing speed. The issues with the cascode diff-amp-based, op-amp topology of Fig. 24.29 are the more limited input common-mode range and the limited voltage swing of node 1. Towards getting better input common-mode voltage range and output swing as well as improved high-speed operation, let's discuss the folded-cascode OTA (see Fig. 22.5).

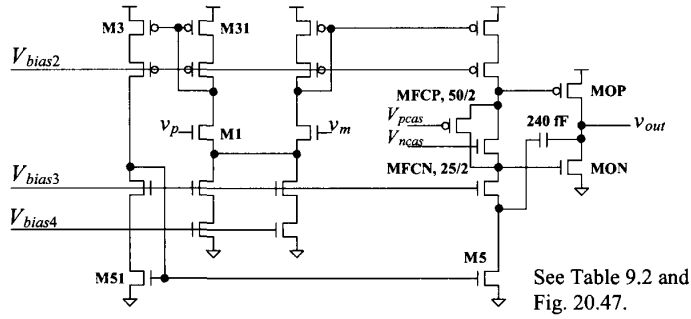


Figure 24.40 Using a class AB output stage with the OTA.

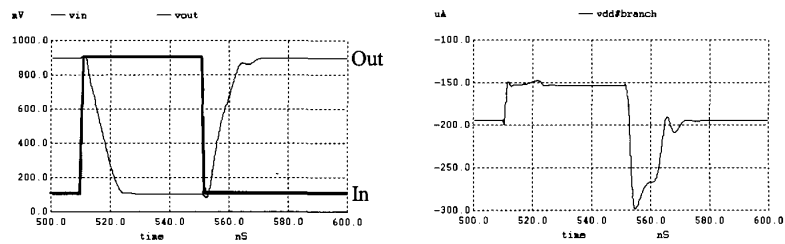


Figure 24.41 The step response and current in the topologies seen in Fig. 24.39 using the op-amp in Fig. 24.40.

IPR2023-00698
Theta EX2019

The Folded-Cascode OTA and Op-Amp

An example of an NMOS diff-amp-based folded cascode OTA is seen in Fig. 24.42. It is assumed that the reader understands the biasing and operation of the circuits seen in Figs. 20.45, 20.46, and 22.5. To avoid labeling MOSFETs with twice the width, we've drawn the schematic in Fig. 24.42 with MOSFETs in parallel in the branches that supply or sink more current (for example, the diff-amp's tail current). All MOSFETs in Fig. 24.42 conduct 10 μA of current under equilibrium conditions. Before moving on, it may be a good idea to pause and make sure that the reader understands how the currents sum in this OTA.

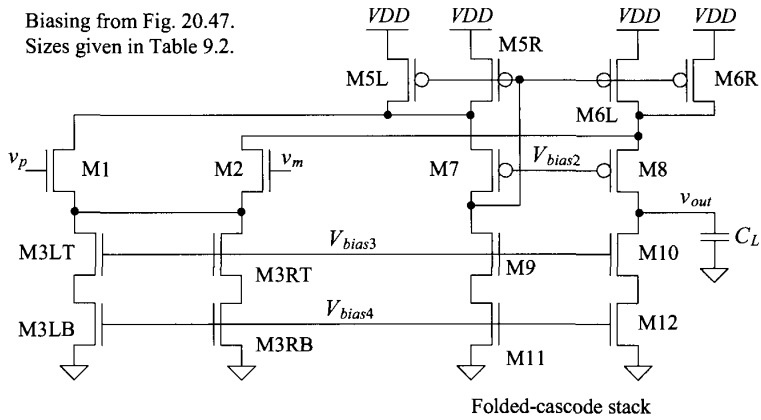


Figure 24.42 A folded-cascode OTA.

To qualitatively describe the operation of the OTA, let's look at what happens when v_p goes high (above v_m). This causes M1 to turn on and M2 to shut off. When M1 turns on, it pulls the drain of M5 down, shutting M7 off. As M7 shuts off, M9 and M11 pull the gate of M5 down. With the gate of M5 being pulled down, the current in M6 increases. At the same time, because the current in M2 is decreasing, the current in M8 is increasing and the output voltage increases. The maximum current the OTA output can supply (out of the OTA to the load) is in the neighborhood of 20 μA, while the maximum current the OTA can sink (into the OTA) is 10 μA (set by the current sink M10/M11).

The same AC equations presented at the beginning of this section also apply to this OTA. For example, the AC drain currents of M1 and M2 can be written as

$$i_{d1} = -i_{d2} = g_{mn} v_{gs1} = -g_{mn} v_{gs2} \tag{24.49}$$

which, again, indicates $v_{gs1} = -v_{gs2}$. However, we know

$$v_p - v_m = v_{gs1} + (-v_{gs2}) \tag{24.50}$$

so

$$i_{d1} = i_{d5} = i_{d6} = (v_p - v_m) \cdot \frac{g_m}{2} \tag{24.51}$$

IPR2023-00698
Theta EX2019

noting that the AC current through M9/M11 (and thus M7) is ideally zero. The current through M8 is then

$$i_{d8} = i_{d6} - i_{d2} = 2 \cdot i_{d1} \tag{24.52}$$

and thus the output voltage (again, the AC current in M10/M11 is ideally zero) is

$$v_{out} = i_{d8} \cdot (R_{ocasn} || R_{ocasp}) \tag{24.53}$$

and thus the gain is

$$A_v = \frac{v_{out}}{v_p - v_m} = g_{mn} \cdot (R_{ocasn} || R_{ocasp}) \tag{24.54}$$

This result should be compared to Eq. (24.46). Equation (24.47) gives the 3-dB frequency of the OTA's open-loop response, while Eq. (24.44) gives the unity-gain frequency. Figure 24.43 shows the open-loop response of the folded-cascode OTA driving a 1 pF load capacitance.

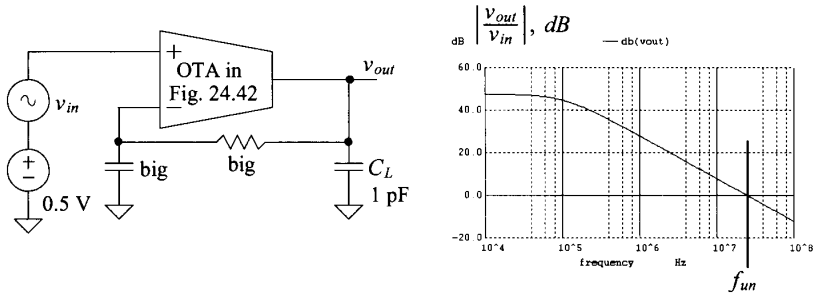


Figure 24.43 Simulating the open-loop response of the OTA in Fig. 24.42.

Figure 24.44 shows the schematic of a folded-cascode op-amp using a class AB output buffer. We've added floating current sources to equalize the voltages across M9/M10 (to minimize the input-referred offset). In the open-loop response of the op-amp (see Fig. 24.45), the load of the op-amp is a 1 pF capacitor. The step response is seen in Fig. 24.46, using the topology of Fig. 24.39 (gain of –1 driving 1 pF and the 10k feedback resistor). Reviewing the previous op-amp responses, we don't see much difference in performance. We can't seem to push the gain-bandwidth product, f_{un}, much above 100 MHz. Again, the only way to increase the speed is to reduce the channel length and/or increase the overdrive voltage.

Figure 24.47 shows what happens if we reduce the lengths of the MOSFETs used in the op-amp in Fig. 24.44 from 2 (100 nm) to 1 (50 nm). The gain-bandwidth product jumps to 400 MHz. However, the low-frequency gain drops to approximately 50 dB. The step response is clean, and the settling time is approximately 5 ns. While we can increase the speed further by redesigning the bias circuit and increasing the currents and overdrive voltage, let's, instead, concentrate on increasing the low-frequency gain of the op-amp. This won't increase the gain-bandwidth product of the op-amp, but it's necessary for precision amplification. To *enhance the gain,* we'll try to increase the OTA output resistance (R_1 or the resistance on node 1) by regulating the drain in the cascode current sources. We've already discussed this technique (see Fig. 20.40), but we haven't applied it to amplifiers. Before leaving this section, let's present one more example.

IPR2023-00698
Theta EX2019

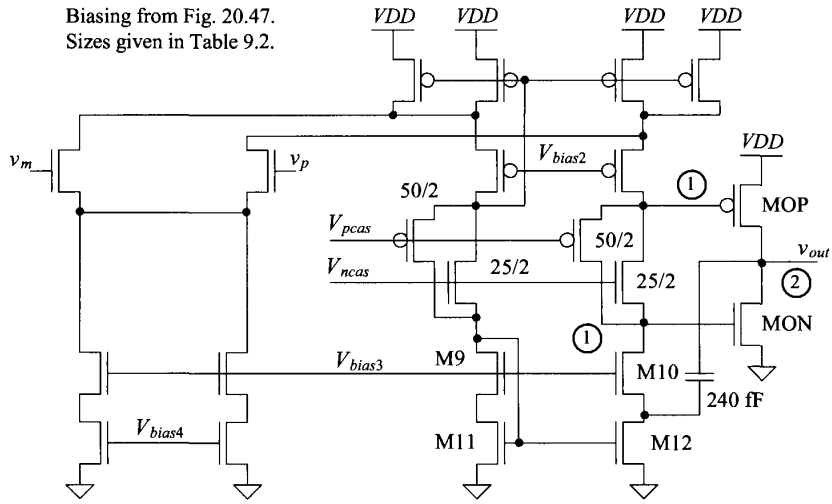


Figure 24.44 Folded-cascode op-amp with class AB output buffer.

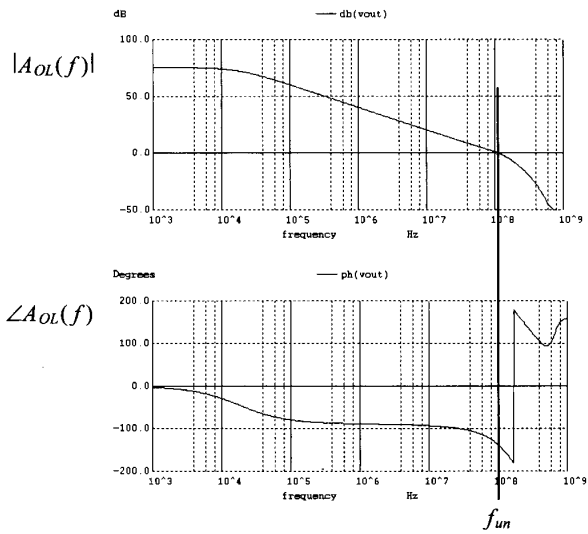
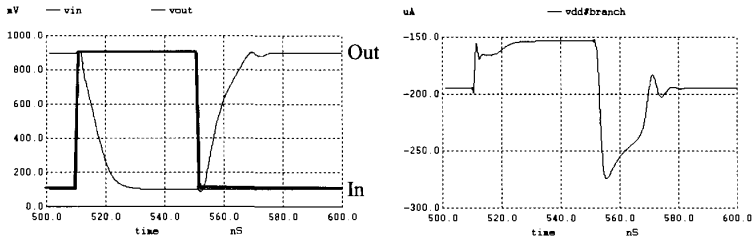


Figure 24.45 Open-loop response of the op-amp in Fig. 24.44 driving a 1-pF load.

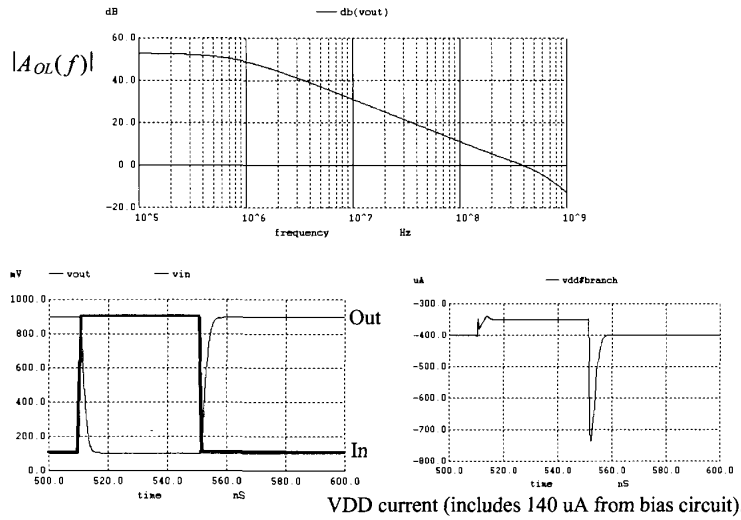
IPR2023-00698
Theta EX2019



VDD current (includes 140 uA from bias circuit).

Figure 24.46 The step response and current in the topologies seen in Fig. 24.39 using the op-amp in Fig. 24.44.

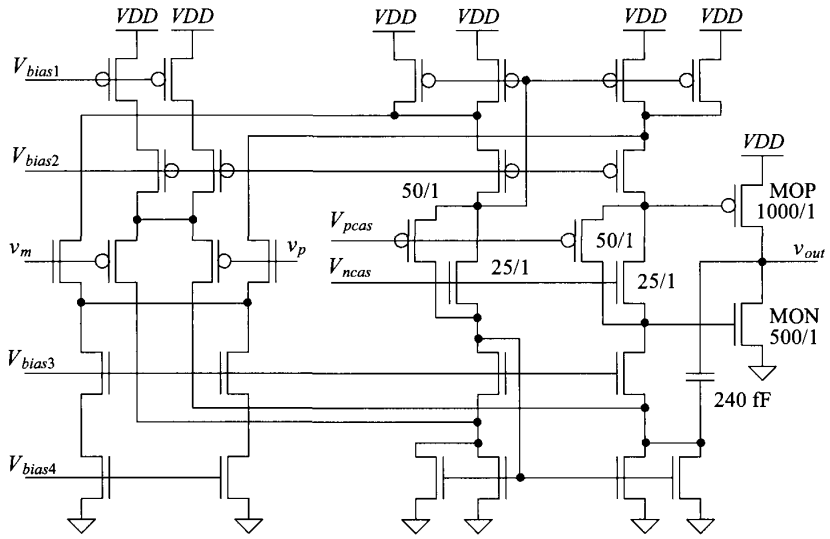
Figure 24.48 shows a wide-swing op-amp based on the topology seen in Fig. 24.44. To get wide-swing operation (which means that the input common-mode voltage extends beyond the power supply rails and the op-amp still functions), we added a PMOS diff-amp stage. Further, to sink the additional current supplied by the PMOS diff-amp when it's on, we add two NMOS devices at the bottom of the folded-cascode stack. When, for example, the PMOS diff-amp shuts off with the input common-mode voltage moving towards *VDD,* the added two devices don't really affect the circuit's operation. The only result is that the V_{GS} of the bottom NMOS devices is slightly smaller when the PMOS diff-amp shuts off. A similar conclusion can be drawn when the NMOS diff-amp shuts off with the common-mode voltage moving towards ground. Note that when both diff-amps are on, the first-stage transconductance becomes, $g_{mn} + g_{mp}$. This means the gain-bandwidth product, f_{un} , moves to $(g_{mn} + g_{mp})/2\pi C_c$. The value of the compensation



VDD current (includes 140 uA from bias circuit).

Figure 24.47 Reducing the length of the MOSFETs used in the op-amp in Fig. 24.44 from 2 to 1 and resimulating open-loop and step responses.

IPR2023-00698
Theta EX2019



Biasing from Fig. 20.47.
Unlabeled NMOS are 50/1.
Unlabeled PMOS are 100/1.

Figure 24.48 An op-amp with an input common-mode range that extends beyond the power supply rails and that can drive heavy loads.

capacitor may need to be increased. When only one diff-amp is on, the transconductance of the other diff-amp goes to zero, resulting in a smaller gain-bandwidth product. The issue here is that the gain of the op-amp is changing with variations in the input common-mode voltage. This, as mentioned before, can lead to distortion and is the reason that the inverting topology, as seen in Fig. 24.39, is preferred (noting that the input common-mode voltage is fixed when using the inverting topology).

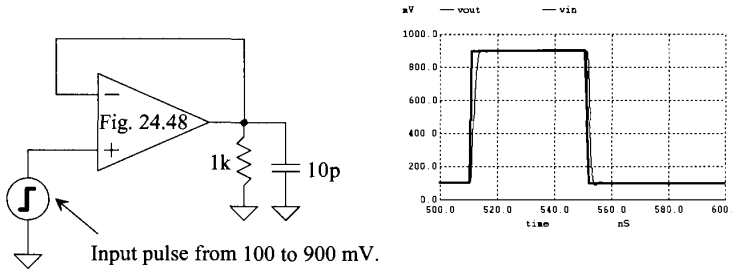


Figure 24.49 Step response of the op-amp in Fig. 24.48 driving 1k and 10 pF. Note that the op-amp is in the follower configuration and that the input common-mode voltage is changing.

IPR2023-00698
Theta EX2019

Finally, in some applications, an op-amp that can supply significant amounts of current, say 100 mA, may be needed. To get such a large current, we may need to increase the widths of the output MOSFETs in the push-pull amplifier to, for example, 10,000 and 5,000 (for the PMOS and NMOS devices, respectively). The problem with this is that the quiescent current that flows in these devices, using the op-amp configuration in Fig. 24.44 as an example, is 1 mA. To reduce this current, we can increase the lengths of the floating current sources used in the folded-cascode section. This increases their gate-source voltage drops and moves the gates of the push-pull MOSFETs towards the power supply rails (shutting them off). Care must be exercised when doing this because shutting off the output MOSFETs can result in moving g_{m2} towards zero, causing, as indicated in Eq. (24.24), f_2 to move down in frequency, thus making the op-amp unstable.

24.4 Gain-Enhancement

While using minimum channel lengths (L = 1) results in the fastest op-amps, the open-loop gain, A_{OLDC}, is considerably reduced. Using gain-enhancement (GE) techniques (regulating the drain node in a cascode structure, see Fig. 20.40), we boost the low-frequency gain by increasing the cascode output resistance. However, the gain-bandwidth product, f_{un}, is still $g_m/2\pi C_c$. (We get the speed by using a smaller compensation capacitor and/or a larger diff-amp transconductance.) Note, in Fig. 24.47 for example, that our gain-bandwidth product went up by the increase in g_m with the reduction in the channel length. Reviewing Eqs. (24.30)–(24.34), we see that if A_{OLDC} goes up, f_{3dB} goes down (and f_{un} is a constant).

In the following discussion, we only concern ourselves with two-stage op-amps in which we are using GE to compensate for the low gain caused by our minimum channel lengths. While GE can be used in an OTA, the fact that the speed of the OTA is limited by the capacitance it is driving keeps them from being as fast as the two-stage op-amps we've discussed (unless, of course, the load capacitance is small).

As seen in Fig. 20.40, we need an additional amplifier to help regulate the drain of the cascode device. Consider the two differential amplifiers (OTAs) seen in Fig. 24.50. The source-followers on the inputs of the diff-amps are used to allow for the amplification of signals near ground (for the P amplifier) or *VDD* (for the N amplifier). The gains of the diff-amps are

$$A_P = g_{mp} \cdot (r_{on} || r_{op}) \text{ and } A_N = g_{mn} \cdot (r_{on} || r_{op}) \tag{24.55}$$

To simplify the equations, we'll assume

$$A_{GE} = A_P = A_N \tag{24.56}$$

The low-frequency gain of the op-amp in Fig. 24.44 is

$$A_{OLDC} = \overbrace{g_{mn} \cdot (R_{ocasn} || R_{ocasp})}^{\text{diff-amp gain}} \cdot \overbrace{(g_{mp} + g_{mn}) \cdot (r_{on} || r_{op})}^{\text{push-pull output stage}} \tag{24.57}$$

Using Eq. (20.67), we can write the open-loop gain *with gain-enhancement* as

$$A_{OLDC,GE} = A_{OLDC} \cdot A_{GE} \tag{24.58}$$

The open-loop gain is increased by the gain of the added amplifier. Figure 24.51 shows the op-amp of Fig. 24.44 with GE.

IPR2023-00698
Theta EX2019

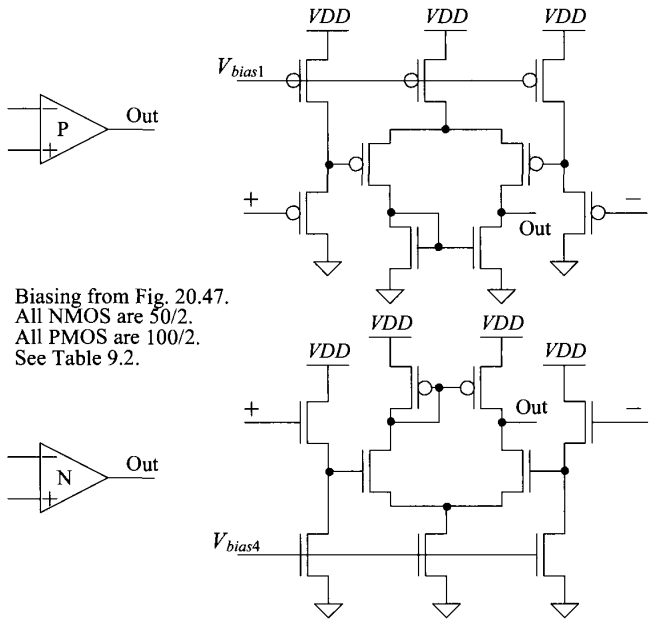


Figure 24.50 Diff-amps with source-follower level shifters for use in GE.

The simulation results in Fig. 24.47 were generated using the op-amp of Fig. 24.44 with the lengths reduced from 2 to 1. The low-frequency, open-loop gain was 53 dB. In Fig. 24.51 we added GE and used the resulting op-amp to generate the simulation results seen in Fig. 24.52 (with the GE compensation capacitors, C_{cGE}, set to zero). The low-frequency, open-loop gain is now 74 dB. The step response is seen as well. The current pulled from *VDD* is now 60 μA higher than the op-amp without GE. By replacing the diff-amps in the N and P amplifiers in Fig. 24.50 with folded-cascode OTAs (like the one seen in Fig. 24.42), we can get a greater increase in open loop gain. However, the stability of the regulated-drain feedback loop becomes even more important.

Bandwidth of the Added GE Amplifiers

Let's model the frequency response of the added amplifiers in Fig. 24.50 by

$$A_{GE}(f) = \frac{A_{GEDC}}{1 + j\frac{f}{f_{3dBGE}}} \tag{24.59}$$

Using this equation with Eqs. (24.30) and (24.58), we can write

$$A_{OLGE}(f) = \frac{A_{OLDC}}{1 + j \cdot \frac{f}{f_{3dB}}} \cdot \overbrace{\frac{A_{GEDC}}{1 + j \cdot \frac{f}{f_{3dBGE}}}}^{A_{GE}(f)} \tag{24.60}$$

IPR2023-00698
Theta EX2019

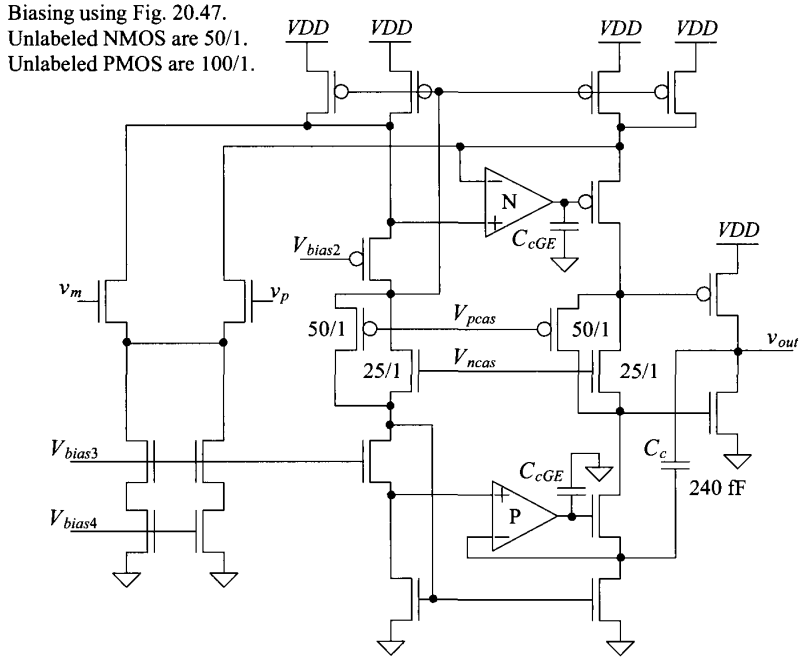


Figure 24.51 Folded-cascode op-amp with class AB output buffer and gain-enhancement.

However, the pole ($f_1 = f_{3dB}$) on the output of the folded-cascode OTA can be written, see Eqs. (20.67) and (24.47), with GE as

$$f_{3dB} = \frac{1}{2\pi(R_{ocasn} || R_{ocasp}) \cdot A_{GE}(f) \cdot C_c} \tag{24.61}$$

assuming the current through C_c is larger than the other currents, the output of the OTA may have to charge. If we substitute Eq. (24.61) into Eq. (24.60) and look at frequencies much larger than f_{3dB} (which from Fig. 24.52 is only a 100 kHz or so), we see that the frequency response of the added GE amplifiers cancels out of $A_{OLGE}(f)$. What this means is that the bandwidth of the added amplifiers doesn't need to be wide. As long as the bandwidths are larger than the f_{3dB} of the op-amp, the GE works as desired. Also, notice that when we substitute Eq. (24.61) into Eq. (24.60), the resulting transfer function shows a *doublet* (a pole and zero at the same frequency). If the pole and zero aren't at exactly the same frequency, then the settling time of the amplifier can be affected by the change in phase shift (noting that in Eq. (20.67) we are approximating the output resistance of the cascode structure). The output resistance of the cascode structures is decreasing with increasing frequency, while, at the same time, the decrease in the output resistance results in an increase in the circuit bandwidth (causing the gain-bandwidth product to be a constant).

IPR2023-00698
Theta EX2019

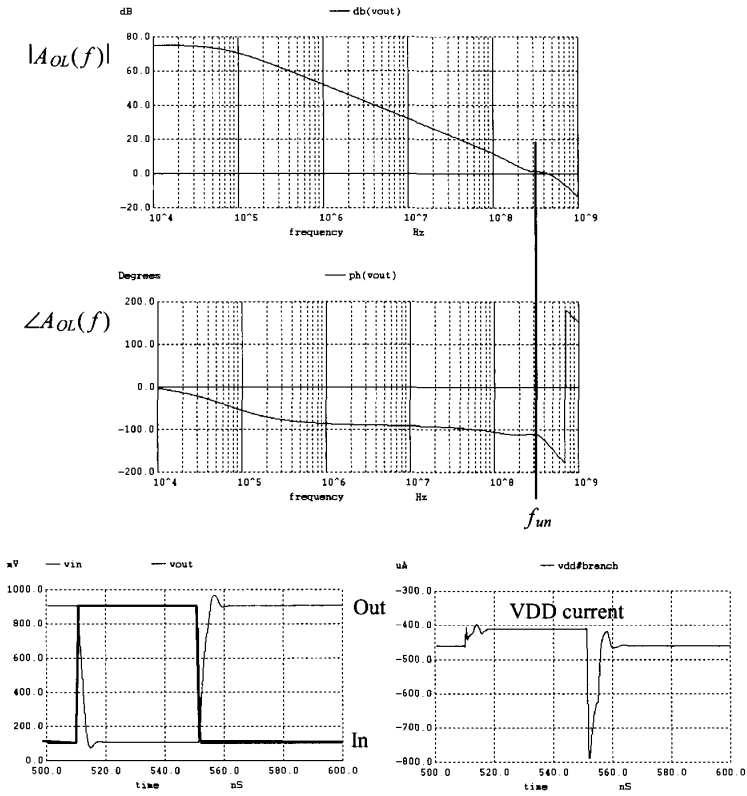


Figure 24.52 Operation of the op-amp in Fig. 24.51 driving a 1-pF load.

Compensating the Added GE Amplifiers

When we generated the simulations in Fig. 24.52, we didn't have compensation capacitors on the output of the added GE OTAs (N and P); that is, C_{cGE} were zero. The problem with this is that the resulting feedback loops aren't necessarily stable. As seen in Fig. 24.52, the step response shows some overshoot (indicating a low phase margin). Also, the frequency response isn't as monotonic as we would like (–20 dB/dec) around f_{un}. As with any feedback structure, compensation is important. *So we need to add capacitors to the outputs of the GE OTAs*. Figure 24.53 shows the response of the op-amp in Fig. 24.51 when C_{cGE}s are set to 240 fF (the GE OTAs then have f_{un} of 100 MHz, as calculated using Eq. (24.44)). The frequency response is more monotonic and the step response doesn't show overshoot and ringing. The gain-bandwidth product of the op-amp is 400 MHz. Note, it may be useful to regenerate Fig. 24.53, using larger values of C_{cGE}, to prove to oneself that indeed the bandwidth of the added GE amplifiers isn't critical. Also, it may be useful to replace the basic diff-amps with higher-gain OTAs to show that the low-frequency gain of the op-amp does indeed increase as Eq. (24.58) shows.

IPR2023-00698
Theta EX2019

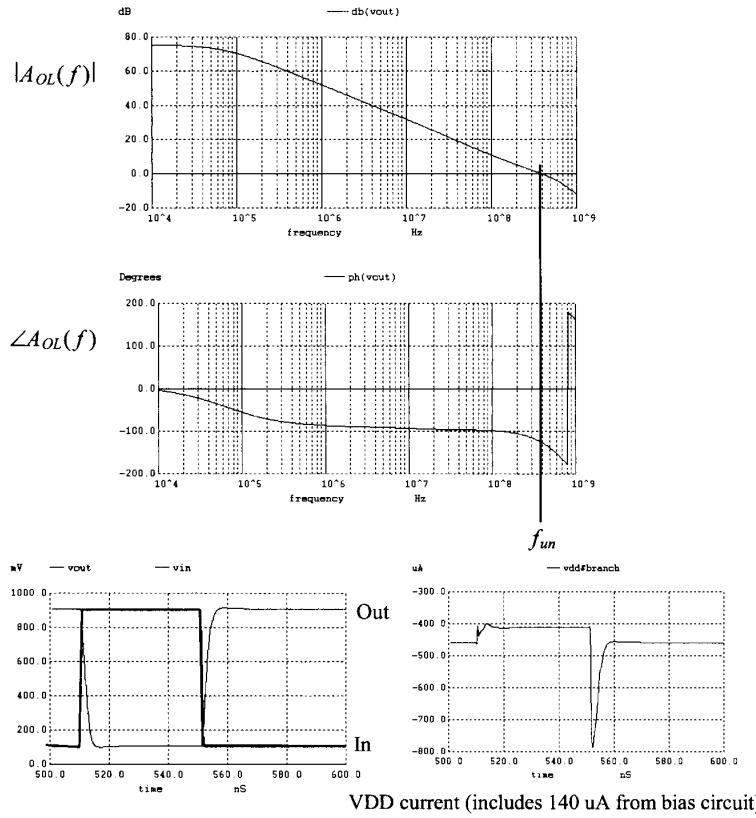


Figure 24.53 Operation of the op-amp in Fig. 24.51 driving a 1-pF load with GE compensation capacitors of 240 fF.

24.5 Some Examples and Discussions

A Voltage Regulator

One of the applications of an op-amp is in regulating a voltage on-chip. Figure 24.54 shows the basic topology. A voltage reference is used with the op-amp to generate a regulated voltage, V_{REG}. If the voltage reference is stable with temperature, the fact that the V_{REG} is a function of a ratio of resistors (so process or temperature changes in the resistance value don't affect the ratio) and the variation in the op-amp's open loop gain is desensitized using feedback makes the regulated voltage stable with process and temperature changes. The ideal (meaning that the op-amp has infinite open-loop gain) regulated voltage is

$$V_{REG} = V_{REF} \cdot \left(1 + \frac{R_A}{R_B}\right) \tag{24.62}$$

IPR2023-00698
Theta EX2019

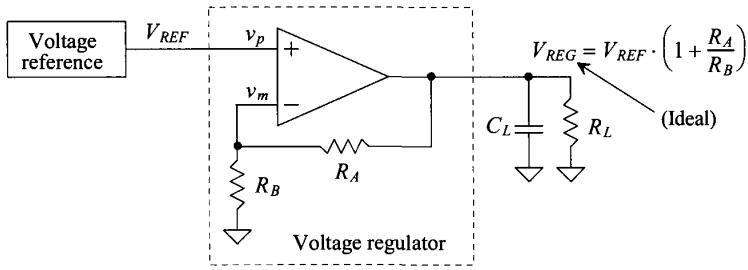


Figure 24.54 Schematic of a voltage regulator.

If the op-amp's open-loop gain is finite, then we can write

$$V_{REG} = A_{OL} \cdot (v_p - v_m) \tag{24.63}$$

and

$$v_m = V_{REG} \cdot \frac{R_B}{R_A + R_B} \text{ and } v_p = V_{REF} \tag{24.64}$$

Solving for the actual regulated voltage, gives

$$V_{REG} = V_{REF} \cdot \frac{1}{\frac{1}{A_{OL}} + \frac{R_B}{R_A + R_B}} \tag{24.65}$$

noting that when A_{OL} is infinite (or very large) this equation simplifies to Eq. (24.62). Equation (24.65) can be very revealing. Let's say that V_{REF} is 500 mV, R_B is an open, and R_A is a short (the op-amp is a voltage follower). Ideally, V_{REG} is also 500 mV. If we use a diff-amp with an open-loop gain of 20, then the actual regulated voltage is

$$V_{REG} = 0.5 \cdot \frac{1}{1.05} = 476\ mV \tag{24.66}$$

This simple example illustrates *why open-loop gain is important.* When op-amps are used in data converters (discussed in Chs. 28-30), they must amplify signals to within mV, which requires op-amp open-loop gains in the thousands.

Figure 24.55 shows a basic regulator topology using the two-stage op-amp. The regulator only sources current and so we've made the PMOS device, on the output, very large. One concern with using the large PMOS device is the op-amp offset (which can be considerable, see Fig. 24.4, for example). Another concern occurs when simulating. In some simulators both NRD and NRS (number of squares of implant area in the source or drain of a MOSFET) defaults to one. If the sheet resistance, rsh, is 50 ohms, then the resistance in series with the MOSFET is 100 ohms (limiting the current the MOSFET can supply). For large MOSFETs, it is generally a good idea to set NRD and NRS to zero. An example of a (large) SPICE MOSFET statement is

```
Mbig   vout   vn1   VDD   VDD   PMOS  L=1  W=10000  NRD=0   NRS=0
```

IPR2023-00698
Theta EX2019

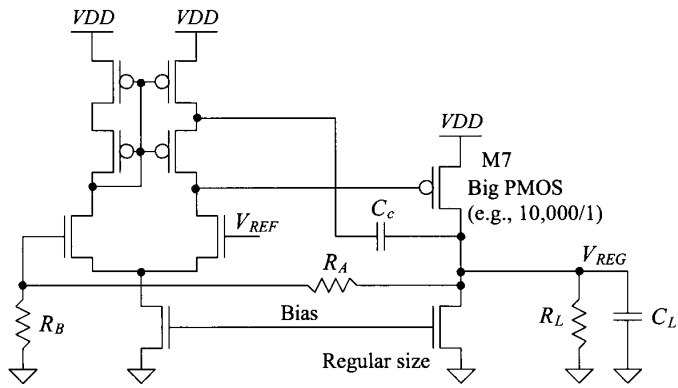


Figure 24.55 Schematic of a voltage regulator.

The point of making M7 big was to ensure that the regulator could supply a significant amount of current. However, notice that the gate of M7 cannot swing any lower (assuming that all MOSFETs remain in saturation) than

$$V_{G7,\min} = V_{REF} - V_{THN} \tag{24.67}$$

For example, if V_{REF} is 500 mV and V_{THN} is 250 mV, then the minimum voltage on the gate of M7 is 250 mV (25% of *VDD*). Ideally, we would like the gate of M7 to be able to swing to ground, fully turning it on. Towards this goal, consider the modified voltage regulator seen in Fig. 24.56. In this figure, we've used the OTA from Fig. 24.33 for the first stage of the op-amp. Now, when M5 turns on and M4 shuts off, the gate of M7 can get yanked down towards ground, allowing M7 to turn fully on.

In any practical regulator, a bypass capacitor (a big capacitor placed from V_{REG} to ground) is used to supply charge to the load for fast current transients. The required bandwidth of the regulator is reduced since the bypass capacitor (part of C_L in Figs. 24.55 and 24.56) takes care of the fast transients. It is desirable to have a large C_L for "filtering" the load's fast current needs. However, from our discussions earlier, this (large load capacitance) makes compensating the op-amp more challenging. In all of our compensation schemes, the pole associated with the output of the first stage was at a frequency much lower than the pole associated with the output stage (this was one of the reasons we used pole-splitting). To isolate the load from the second-stage gain, we might insert an NMOS source-follower (SF) between the load and the output of the second stage. The problem with this is that we won't be able to turn the SF on enough to supply significant amounts of current. Next, we might try replacing the common-source amplifier (M7) with a class AB stage (like the topology seen in Fig. 24.29). Here, again, the output won't be able to supply as much current as what we get with a full *VDD* across M7's V_{SG}. (Note that one of the reasons for using a voltage regulator is to hold V_{REG} constant even if *VDD* approaches V_{REG}.) If we must use a large load capacitor to supply charge for the fast variations in the load current and if the topology in Fig. 24.56 is used, then *we must try to compensate the op-amp with the load capacitance.*

IPR2023-00698
Theta EX2019

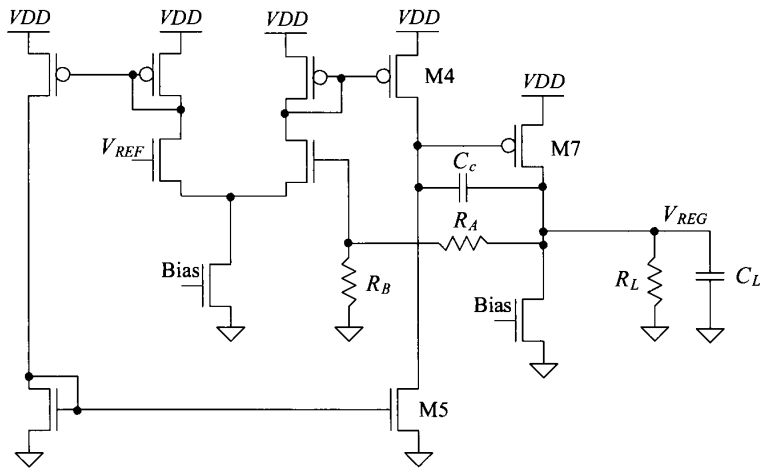


Figure 24.56 Schematic of a voltage regulator where the gate of M7 can be pulled all the way to ground.

If we are compensating the op-amp (the voltage regulator) using the load capacitance, then we've got to make sure that there is a minimum value of C_L. Further, increasing the load capacitance should make the regulator more stable, slowing the response of the op-amp. The reduction in the speed of the op-amp is offset by the fact that a larger load capacitance can supply more charge to a fast current transient before the op-amp must respond.

Why use a compensation capacitor, C_c , if we are compensating with the load capacitance? The compensation capacitor pushes the pole associated with the output node to a higher frequency (exactly what we don't want to do). We still use the capacitor (but with a smaller value) for large-signal reasons. If V_{REG} drops suddenly, the decrease in voltage is fed back directly to the gate of M7 through C_c. This turns M7 (quickly) on and allows it to pull V_{REG} back up (bypassing the slower feedback action of the op-amp). Note then that C_c is *not used* for compensating the op-amp (and actually makes the regulator more unstable). Further note that increasing the gain of the op-amp by making the first-stage gain larger (by, say, using a cascode diff-amp to increase the output resistance of the diff-amp, R_1) also hurts the regulator's stability. (By using a larger R_1, the pole associated with the output of the first stage is moved lower in frequency.)

To estimate the location of the output pole, assuming that it is the dominant pole in the system, let's calculate the closed-loop output resistance of the op-amp $R_{out,CL}$ (the open- loop output resistance is R_2). We know that the low-frequency, open-loop gain of the op-amp is

$$A_{OLDC} = \frac{v_{out}}{v_p - v_m} = g_{m1} R_1 g_{m2} R_2 \qquad (24.68)$$

If we replace the load in Fig. 24.54 with a test voltage, v_t , and note that V_{REF} (= v_p = 0) is AC ground, then with the help of Fig. 24.57 we can write

IPR2023-00698
Theta EX2019

$$i_t = \frac{v_t}{R_2} + \frac{v_t}{R_A + R_B} + g_{m1}R_1 g_{m2} \cdot \frac{v_t}{A_{CL}} \tag{24.69}$$

If we assume that the current through the feedback path ($R_A + R_B$) is small, then the output resistance of the regulator is

$$R_{out,CL} \approx \frac{v_t}{i_t} = \frac{1}{\left(\frac{1}{R_2} + \frac{g_{m1}R_1 g_{m2}}{A_{CL}}\right)} \approx \frac{A_{CL}}{A_{OLDC}} \cdot R_2 \tag{24.70}$$

Our output pole is located at

$$f_2 = \frac{1}{2\pi R_{out,CL} \cdot C_L} = \frac{A_{OLDC}}{2\pi R_2 C_L \cdot A_{CL}} \tag{24.71}$$

For this pole to compensate the op-amp, we want to use a low value of open-loop gain, A_{OLDC}. Again, this means that we don't want a large value of resistance on node 1 (the output of the diff-amp). The problem with this is that the lower the open-loop gain, the more difficult it is to regulate the output voltage, as indicated in Eq. (24.66).

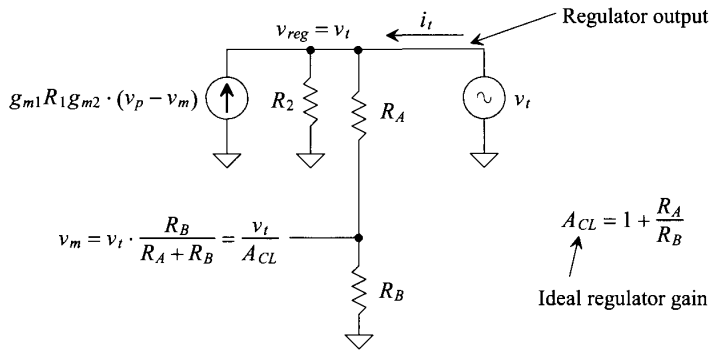


Figure 24.57 Determining the closed-loop output resistance of a two-stage op-amp.

To get an idea for the minimum size load capacitance required for a stable op-amp, let's use the parameters from Table 9.2. To begin, note that the more current the regulator supplies to a load (the more current flowing through M7), the smaller the value of R_2. This (decreasing R_2) means that the regulator becomes more likely to be unstable (f_2 increases) as it supplies more current. However, at the same time, the open-loop gain, A_{OLDC}, drops and the ability of the op-amp to regulate the output declines.

Let's assume M7's r_o and g_m are the values given in Table 9.2. Because M7 is so wide, this means (using these values for r_o and g_m) that it is conducting very little current. The value of f_1 can be estimated (assuming the $C_{gs7} >> C_c$ and the width of M7 is 10,000 or 100 times larger than the widths of the PMOS in Table 9.2) as

$$f_1 \approx \frac{1}{2\pi R_1 C_{gs7}} = \frac{1}{2\pi \cdot (167k||333k) \cdot 100 \cdot 8.34f} = 1.7\ MHz \tag{24.72}$$

IPR2023-00698
Theta EX2019

The DC open-loop gain of the op-amp is calculated using Eq. (24.68) as 277 (no load). If we use the op-amp in the follower configuration where R_B is infinite and R_A is zero (A_{CL} is 1), then (knowing $R_2 = 167k||333k = 111k$)

$$f_2 = \frac{A_{OLDC}}{2\pi R_2 C_L \cdot A_{CL}} = \frac{277}{2\pi(111k)\cdot C_L} \tag{24.73}$$

Knowing we want $f_2 << f_1$, A_{OLDC} = 277, and f_1 = 1.7 MHz, we can set the unity-gain frequency to 1 MHz (making sure that it is less than f_1). If we want a –20 dB/decade roll-off above f_2 (now the low-frequency pole) then, as seen in Eq. (24.33), we can write

$$f_{un} = A_{OLDC}\cdot f_2 = \frac{A_{OLDC}^2}{2\pi R_2 C_{L\min}\cdot A_{CL}} = 1\ MHz \tag{24.74}$$

Solving for the minimum value of C_{Lmin} using these equations gives 110 nF. Clearly, this value is too large if we want the regulator to be completely on-chip. Rewriting Eq. (24.74)

$$C_{L\min} = \frac{(g_{m1}R_1 g_{m2})^2 R_2}{2\pi\cdot f_{un}\cdot A_{CL}} \tag{24.75}$$

Note that linearly increasing the current through M7 causes R_2 to decrease linearly (assuming that M7's output resistance dominates R_2) and g_{m2}^2 to increase linearly (see Eqs. (9.6) and (9.22)). The result is that C_{Lmin} doesn't change. However, notice what happens if we include R_L in the calculations. If $R_L << r_{o7}$, then we can write

$$\frac{C_{L\min}}{R_{L\max}} = \frac{(g_{m1}R_1 g_{m2})^2}{2\pi\cdot f_{un}\cdot A_{CL}} \tag{24.76}$$

If we select $C_{L\min} = 1000\,pF$, then $R_{L\max} = 1\,k\Omega$. We use the variable R_{Lmax} to indicate that this is the maximum value of load resistance possible for a stable regulator (we want the regulator to always supply at least V_{REG}/R_{Lmax} current). If the regulator's output voltage is 500 mV, it must be supplying at least 500 μA of current. This resistor keeps the open-loop gain low when the regulator is supplying small amounts of current. We might think, after looking at Eq. (24.73), that by supplying more current (R_L or R_2 decreasing further), the location of f_2 increases and the regulator moves towards instability. However, for further decreases in the load resistance, the current supplied by M7 increases and thus so does its g_m. What happens is this: we get the canceling of effects. As the current supplied to the output increases linearly with a linear decrease in R_L , the value of g_m^2 increases linearly and the effects cancel.

A final comment: because this design is different from the previous op-amp designs, thorough characterization using simulations is needed. Steps in the load current as well as a wide range of load capacitances and resistances should be simulated to ensure that the regulator remains stable under all possible *VDDs*, temperatures, and process shifts.

Bad Output Stage Design

In this section we'll discuss some problems with biasing in Class AB output stages. To begin, consider the op-amp design seen in Fig. 24.58. This topology is a diff-amp driving an inverter. The problem with this design is that the current flowing in the push-pull output stage (an inverter) isn't set by a bias circuit. Because of this, the current may be

IPR2023-00698
Theta EX2019

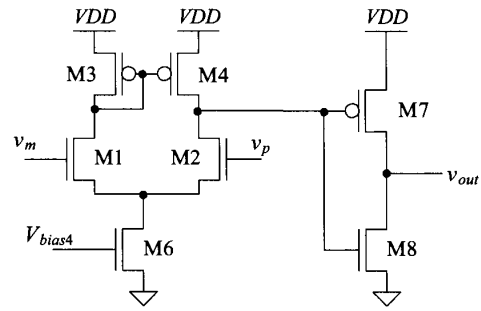


Figure 24.58 Bad op-amp design.

significant and vary greatly with both temperature and process shifts. Further, notice that when the inputs to the op-amp are equal, the drain of M4 is at the same potential as the drain of M3 (the currents in M3 and M4 are equal). This, as we've already discussed, can be used to set the bias current in M7 (we treat M7, for biasing purposes, as if its gate were tied to the gates of M3 and M4). If we call the gate potential of M7 (and thus the gate potential of M8) $VDD - V_{SG}$, then for M8 to be in saturation

$$V_{out} \geq VDD - V_{SG} - V_{THN} \tag{24.77}$$

If VDD is 1 V, $V_{SG} = 350$ mV, and $V_{THN} = 250$ mV, then V_{out} must be greater than 400 mV for M8 to be saturated. What this means is that for wide output swings, M8 will move into triode and the push-pull output amplifier gain will drop (killing the overall gain of the op-amp). What we need to do is drop the gate potential of M8 down so that its drain can swing to a lower voltage without it trioding. Towards this goal, consider adding a source-follower level shifter to the output of the op-amp, as seen in Fig. 24.59. Now the output voltage in the amplifier can swing lower without M8 trioding. However, we still aren't controlling the current in the output stage. By not controlling this current, as we've already discussed, we produce an op-amp with poor systematic, input-referred offset voltage. If, for example, we have an input-referred offset of 10 mV and the diff-amp gain is 20, then the change in the output voltage on the gate of M7 is 200 mV. This can cause

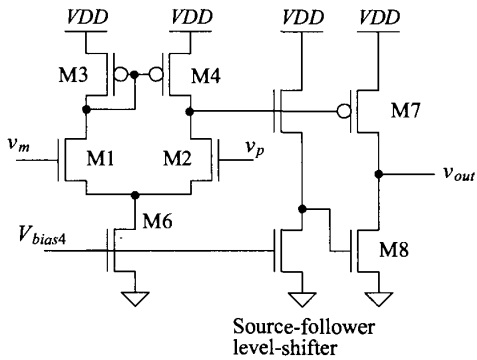


Figure 24.59 Another bad op-amp design.

IPR2023-00698
Theta EX2019

the output current flowing in M7/M8 to be hundreds of μA above the current in the diff-amp.

We might try connecting the gate of M8 to the gates of M3/M4. When the noninverting input voltage, v_p , goes high, M2 turns on and causes the gate potential of M7 to drop (turning M7 further on). At the same time, the gate potential of M3 increases, causing M8 to also turn further on. The result is that M7 and M8 both turn on and fight each other for control of v_{out}. For proper operation, the gates of the output push-pull stage should move in the same direction.

To precisely control the current in the output stage, consider the topology seen in Fig. 24.60. Since, for biasing purposes, we treat the drain of M4 as if it's at the same potential as the gate of M3, the currents in M7, M9, and M10 are equal when the op-amp inputs are at the same potential. We also treat M8 as if its gate were tied to the gate of M11. The result is that the quiescent currents in M7 and M8 are equal and set by the biasing of M6. We seem to have solved the problem with biasing the class AB output stage.

However, a different problem arises. We now have three high-impedance nodes: the output of the diff-amp (a pole we've called f_1), the output of the op-amp (a pole we've called f_2), and, now, the gate of M8. If we try to make the node at the gate of M7 the dominant node, then the path that the op-amp inputs sees through M3 to M9 to the gate of M8 (the other high-impedance node) isn't affected. If we add compensation capacitors from the op-amp output to both the gates of M7 and M8, the current fed back isn't necessarily balanced. This imbalance may only be noticed when the op-amp's large signal step response is simulated. (It is generally not seen in an AC simulation.) The result is that the op-amp's settling time and stability become marginal in some situations. Having said this, the topology can still be useful in some situations (e.g., low power or low *VDD*).

In an attempt to improve or simplify the op-amp in Fig. 24.60, we might try to make M12 diode connected and remove M10/M11 from the circuit. With this change, however, we don't get class AB action (M9 mirrors the current in M3, which flows in the diode-connected M12 and is mirrored by M8). Finally, again, why not eliminate M9 and M12 followed by connecting the gate of M8 to the diode-connected M11? The gate potentials of M8 and M7 would then move in opposite directions, causing M7 to fight with M8 for control of the output voltage.

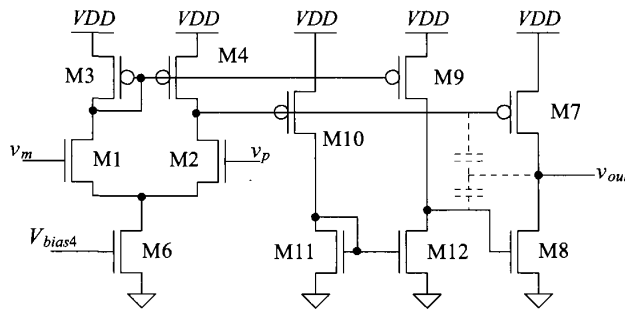


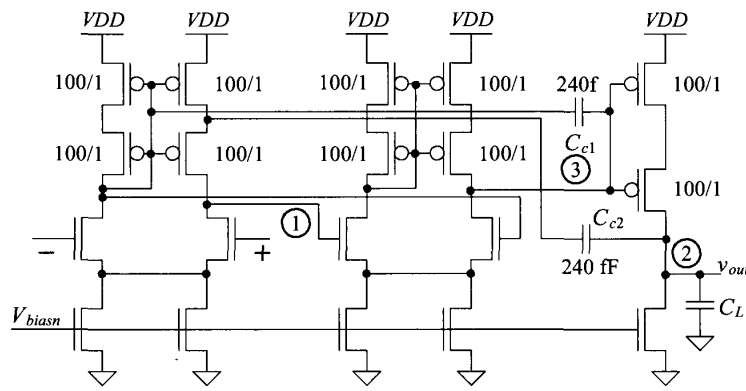
Figure 24.60 Controlling the current in a class AB output stage.

IPR2023-00698
Theta EX2019

Three-Stage Op-Amp Design

In some applications we need an op-amp that can operate with a low power-supply voltage or with minimal power from *VDD*. The bias circuit we developed in Fig. 20.47 can pull more current from *VDD* than the op-amp it biases. (So, if possible, use the same bias circuit for multiple op-amps, if power and layout area are of concern.) If we look at the heart of this bias circuit (Fig. 20.47) and its operation, the Beta-multiplier in Figs. 20.22 and 20.23, we see that *VDD* can be decreased down to 500 mV with the bias currents relatively unaffected. The problem in using a *VDD* which is half of the normal *VDD* (= 1 V) is that our class AB output stage (Fig. 24.44 for example) fails to operate. For proper operation of the class AB output stage using the floating current sources (biased with V_{ncas} and V_{pcas}), we need a *VDD* > 800 mV (perhaps higher over process corners and temperature). Also, as already discussed, if we connect a resistive load to the output of the two-stage op-amp in Fig. 24.2, the gain is reduced.

Towards keeping a large gain and lowering the power supply voltage, consider the three-stage op-amp in Fig. 24.61. This design is a cascade of two diff-amps followed by a common-source amplifier. If we connect a resistive load to the output of the common-source stage, the overall op-amp gain remains relatively high due to the cascaded gain of the diff-amps. The concern with this topology (an op-amp with more than two stages), as alluded to earlier, is compensation. We still compensate the op-amp so that the pole associated with node 1 is dominant. The nesting of the compensation capacitors within the op-amp in Fig. 24.61 is sometimes called *nested Miller compensation*. We'll avoid the use of this label here since we aren't strictly using Miller compensation (Fig. 24.8 among others) but rather the indirect compensation method discussed earlier (see Fig. 24.18 and the associated discussion). Note how the current through C_{c1} feeds back to node 1 through the diode-connected load of the input diff-amp. This is necessary to ensure that the signal through C_{c1} adds with the signal fed back directly to node 1 through C_{c2}.



Bias circuit in Fig. 20.22.
See Table 9.2.

Figure 24.61 A three-stage op-amp.

IPR2023-00698
Theta EX2019

Because of pole-splitting, the location of the pole at node 3 (the output of the second diff-amp) is pushed out to, see Eq. (24.24),

$$f_3 = \frac{g_{mn} \cdot C_{c1}}{2\pi \cdot C_1(C_{L2} + C_{c1})} \approx \frac{g_{mn}}{2\pi C_1} \tag{24.78}$$

a large frequency. Here we are assuming that the transconductance of the second diff-amp is g_{mn} and that C_{c1} is much larger than the input capacitance of common-source amplifier (which is the second diff-amp's C_{L2}). The location of the pole on the output of the amplifier is still given by

$$f_2 = \frac{g_{mp} \cdot C_{c2}}{2\pi \cdot C_1(C_L + C_{c2})} \tag{24.79}$$

where g_{mp} is the transconductance of the common-source amplifier. Note that we aren't discussing the LHP zeroes whose locations are specified using Eq. (24.22). We'll discuss these zeroes in a moment. We'll compensate, as usual, the op-amp so that the unity-gain frequency is less than the frequencies of the poles at f_2 and f_3.

The low-frequency, open-loop gain of the op-amp is given by

$$A_{OLDC} = \overbrace{g_{mn}(r_{on}||r_{op})}^{\text{first diff-amp's gain, } A_1} \cdot \overbrace{g_{mn}(r_{on}||r_{op})}^{\text{second diff-amp's gain, } A_2} \cdot \overbrace{g_{mp}(r_{on}||r_{op})}^{\text{common-source gain, } A_3} \tag{24.80}$$

At frequencies above the op-amp's f_{3dB} and below the gain-bandwidth product (the unity-gain frequency, f_{un}), the gain of the first-stage is rolling off. For all intents and purposes, the AC signal on the output of the first diff-amp is considerably smaller than the AC outputs of the second or third stages. In this situation the current fed back to node 1, knowing that the output voltage of the second diff-amp (at node 3) is v_{out}/A_3 , is

$$i_{Cctot} \approx \frac{v_{out}}{1/j\omega C_{c2}} + \frac{v_{out}/A_3}{1/j\omega C_{c1}} \tag{24.81}$$

The output current of the diff-amp is

$$i_{out1} = g_{mn}(v_p - v_m) \tag{24.82}$$

Because this current must equal the current fed back to node 1, we can equate Eqs. (24.81) and (24.82) to write

$$\frac{v_{out}}{v_p - v_m} = \frac{g_{mn}}{j\omega(C_{c2} + C_{c1}/A_3)} \tag{24.83}$$

The unity-gain frequency is then

$$f_{un} = \frac{g_{mn}}{2\pi(C_{c2} + C_{c1}/A_3)} \approx \frac{g_{mn}}{2\pi C_{c2}} \tag{24.84}$$

noting that the unity-gain frequency is a function of the third-stage gain. Here, we'll set C_{c2} to 240 fF to get a unity-gain frequency of 100 MHz. We still need C_{c1} to push node 3's pole to a higher frequency (pole splitting). We'll also set C_{c1} to 240 fF.

Figure 24.62 shows the frequency response of the op-amp in Fig. 24.61 without driving a load capacitance. As designed for using Eq. (24.84), the unity-gain frequency is 100 MHz. Also, as specified in Eqs. (24.22) and (24.25), we have zeroes around the unity-gain frequency. It appears from the frequency response in Fig. 24.62 that the zeroes

IPR2023-00698
Theta EX2019

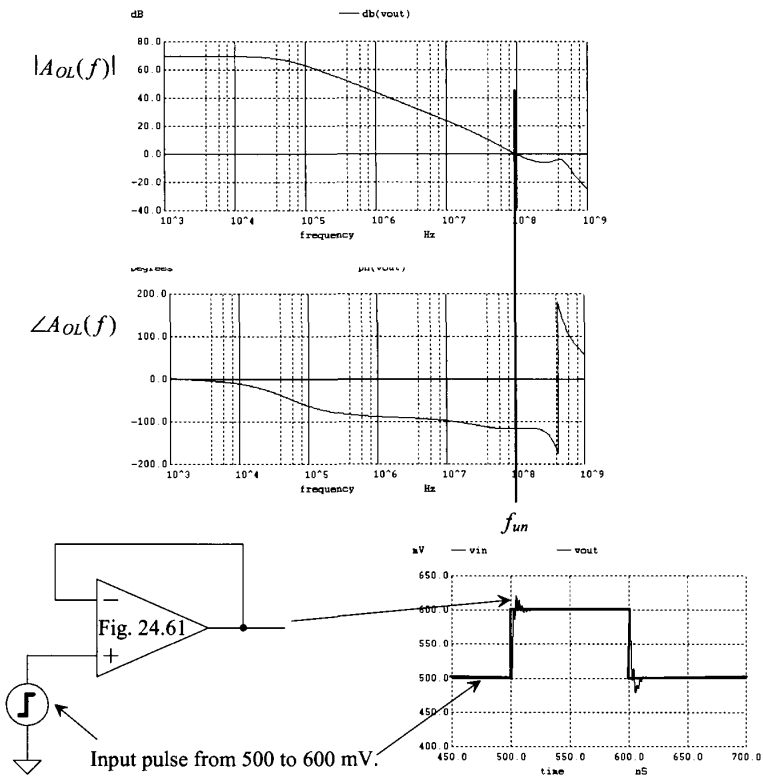


Figure 24.62 Unloaded AC and step responses for the op-amp in Fig. 24.61.

occur above the unity-gain frequency and cause the open-loop response to flatten out. What this indicates is that one of the zeroes is canceling one of the higher-frequency poles. Note that the PM is 45°. The step response seen in Fig. 24.62 shows the characteristic ringing associated with a PM of 45°.

If we drive a load capacitance, then we expect the third-stage gain, A_3, to start to decrease at a lower frequency (it starts to decrease sooner with increasing frequency). Looking at Eq. (24.84), we then expect f_{un} to drop in value. Figure 24.63 shows the simulation where the op-amp in Fig. 24.61 is driving a 1 pF load. As expected, the addition of the 1 pF load causes f_{un} to drop in value. The PM is approximately 25°. With such a low PM, we expect the output to show significant ringing (and, as seen in Fig. 24.63, it does). To increase the PM, we can increase the value of the compensation capacitor. We can increase C_{c2} to, say, 2400 fF, for a unity gain frequency of (roughly) 10 MHz. We don't necessarily need to increase C_{c1} to decrease f_{un}, as seen in Eq. (24.84).

IPR2023-00698
Theta EX2019

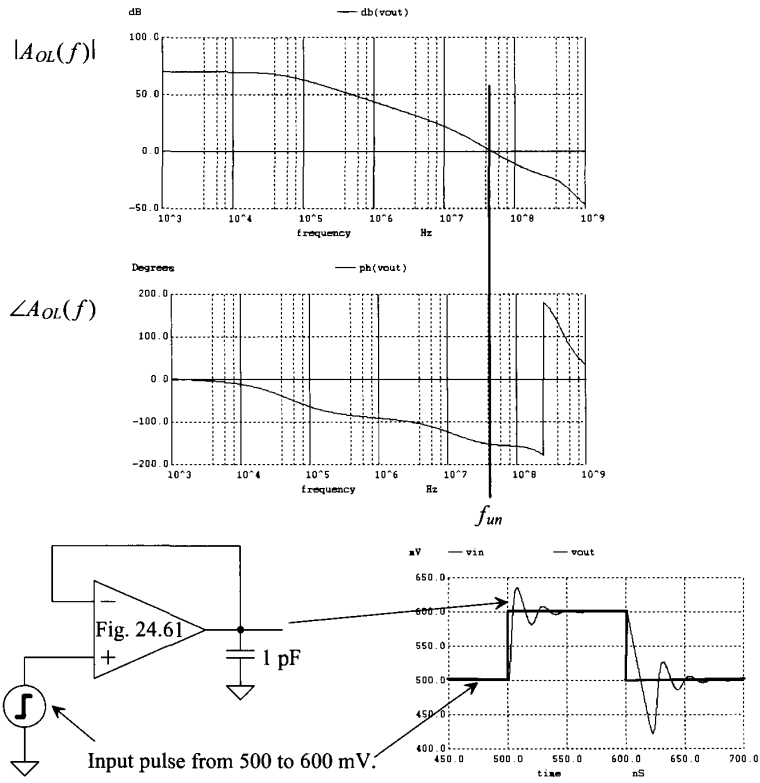


Figure 24.63 AC and step responses for the op-amp in Fig. 24.61 driving 1 pF.

Finally, why did we choose the topology seen in Fig. 24.61 for our three-stage op-amp example? The main reason for using this topology comes from biasing the common-source stage. If we were to replace the second diff-amp with a common-source stage (so our last two stages in the op-amp are common-source amplifiers), we wouldn't have a known second stage output voltage to bias the final common-source stage. The exact output voltage of a common-source amplifier, without feedback to set it, is unknown (the two drains of the NMOS and PMOS will either float up or float down, depending on which device is sourcing or sinking the most current). By using the diff-amp stages, as seen in Fig. 22.8, we ensure that each stage is biased with a known current. We don't use a diff-amp for the final stage of the op-amp because of the diff-amp's limited output swing.

Another benefit of using two diff-amps on the input of the op-amp is the increase in the *CMRR*. The overall common-mode gain is the product of each diff-amp's common-mode gain. If a single diff-amp has a *CMRR* of 80 dB, then the cascade of two diff-amps results in a *CMRR* of 160 dB.

IPR2023-00698
Theta EX2019

ADDITIONAL READING

[1] V. Saxena, *Indirect Feedback Compensation Techniques for Multi-Stage Operational Amplifiers*, Master's Thesis, 2007. Available at CMOSedu.com. Excellent treatment of compensating op-amps for high-speed and low-power.

[2] R. Harjani, R. Heineke, and F. Wang, "An integrated low-voltage class AB CMOS OTA," *IEEE Journal of Solid-State Circuits,* vol. 34, pp. 134–142, 1999.

[3] L. Moldovan and H. H. Li, "A rail-to-rail, constant gain, buffered op-amp for real time video applications," *IEEE Journal of Solid-State Circuits,* vol. 32, pp. 169–176, February 1997.

[4] F. You, S. H. K. Embabi, and E. Sánchez-Sinencio, "A multistage amplifier topology with nested Gm-C compensation for low-voltage application," *IEEE International Solid-State Circuits Conference,* vol. XL, pp. 348–349, Feb. 1997.

[5] W. S. Wu, W. J. Helms, J. A. Kuhn, and B. E. Byrkett, "Digital-compatible high-performance operational amplifier with rail-to-rail input and output ranges," *IEEE Journal of Solid-State Circuits,* vol. 29, pp. 63–66, January 1994.

[6] R. Hogervorst, J. P. Tero, R. G. H. Eschauzier, and J. H. Huijsing, "A Compact Power-Efficient 3 V CMOS Rail-to-Rail Input/Output Operational Amplifier for VLSI Cell Libraries," *IEEE Journal of Solid State Circuits,* vol. 29, pp. 1505–1513, December 1994.

[7] J. Ramírez-Angulo and E. Sánchez-Sinencio, "Active compensation of operational transconductance amplifier filters using partial positive feedback," *IEEE Journal of Solid-State Circuits,* vol. 25, pp. 1024–1028, August 1990.

[8] K. Bult and G. J. G. M. Geelen, "A Fast-Settling CMOS Op-Amp for SC Circuits with 90-dB DC Gain," *IEEE Journal of Solid State Circuits,* vol. 25, pp. 1379–1384, December 1990.

[9] S. M. Mallya and J. H. Nevin, "Design procedures for a fully differential folded-cascode CMOS operational amplifier," *IEEE Journal of Solid-State Circuits,* vol. 24, pp. 1737–1740, December 1989.

[10] M. Banu, J. M. Khoury, and Y. Tsividis, "Fully Differential Operational Amplifiers with Accurate Output Balancing," *IEEE Journal of Solid State Circuits,* vol. 23, No. 6, pp. 1410–1414, December 1988.

[11] B. K. Ahuja, "An Improved Frequency Compensation Technique for CMOS Operational Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 18, pp. 629–633, December 1983.

[12] P. R. Gray and R. G. Meyer, "MOS Operational Amplifier Design: A Tutorial Overview," *IEEE Journal of Solid-State Circuits,* vol. 17, pp. 969–982, 1982.

[13] B. Y. Kamath, R. G. Meyer, and P. R. Gray, "Relationship Between Frequency Response and Settling Time of Operational Amplifiers," *IEEE Journal of Solid State Circuits,* vol. SC-9, pp. 347–352, December 1974.

IPR2023-00698
Theta EX2019

PROBLEMS

24.1 Suggest, and verify with simulations, a method for reducing the minimum input common-mode voltage of the op-amp in Fig. 24.2.

24.2 Redesign the bias circuit for the op-amp in Fig. 24.2 for minimum power. Compare the power dissipation of your new op-amp design (actually the bias circuit) to the design in Fig. 24.2. Using your redesign, generate the plots seen in Fig. 24.3.

24.3 Show, using simulations, how a 1% mismatch in the widths of M1 and M2 in the op-amp of Fig. 24.2 affect the op-amp's input-referred offset voltage. Compare this offset to the offset caused by a 1% mismatch in the widths of M3 and M4. Quantitatively explain why one is worse than the other.

24.4 Simulate the use of the "zero-nulling" circuit in Fig. 24.15 in the op-amp of Fig. 24.8. Show AC, operating-point, and transient (step) operation of the resulting op-amp. Verify with the .op analysis that the gate of MP1 is at the same potential as the gate of M7 in quiescent conditions.

24.5 When we derived Eq. (24.24), we lumped C_c into C_2 and neglected the effects of MCG. A more accurate model for the indirect compensation, assuming $C_L >>$ the output capacitance of the amplifier, is seen in Fig. 24.64. Using this model, estimate the location of the output pole.

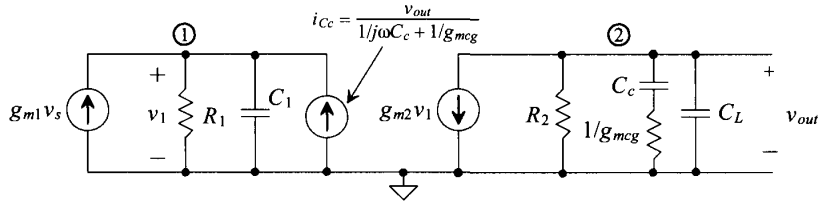


Figure 24.64 Model used to estimate bandwidth when indirect feedback current. See Problem 24.5.

24.6 Regenerate Fig. 24.19 using a 2.4 pF compensation capacitor. On the resulting simulation output, label the location f_1 , f_2 , f_z , and f_{un}. How do the simulated results compare to the hand calculated values?

24.7 For the op-amp in Fig. 24.21, determine the *CMRR* using hand calculations. Verify your hand calculations using simulations. How does the *CMRR* change based on the DC common-mode voltage?

24.8 Simulate the *PSRR*s for the op-amp in Fig. 24.8 (with an R_z of 6.5k and a C_c of 2.4 pF) and compare the results to the op-amp in Fig. 24.21 when C_c is set to (also) 2.4 pF (so each op-amp has the same gain-bandwidth).

24.9 Simulate the operation of the op-amp in Fig. 24.28. Show the open-loop frequency response of the op-amp. What is the op-amp's PM? Show the op-amp's step response when it is put into a follower configuration driving a 100 fF load with an input step in voltage from 100 mV to 900 mV.

IPR2023-00698
Theta EX2019

24.10 The op-amp seen in Fig. 24.29 has a gain-bandwidth product (f_{un}) of about 100 MHz. Suppose that this op-amp is used in the amplifier seen in Fig. 24.65 (gain of −5). Estimate the amplifier's closed loop bandwidth (where the output of the amplifier is −3 dB down from its low-frequency value) using Eq. (24.34). Verify your results using simulations (transient analysis). What are the maximum and minimum voltages allowable for an input sinewave if the output voltage of the amplifier must lie between 100 and 900 mV?

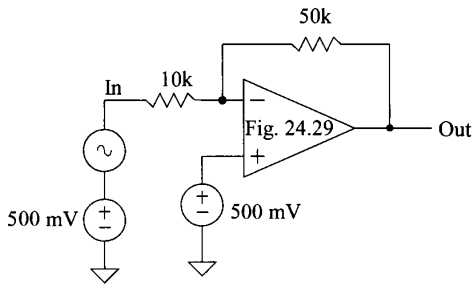


Figure 24.65 Amplifier used for Problem 24.10.

24.11 Suppose it is decided to eliminate the 500 mV common-mode voltage in the amplifier seen in Fig. 24.65 and use ground, as seen in Fig. 24.66. Knowing that the input voltage can only fall between ground and *VDD*, what is the problem one will encounter?

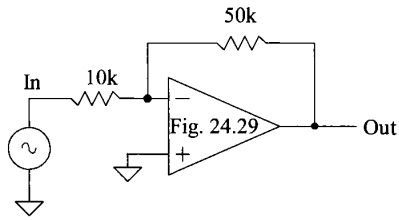


Figure 24.66 Amplifier used for Problem 24.11. What's wrong with this topology?

24.12 To limit the current flowing in MOP or MON in the op-amp of Fig. 24.29 (to protect the op-amp from destruction if its output is shorted to ground, for example), we may add 100 Ω resistors, as seen in Fig. 24.67. What is the maximum amount of current this modified op-amp can source/sink? How is the closed-loop output resistance of the op-amp affected. (Hint: see Eq. (24.70).) Simulate the operation of the op-amp using the topology seen in Fig. 24.31. Is there any noticeable difference between the simulation output in Fig. 24.31 and the output with the 100 Ω resistors present?

IPR2023-00698
Theta EX2019

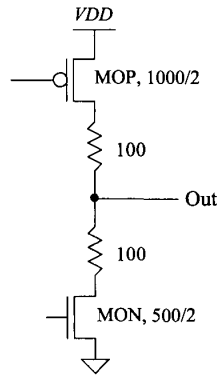


Figure 24.67 Adding resistors to the output of the op-amp in Fig. 24.29 for short circuit protection. See Problem 24.12.

24.13 Resimulate the OTA in Fig. 24.33 driving a 1 pF load (to determine f_{un}) if K = 10. How do the simulation results compare to the hand calculations using Eq. (24.41)? Estimate the parasitic poles associated with the gates of M4 and M5 (for example, the pole associated with the gate of M4 is $\approx 1/g_{m41} \cdot C_{sg4}$). Are these poles comparable to f_{un}?

24.14 Using the OTA in Fig. 24.35, design a lowpass filter with a 3 dB frequency of 1 MHz.

24.15 Suppose M8T in the op-amp of Fig. 24.37 is removed and replaced with a short from the output to the drain of M8B. How will the gain be affected? Verify your answer with SPICE.

24.16 Suppose, to simulate the open-loop gain of an OTA, the big resistor and capacitor used in Fig. 24.43 are removed and the inverting input is connected to 500 mV. Will this work? Why or why not? What happens if the OTA doesn't have an offset voltage? Will it work then?

24.17 Why is the noninverting topology (Fig. 24.49) inherently faster than the inverting topology (Fig. 24.39). What are the feedback factors, β, for each topology. Use the op-amp in Fig. 24.48 to compare the settling times for a +1 and a −1 amplifier driving 10 pF.

24.18 Suppose, to reduce power, the lengths of the current sources used in the amplifiers seen in Fig. 24.50 are increased from 2 to 10. Will the AC performance of the op-amp used to generate Fig. 24.53 change with this modification? Why or why not? Verify your answer using a SPICE simulation. Compare the currents used in the modified op-amp (with the lower power GE diff-amps) to the current seen in Fig. 24.53.

24.19 To increase the gain of the op-amp in Fig. 24.51, we may replace the GE diff-amps with folded-cascode OTAs. Will we need source-follower level-shifters

IPR2023-00698
Theta EX2019

in the new design? Regenerate the simulation data seen in Fig. 24.53 using the folded-cascode OTAs.

24.20 Suppose an op-amp is to be used to amplify 250 mV to 500 mV with an error less than 1 mV. Estimate the minimum required op-amp open loop gain.

24.21 Design a voltage regulator to supply at least 50 mA of current at 500 mV with a *VDD* as low as 600 mV. Assume that a 500 mV voltage reference is available and that the load capacitance is, minimum, 1,000 pF. How does the design respond to a load current pulse from 0 to 50 mA? Use SPICE to verify your design.

24.22 Using the nominal sizes from Table 9.2 and the bias circuit in Fig. 20.47, simulate, using an .op analysis, the operation of the op-amp in Fig. 24.58 in the configuration seen in Fig. 24.9. What is the current flowing in M7 and M8 when *VDD* is 1 V? is 1.2 V?

24.23 Repeat Problem 24.22 for the op-amp in Fig. 24.59.

24.24 Using the information from Table 9.2 and the bias circuit in Fig. 20.47, demonstrate the operation (AC and transient) of the op-amp in Fig. 24.60 driving a 100 fF load. Show that the bias current pulled from *VDD* is relatively constant for a *VDD* of 1 or 1.2 V (put the op-amp in the configuration seen in Fig. 24.9).

24.25 Replace the common-source output stage in the op-amp of Fig. 24.61 with a class AB output stage like the one seen in Fig. 24.60. Simulate the operation of the amplifier (AC and transient).

IPR2023-00698
Theta EX2019

Chapter
25

Dynamic Analog Circuits

In Chapter 14 we discussed dynamic logic gates. Dynamic logic is useful for reducing power dissipation and the number of MOSFETs used to perform a given circuit operation (layout area). Dynamic analog circuits exploit the fact that information can be stored on a capacitor or gate capacitance of a MOSFET for a period of time. In this chapter, we discuss analog circuits such as sample and holds, current mirrors, amplifiers, and filters using dynamic techniques.

25.1 The MOSFET Switch

A fundamental component of any dynamic circuit (analog or digital) is the switch (Fig. 25.1). An important attribute of the switch, in CMOS, is that under DC conditions the gate of the MOSFET does not draw a current. Therefore, neglecting capacitances from the gate to the drain/source, we find that the gate control signal does not interfere with information being passed through the switch. Figure 25.2 shows the small-signal resistance of the switches of Fig. 25.1 plotted against input voltage. The benefits of using the CMOS transmission gate are seen from this figure, namely, lower overall resistance. Another benefit of using the CMOS TG is that it can pass a logic high or a logic low without a threshold voltage drop. The largest voltage that an NMOS switch can pass is $VDD - V_{THN}$, while the lowest voltage a PMOS switch can pass is V_{THP}.

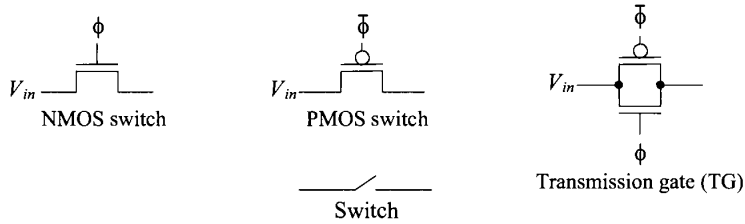


Figure 25.1 MOSFETs used as switches.

IPR2023-00698
Theta EX2019

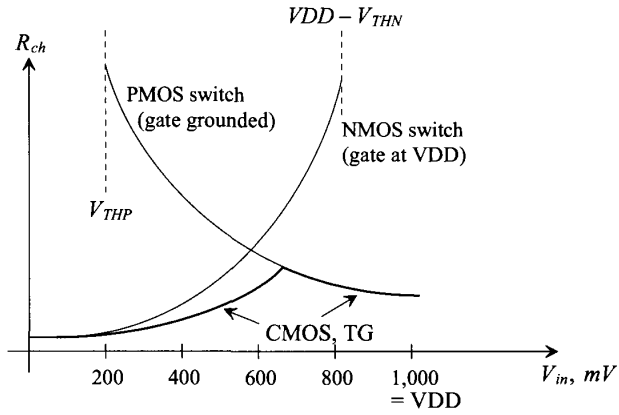


Figure 25.2 Small-signal on-resistance of MOSFET switches.

While MOS switches may offer substantial benefits, they are not without some detraction. Two nonideal effects typically associated with these switches may ultimately limit the use of MOS switches in some applications (particularly sampled-data circuits such as data converters). These two effects are known as *charge injection* and *clock feedthrough*.

Charge Injection

Charge injection can be understood with the help of Fig. 25.3. When the MOSFET switch is on and V_{DS} is small, the charge under the gate oxide resulting from the inverted channel is (from Ch. 6) Q'_{ch}. When the MOSFET turns off, this charge is injected onto the capacitor and into V_{in}. Because V_{in} is assumed to be a low-impedance, source-driven node, the injected charge has no effect on this node. However, the charge injected onto C_{load} results in a change in voltage across it. Note that we also have charge injection (in the opposite direction) when we turn the switch on. However, the fact that the input voltage is connected to C_{load} through the channel resistance makes this error unimportant (the voltage across C_{load} charges to V_{in} through the MOSFET's channel resistance).

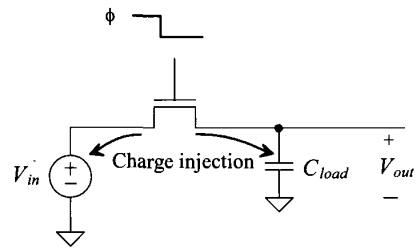


Figure 25.3 Simple configuration using an NMOS switch to show charge injection.

IPR2023-00698
Theta EX2019

Although the charge injection mechanism is itself a complex one, many studies have sought to characterize and minimize its effects. It has been shown that if the clock signal turns off fast, the channel charge distributes fairly equally between the adjacent nodes. Thus, half of the channel charge is distributed onto C_{load}. From Ch. 6, the charge/unit area of an inverted channel can be approximated as

$$Q_I'(y) = C_{ox}' \cdot (V_{GS} - V_{THN}) \tag{25.1}$$

The total charge in the channel must then be multiplied by the area of the channel resulting in

$$Q_I(y) = C_{ox}' \cdot W \cdot L \cdot (V_{GS} - V_{THN}) \tag{25.2}$$

Therefore, the change in voltage across C_{load} (if an NMOS switch is used) is

$$\Delta V_{load} = -\frac{C_{ox}' \cdot W \cdot L \cdot (V_{GS} - V_{THN})}{2C_{load}} \tag{25.3}$$

which can be written as

$$\Delta V_{load} = -\frac{C_{ox}' \cdot W \cdot L \cdot (VDD - V_{in} - V_{THN})}{2C_{load}} \tag{25.4}$$

if it is assumed that the clock swings between *VDD* and ground. The threshold voltage, Eq. (6.19), can also be substituted into Eq. (25.4) to form

$$\Delta V_{load} = -\frac{C_{ox}' \cdot W \cdot L \cdot (VDD - V_{in} - [V_{THN0} + \gamma(\sqrt{|2V_{fp}| + V_{in}} - \sqrt{|2V_{fp}|})])}{2C_{load}} \tag{25.5}$$

Note that Eq. (25.5) illustrates the problem associated with charge injection. The change in voltage across C_{load} is nonlinear with respect to V_{in} due to the threshold voltage. Thus, it can be said that if the charge injection is signal-dependent, harmonic distortion results. In sampled-data systems, charge injection results in nonlinearity errors. In the case where the charge injection is signal-independent, a simple offset occurs, which is much easier to manage than harmonic distortion. These will be discussed in more detail in Chs. 28 and 29, but it should be obvious here that charge injection effects should be minimized as much as possible.

Capacitive Feedthrough

Consider the schematic of the NMOS switch shown in Fig. 25.4. Here the capacitances between the gate/drain and gate/source of the MOSFET are modeled with the assumption that the MOSFET is operating in the triode region. When the gate clock signal, ϕ, goes high, the clock signal feeds through the gate/drain and gate/source capacitances. However, as the switch turns on, the input signal, V_{in}, is connected to the load capacitor through the NMOS switch. The result is that C_{load} is charged to V_{in} and the capacitive feedthrough has no effect on the final value of V_{out}. However, now consider what happens when the clock signal makes the transition low, that is, the n-channel MOSFET turns off. A capacitive voltage divider exists between the gate-drain (source) capacitance and the load capacitance. As a result, a portion of the clock signal, ϕ, appears across C_{load} as

$$\Delta V_{load} = \frac{C_{overlap} \cdot VDD}{C_{overlap} + C_{load}} \tag{25.6}$$

IPR2023-00698
Theta EX2019

where $C_{overlap}$ is the overlap capacitance value,

$$C_{overlap} = C'_{ox} \cdot W \cdot LD \tag{25.7}$$

and LD is the length of the gate that overlaps the drain/source.

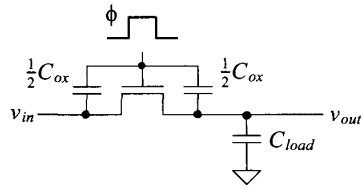


Figure 25.4 Illustration of capacitive feedthrough.

Reduction of Charge Injection and Clock Feedthrough

Many methods have been reported that reduce the effects of charge injection and capacitive feedthrough. One of the most widely used is the dummy switch, as seen in Fig. 25.5. Here, a switch, M2, with its drain and source shorted is placed in series with the desired switch M1. Notice that the clock signal controlling the dummy switch is the complement of the signal controlling M1, and in addition, should also be slightly delayed.

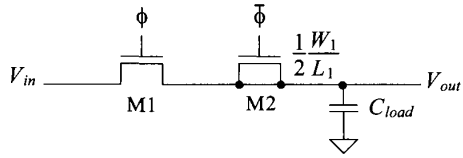


Figure 25.5 Dummy switch circuit used to minimize charge injection.

When M1 turns off, half of the channel charge is injected toward the dummy switch, thus explaining why the size of M2 is one-half that of M1. Although M2 is effectively shorted, a channel can still be induced by applying a voltage on the gate. Therefore, the charge injected by M1 is essentially matched by the charge induced by M2, and the overall charge injection is canceled. Note what happens when M2 turns off. It will inject half of its charge in both directions. However, because the drain and source are shorted and M1 is on, all of the charge from M2 will be injected into the low-impedance, voltage-driven source, which is also charging C_{load}. Therefore, M2's charge injection will not affect the value of voltage on C_{load}.

Another method for counteracting charge injection and clock feedthrough is to replace the switch with a CMOS transmission gate (TG). This results in lower changes in V_{out} because the complementary signals that are used will act to cancel each other.

IPR2023-00698
Theta EX2019

However, this approach requires precise control on the complementary clocks (the clocks must be switched at exactly the same time) and assumes that the input signal, V_{in}, is small, since the symmetry of the turn-on and turn-off waveforms depend on the input signal.

Fully-differential circuit topologies are used to cancel these effects to a first order, as seen in Fig. 25.6. Since the nonideal charge injection and clock feedthrough effects appear as a common-mode signal to the amplifier, they will be reduced by the CMRR of the amplifier. However, the second-order effects resulting from the input signal amplitude dependence will ultimately limit the dynamic range of operation, neglecting coupled and inherent noise in the dynamic circuits. This subject will be discussed in more detail in the following sections and in the next chapter.

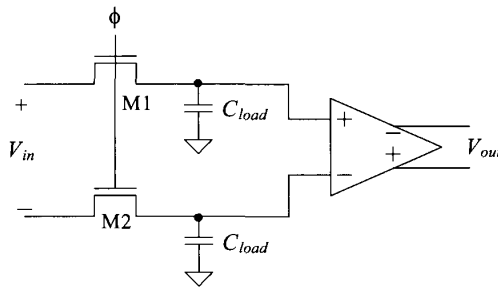


Figure 25.6 Using a fully-differential circuit to minimize charge injection and clock feedthrough.

kT/C Noise

In Ch. 8 we saw that the maximum RMS output noise generated from a simple *RC* circuit was $\sqrt{kT/C}$ (see Table 8.1). If we think of the MOSFET in Fig. 25.7 as a resistor (when the MOSFET is on), then we can add this RMS noise source in series with the output of the capacitor. The noise can be regarded as a sampled (random) voltage onto the capacitor each time the switch is turned on. The RMS noise generated, at room temperature, when using a 1 pF capacitor is 64 μV, while a 100 fF capacitor results in a noise voltage of 200 μV. In other words, the larger the capacitor, the smaller the noise voltage sampled on to the storage capacitor. For high-speed systems, it is desirable to use small capacitors since they take less time to charge. When designing a high-speed and low-noise circuit, trade-offs must be made when selecting the capacitor size.

Note, as mentioned above, the MOSFET is thought of as a resistor (and so *kT/C* noise in the circuit in Fig. 25.7 is from MOSFET thermal noise). If we review the noise mechanisms found in a MOSFET in Ch. 9, we also see that Flicker noise may be present. However, for a MOSFET's drain current to contain Flicker noise, the MOSFET must be conducting a DC current. Because the MOSFET in Fig. 25.7 doesn't conduct any DC current after C_H is charged, Flicker noise doesn't affect the sampled voltage.

IPR2023-00698
Theta EX2019

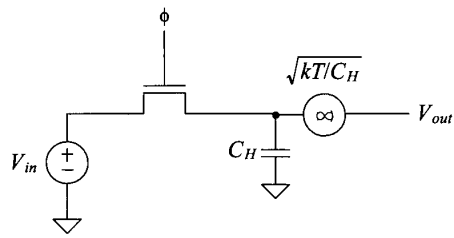


Figure 25.7 How kT/C noise adds to a sampled signal.

25.1.1 Sample-and-Hold Circuits

An important application of the switch is in the sample-and-hold circuit. The sample-and-hold circuit finds extensive use in data converter applications as a sampling gate. A variety of topologies exist, each with their own benefits. The simplest is shown in Fig. 25.8. A narrow pulse is applied to the gate of the MOSFET, enabling v_{in} to charge the hold capacitor, C_H. The width of the strobing gate pulse should allow the capacitor to fully charge before being removed. The op-amp simply acts as a unity gain buffer, isolating the hold capacitor from any external load. This circuit suffers from the clock feedthrough and charge injection problems mentioned in the previous discussion.

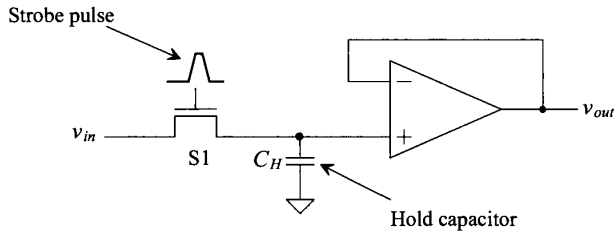


Figure 25.8 A basic sample-and-hold circuit (more correctly called a track-and-hold, see Fig. 28.5b).

A possible improvement in the basic S/H circuit is seen in Fig. 25.9. Here, two amplifiers buffer the input and the output. Notice that switch S_2 ensures that amplifier A1 is stable while in hold mode. If the switch were not present, A1 would be open loop during hold mode and would swing to one of the rails. During the next sample mode, it would then be slew-limited while going from the supply to the value of v_{in}. However, with the addition of S_2, the output of A1 tracks v_{in} even while in hold mode. The switch S_3 also disconnects A1 from the output during hold mode. This S/H has its disadvantages, however. The capacitor is still subjected to charge injection and clock-feedthrough problems. In addition, during sample mode, the circuit may become unstable since there are now two amplifiers in the single-loop feedback structure. Although compensation

IPR2023-00698
Theta EX2019

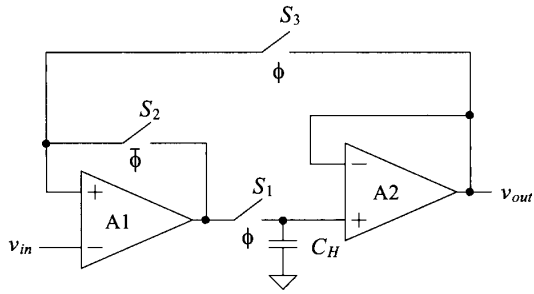


Figure 25.9 A closed-loop S/H circuit.

capacitors can be added to stabilize its performance, the size and placement of the capacitors depend solely on the type and characteristics of the op-amps.

Another S/H circuit is seen in Fig. 25.10. Here, a transconductance amplifier is used to charge the hold capacitor. A control signal turns the amplifier, A1, on or off digitally, thus eliminating the need for the switches S_2 and S_3 (from Fig. 25.9). Since CMOS op-amps are well suited for high-output impedance applications, this configuration would seem to be a popular one. However, the speed of this topology is dictated by the maximum current output of the transconductance amplifier and the size of the hold capacitor.

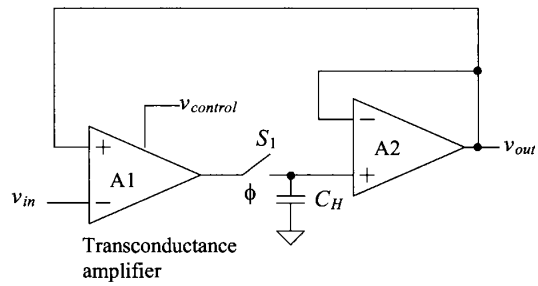


Figure 25.10 A closed-loop S/H circuit using a transconductance amplifier.

A third S/H circuit can be seen in Fig. 25.11. The advantage of this circuit may not be completely obvious and warrants further explanation. First, notice that the hold capacitor is actually in the feedback path of the amplifier, A2, with one side connected to the output of the amplifier and the other connected to a virtual ground. When switch S_1 turns off, any charge injected onto the hold capacitor results in a slight change in the output voltage. However, now that one side of the switch is at virtual ground, the change in voltage is no longer dependent on the threshold voltage of the switch itself. Therefore,

IPR2023-00698
Theta EX2019

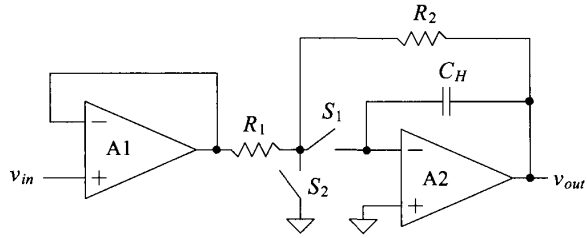


Figure 25.11 A closed-loop S/H circuit using a transconductance amplifier.

the charge injection will be independent of the input signal and will result as a simple offset at the output. An offset error is much easier to tolerate than a nonlinearity error, as will be seen in Chs. 28 and 29.

When sampling, S_1 is closed and S_2 is open, and the equivalent circuit is simply a low-pass filter with a buffered input. The overall transfer function becomes

$$\frac{v_{out}}{v_{in}} = -\frac{R_2}{R_1} \cdot \frac{1}{(sR_2C_H + 1)} \tag{25.8}$$

Therefore, this circuit performs a low-pass filter function while sampling. The buffer A1 can be eliminated when we desire a low-input impedance. Once hold mode commences, the output will stay constant at a value equal to v_{in}, while the switch S_2 isolates the input from the hold capacitor. One important issue to note here is that A2 will need to be a buffered CMOS amplifier because of the resistive load attached at v_{out} during hold mode. Notice also that during both sample mode and hold mode, there is only one op-amp in each feedback loop, so this S/H topology is much more stable than the closed-loop structure introduced in Fig. 25.10.

25.2 Fully-Differential Circuits

As we saw in Fig. 25.6 and the associated discussion, using a fully-differential topology (an op-amp with both differential inputs and outputs) can reduce the effects from imperfect switches. However, using a fully-differential topology requires the use of a *common-mode-feedback* (CMFB) circuit. In the following we present an overview of fully-differential op-amps and their application in a sample-and-hold circuit.

Gain

The differential output op-amp symbol is shown in Fig. 25.12. The open-loop gain of the op-amp is related to its inputs and outputs by

$$v_{out} = v_{op} - v_{om} = A_{OL} \cdot (v_p - v_m) \tag{25.9}$$

This should be compared to the single-ended output op-amp (the op-amp we have been discussing up to this point), which has a gain given by

$$v_{out} = v_{op} = A_{OL} \cdot (v_p - v_m) \tag{25.10}$$

IPR2023-00698
Theta EX2019

If we ignore v_{om}, then the differential output op-amp behaves just like a single-ended op-amp. For linear applications, the op-amp is used with a feedback network, the inputs are related (assuming the open-loop gain, A_{OL} is large) by

$$v_p \approx v_m \tag{25.11}$$

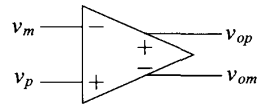


Figure 25.12 Differential output op-amp.

Common-Mode Feedback (CMFB)

Normally, the average of the op-amp outputs is called the common-mode output voltage. The two op-amp outputs swing around the common-mode voltage. If the largest op-amp output voltage is *VDD* and the minimum output voltage is ground, then the common-mode voltage is

$$V_{CM} = \frac{VDD}{2} \tag{25.12}$$

Figure 25.13a shows a simple differential output op-amp gain configuration (inverting or noninverting, depending on which output is used as the positive output). Because the input voltages to the circuit are equal, the output voltages of the op-amp should remain at V_{CM}. Due to the op-amp's high gain and the negative feedback, we always have $v_p \approx v_m$ (Eq. [25.11]). However, if $v_{op} = v_{om} = VDD$ or $v_{op} = v_{om}$ = ground or $v_{op} = v_{om}$ = anything, then $v_p = v_m$ and Eq. (25.11) is satisfied. *This is a problem* and the reason why we need a CMFB circuit. In Fig. 25.13b, we add a CMFB circuit to monitor the outputs of the op-amp and make adjustments in the op-amp (how these adjustments are made is discussed in the next chapter) to keep the two outputs balanced around V_{CM} (VDD/2).

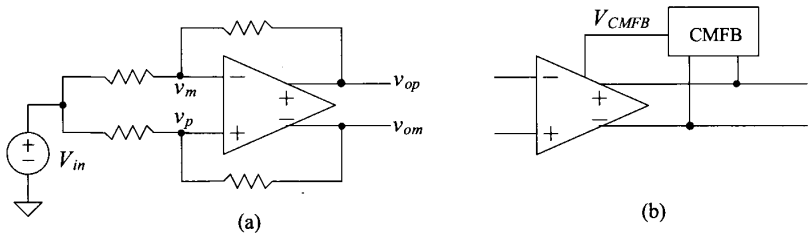


Figure 25.13 (a) Simple gain configuration using differential output op-amp and (b) the use of a common-mode feedback circuit to adjust the common-mode output voltage.

IPR2023-00698
Theta EX2019

Coupled Noise Rejection

Differential topologies are required to minimize the effects of charge injection and capacitive feedthrough from switches. They also are used to help reduce the effects of coupled noise. Consider the cascade of differential output op-amps, with no feedback network shown, seen in Fig. 25.14. The stray capacitance between the interconnecting metal lines (the metal lines that carry the signals between op-amps) and the substrate or any other noise source are shown. If the metal lines are run close to one another, then the noise voltage will couple even amounts (ideally) of noise into each signal wire. Since the diff-amp, on the input of the op-amp, rejects common signals (signals that are present on both inputs), the coupled noise is not passed to the next op-amp in the string. Variations on the power supply rails are rejected as well. If the differential op-amp is symmetric, then changes on the power supply couple evenly into both outputs, having little effect on the difference, or desired signal, coming out of the op-amp. For these reasons, that is, good coupled noise rejection and PSRR, the differential op-amp is a necessity in any dynamic analog integrated circuit.

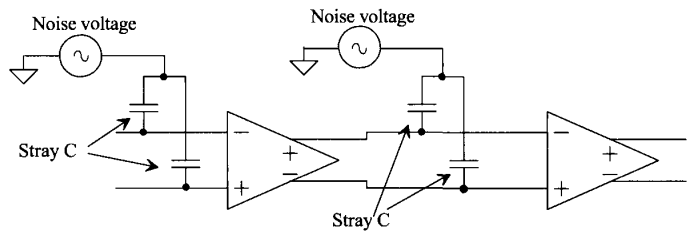


Figure 25.14 Differential output op-amps showing parasitic capacitance and noise.

Other Benefits of Fully-Differential Op-Amps

Another benefit of fully-differential circuits is a doubling in output voltage swing. For example, if *VDD* is 1 V, then the output of a single-ended op-amp can swing from 1 V to ground. In a fully-differential topology, both v_{op} and v_{om} can swing from 0 to 1. Using Eq. (25.9), we can write

$$v_{out,\max} = v_{op} - v_{om} = 1 - 0 = 1\ V \text{ and } v_{out,\min} = v_{op} - v_{om} = 0 - 1 = -1\ V \quad (25.13)$$

The output swing is 2*VDD* when using fully-differential topologies.

A further benefit of fully-differential topologies is the fact that the input common-mode voltage of the op-amp remains at V_{CM} (and so the op-amp's first stage diff-amp's common-mode voltage range requirements are easy to meet).

25.2.1 A Fully-Differential Sample-and-Hold

Figure 25.15 shows a fully-differential sample-and-hold circuit and the associated clock waveforms that eliminate clock feedthrough and charge injection to a first order. The switches in this figure are closed when their controlling clock signals are high. The basic operation can be understood by considering the state of the circuit at t_0. At this time, the

IPR2023-00698
Theta EX2019

input signals charge the sampling capacitors. The bottom plates of the capacitors (poly1) are tied directly to the input signals, for reasons that will be explained below. The op-amp is operating in a unity-follower configuration in which both inputs of the op-amp are held at V_{CM}. At this particular instance in time, prior to t_1, the amplifier is said to be operating in the sample mode of operation.

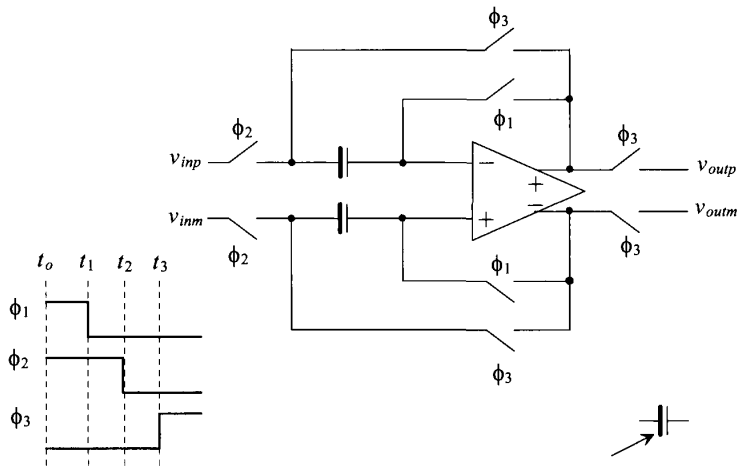


Figure 25.15 Sample-and-hold using differential topology.

At t_1, the ϕ_1 switches turn off. The resulting charge injection and clock feedthrough appear as a common-mode signal on the inputs of the op-amp and are ideally rejected. Since the top plates of the hold capacitors (the inputs to the op-amp) are always at V_{CM}, at this point in time the charge injection and clock feedthrough are independent of the input signals. This produces an increase in the dynamic range of the sample-and-hold (the minimum measurable input signal decreases). The voltage on the inputs of the op-amp (the top plate of the capacitor) between t_1 and t_2 is $V_{OFF1} + V_{CM}$, a constant voltage. Note that the op-amp is operating open loop at this time so the time between t_1 and t_3 should be short.

At t_2 the ϕ_2 switches turn off. At this point in time, the voltages on the bottom plates of the sampling (or hold) capacitors (poly1) are v_{inp} and v_{inm}. The voltages on the top plates of the capacitors (connected to the op-amp) are $V_{OFF1} + V_{OFF2} + V_{CM}$ (assuming that the storage capacitors are much larger than the input capacitance of the op-amp). The term V_{OFF2} is ideally a constant that results from the charge injection and capacitive feedthrough from the ϕ_2 switches turning off. The time between t_1 and t_2 should be short compared to variations in the input signals.

IPR2023-00698
Theta EX2019

At time t_3 the ϕ_3 switches turn on and the op-amp behaves like a voltage follower; the circuit is said to be in the hold mode of operation. The charge injection and clock feedthrough resulting from the ϕ_3 switches turning on causes the top plate of the capacitor to become $V_{OFF1} + V_{OFF2} + V_{OFF3} + V_{CM}$, again assuming that the storage capacitors are much larger than the input capacitance of the op-amp. The outputs of the sample-and-hold are v_{inp} and v_{inm}, assuming infinite op-amp gain since these offsets appear as a common-mode voltage on the input of the op-amp. Note that the terms V_{OFF2} and V_{OFF3} are dependent on the input signals.

Connecting the Inputs to the Bottom (Poly1) Plate

The reason for connecting the input signals to the bottom plate of the capacitor can be explained with the help of Fig. 25.16. This figure is a simplified, single-ended version of Fig. 25.15 where the capacitance, C_p, is the parasitic capacitance from the bottom plate to the substrate. With regard to Fig. 25.16a, coupled noise from the substrate sees either the input voltage from the op-amp driving the sample-and-hold (in the sample mode) or the output voltage of the op-amp used in the sample-and-hold itself (in the hold mode). Since the op-amps in either mode directly set this voltage, the substrate noise has, ideally, little effect on the circuit's operation.

In Fig. 25.16b, coupled substrate noise feeds directly into the input of the op-amp and can thus drastically affect the output of the sample-and-hold. Another more subtle problem occurs in the circuit of Fig. 25.16b. When the circuit makes the transition to the hold mode at t_3, the output of the op-amp should quickly change to the voltage sampled on the input capacitors. The time it takes the output of the op-amp to change and settle to this final voltage is called the *settling time*. The parasitic capacitance on the input of the op-amp in (b) reduces the feedback factor from unity to $C_H/(C_p + C_H)$. This slows the settling time and causes a gain error in the circuit's transfer function. For these reasons, the parasitic capacitance on the top plate of the capacitor should be small. Nothing should be laid out over or in near proximity of poly2.

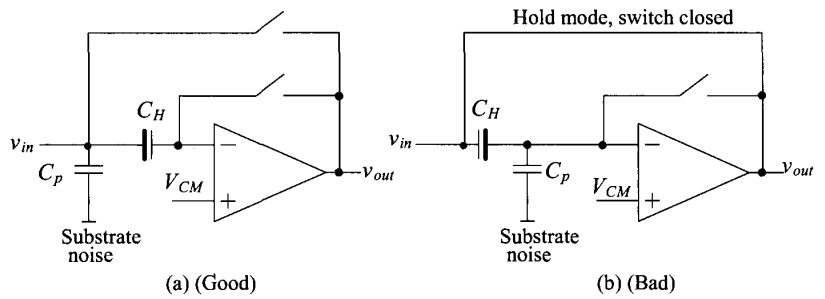


Figure 25.16 Explanation for connecting the bottom plate of the capacitor to the input.

IPR2023-00698
Theta EX2019

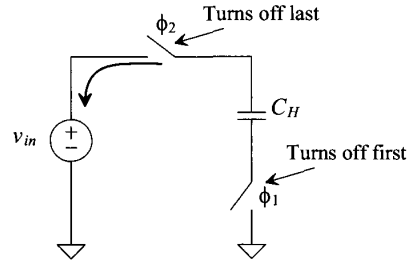


Figure 25.17 Bottom plate sampling.

Bottom Plate Sampling

Turning the ϕ_1 controlled switches off in Fig. 25.15 slightly before the ϕ_2 controlled switches is sometimes called *bottom plate sampling*. Figure 25.17 illustrates why. When the ϕ_1 switches turn off, the charge injected into C_H is a constant independent of the input signal amplitude. The resulting offset voltage across C_H is then constant. When the ϕ_2 switches turn off, the resulting charge injection then takes the path of least resistance, that is, into the input source v_{in}. The voltage across C_H is then, ideally, independent of the input signal voltage.

SPICE Simulation

Let's simulate the operation of the sample-and-hold in Fig. 25.15. To begin, let's use voltage-controlled voltage sources and a DC voltage to model the op-amp (again, we'll discuss differential output op-amps in the next chapter). The ideal fully-differential output op-amp SPICE model is seen in Fig. 25.18. The SPICE netlist statements used to model this circuit (we'll assume an open-loop gain of 10^6 and *VDD* of 1 *V*) are

```
E1    vop   vcm   vp    vm    1e6
E2    vcm   vom   vp    vm    1e6
Vcm   vcm   0     DC    500m
```

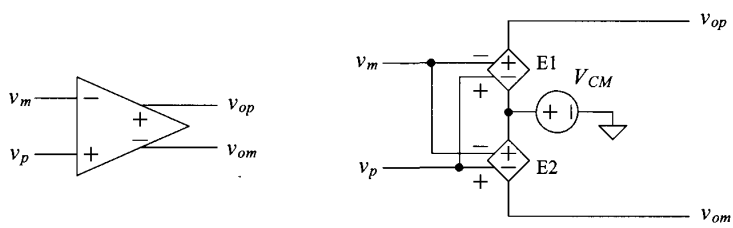


Figure 25.18 SPICE modeling a differential input/output op-amp with common-mode voltage.

IPR2023-00698
Theta EX2019

For the switches we'll use NMOS devices as seen in the simulation schematic of Fig. 25.19. For the input signals, we've used a 5 MHz sinewave signal with a peak amplitude of 200 mV. Note that each of the inputs is referenced to the common-mode voltage, V_{CM} (which is 500 mV here).

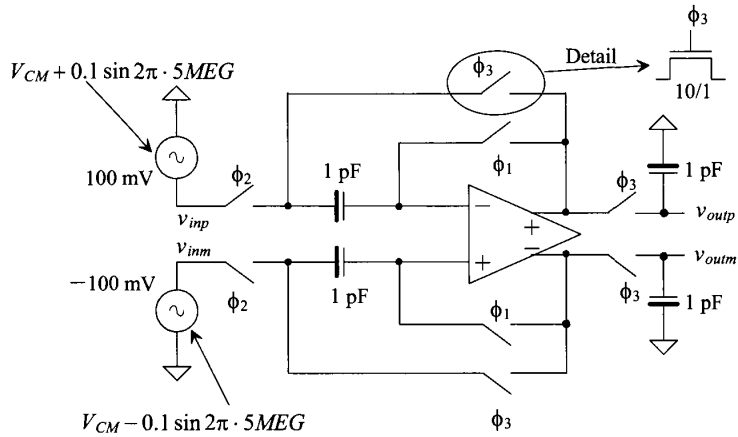


Figure 25.19 Simulating the operation of the fully differential sample-and-hold.

The simulation results are seen in Fig. 25.20. Let's first focus on the clock signals. Notice how ϕ_3 isn't high at the same time as ϕ_1 or ϕ_2. If all switches were "on" at the same time, the charge stored (or held) on the 1 pF capacitors at the output of the circuit would change (remember when the ϕ_1 switches close the op-amp is in the follower configuration with the inputs and outputs pulled to V_{CM}). We make sure to disconnect these output capacitors before putting the circuit into the sample mode. The output of the sample-and-hold seen in Fig. 25.20 doesn't exactly match the input signal. There is a one-clock cycle delay and we can see the effects of the finite clock frequency (finite sample points). Note, these topics are discussed in detail in the book entitled *CMOS Mixed-Signal Circuit Design.*

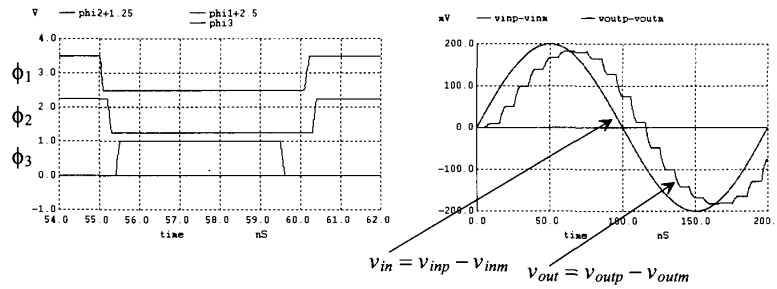


Figure 25.20 Simulating the operation of the sample-and-hold in Fig. 25.19.

IPR2023-00698
Theta EX2019

25.3 Switched-Capacitor Circuits

Consider the circuit shown in Fig. 25.21a. This dynamic circuit, named a *switched-capacitor resistor*, is useful in simulating a large value resistor, generally >1 MΩ. The clock signals ϕ_1 and ϕ_2 form two phases of a nonoverlapping clock signal with frequency f_{clk} and period T as seen in the figure. Let's begin by considering the case when S1 is closed. When ϕ_1 is high, the capacitor C is charged to v_1. The charge, q_1, stored on the capacitor during this interval, Fig. 25.21c, is

$$q_1 = Cv_1 \tag{25.14}$$

while if S2 is closed, the charge stored on the capacitor is

$$q_2 = Cv_2 \tag{25.15}$$

If v_1 and v_2 are not equal, keeping in mind that S1 and S2 cannot be closed at the same time due to the nonoverlapping clock signals, then a charge equal to the difference between q_1 and q_2 is transferred between v_1 and v_2 during each interval T. The difference in the charge is given by

$$q_1 - q_2 = C(v_1 - v_2) \tag{25.16}$$

If v_1 and v_2 vary slowly compared to f_{clk}, then the average current transferred in an interval T is given by

$$I_{avg} = \frac{C(v_1 - v_2)}{T} = \frac{v_1 - v_2}{R_{sc}} \tag{25.17}$$

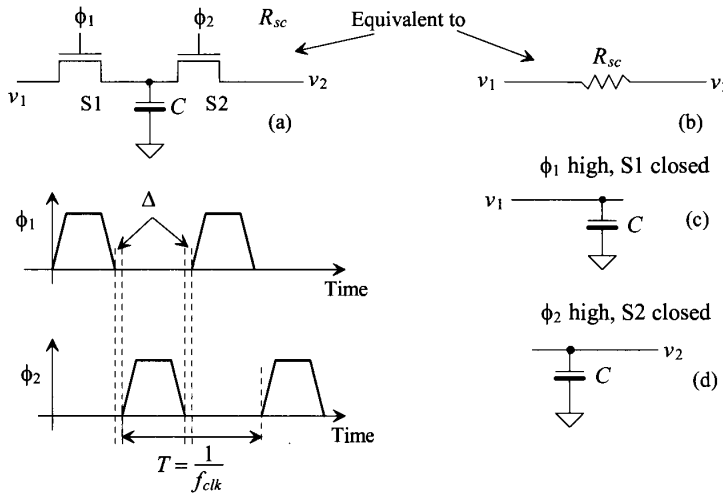


Figure 25.21 Switched-capacitor resistor (a) and associated waveforms and (b, c, d) the equivalent circuits.

IPR2023-00698
Theta EX2019

The resistance of the switched-capacitor circuit is given by

$$R_{sc} = \frac{T}{C} = \frac{1}{C \cdot f_{clk}} \tag{25.18}$$

In general, the signals v_1 and v_2 should be bandlimited to a frequency at least ten times less than f_{clk} (more on this later). Note that we derived a similar result back in Ch. 17, see Eq. (17.29).

Example 25.1

Using switched-capacitor techniques, implement the circuit shown in Fig. 25.22a so that the product of RC is 1 ms, that is, the 3-dB frequency of $|v_{out}/v_{in}|$ is 159 Hz.

The switched-capacitor implementation of this circuit is shown in Fig. 25.22b. The product of RC may now be written in terms of Eq. (25.18) as

$$RC_2 = \frac{C_2}{C_1} \cdot \frac{1}{f_{clk}} \tag{25.19}$$

This result is important! The product RC_2 is determined by f_{clk}, which may be an accurate frequency derived from a crystal oscillator and the ratio of C_2 to C_1, which will be within 1% on a chip. This means that even if the values of the capacitors change by 20% from wafer to wafer, the ratio of the capacitors relative to one another, on the same wafer, will remain constant within 1%.

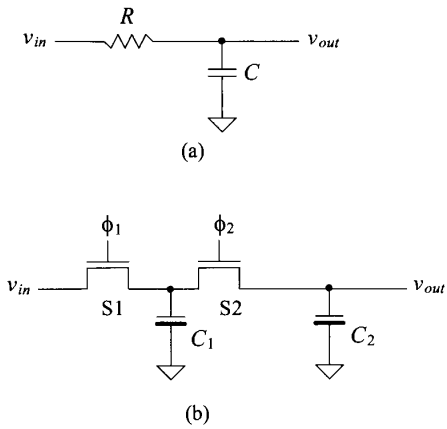


Figure 25.22 (a) Circuit used in Ex. 27.1 and (b) its implementation using a SC resistor.

It is also desirable to keep C_1 larger than the associated parasitics present in the circuit (e.g., depletion capacitances of the source/drain implants and the stray capacitances to substrate). For the present example, we will set C_1 to 1 pF. The selection of f_{clk} is usually determined by what is available. For the present design, a value of 100 kHz will be used. This selection assumes that the energy present in v_{in} at frequencies above 10 kHz is negligible. The value of C_2 is determined solving Eq. (25.19) and is 100 pF. Note that the value of the switched-capacitor

IPR2023-00698
Theta EX2019

resistor is 10 MΩ. Implementing this resistor in a CMOS process using n-well with a sheet resistance of 1,000 ohms/square would require 10,000 squares! The resulting delay through the n-well resistor, because of the capacitance to substrate, may cause a significant phase error in the transfer function. ■

25.3.1 Switched-Capacitor Integrator

Because the switched-capacitor resistor of Fig. 25.21a is sensitive to parasitic capacitances, it finds little use, by itself, in analog switched-capacitor circuits. Consider the circuit of Fig. 25.23a. This circuit, a switched-capacitor integrator, is the heart of the circuits we will be discussing in the remainder of the section. The portion of the circuit consisting of switches S1 through S4 and C_I forms a switched-capacitor resistor with a value given by

$$R_{sc} = \frac{1}{C_I f_{clk}} \tag{25.20}$$

The equivalent continuous time circuit for the switched-capacitor integrator is shown in Fig. 25.23b. Notice that v_{in} is now negative. It may be helpful in the following discussion to remember that the combination of switches and C_I of the switched-capacitor integrator can be thought of as a simple resistor. The transfer function of the switched-capacitor integrator is given by

$$\frac{v_{out}}{v_{in}} = \frac{1/j\omega C_F}{R_{sc}} = \frac{1}{j\omega\left(\frac{C_F}{C_I} \cdot \frac{1}{f_{clk}}\right)} \tag{25.21}$$

Again, the ratio of capacitors is present, allowing the designer to precisely set the gain of the amplifier and the integration time constant.

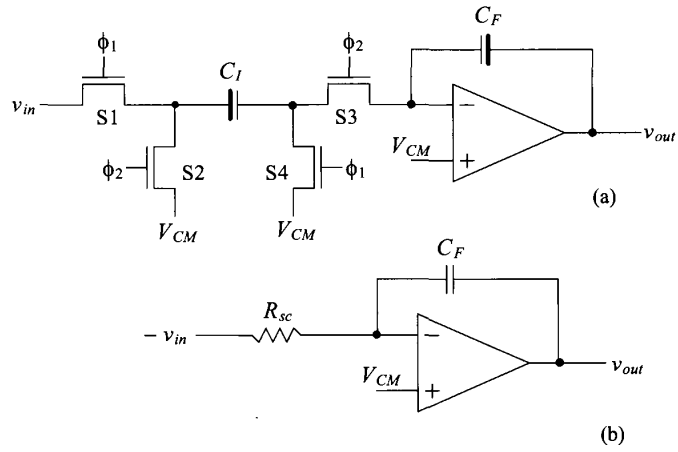


Figure 25.23 (a) A stray insensitive switched-capacitor integrator (noninverting) and (b) the equivalent continuous time circuit.

IPR2023-00698
Theta EX2019

Parasitic Insensitive

The switched-capacitor integrator in Fig. 25.23 is not sensitive to parasitic or stray capacitances. This can be understood with the use of Fig. 25.24. To begin, if we realize that C_{p2} (the parasitic capacitance on the right side of C_I) is always connected to V_{CM} either through S4 or through the connection to the inverting input of the op-amp, then C_{p2} doesn't see a change in the charge stored on it. Next, the capacitance C_{p1} is charged to v_{in} when S1 is closed and then charged to V_{CM} when S2 closes. Since none of the charge stored on C_{p1} when S1 is closed is transferred to C_I, it does not affect the integrating function. A practical minimum for C_I is 100 fF set by kT/C noise (see Table 8.1).

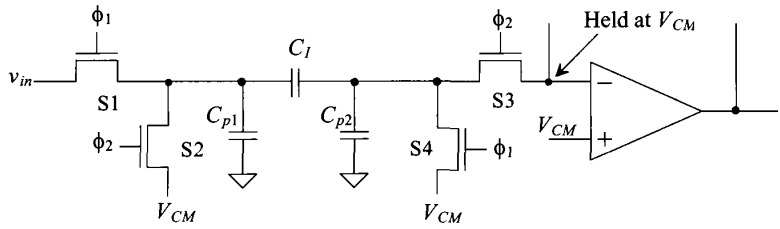


Figure 25.24 Parasitic capacitances associated with a switched-capacitor resistor.

Other Integrator Configurations

An inverting integrator configuration can be formed by simply swapping the clock signals used with S1 and S2 in Fig. 25.23 (the noninverting configuration). The gate of S1 is connected, for the inverting configuration, to ϕ_2 while the gate of S2 is connected to ϕ_1. The gain of this configuration is given by

$$\frac{v_{out}}{v_{in}} = -\frac{1}{j\omega\left(\frac{C_F}{C_I}\cdot\frac{1}{f_{clk}}\right)} \tag{25.22}$$

An example of a switched-capacitor integrator circuit that combines input signals is shown in Fig. 25.25a. Remembering that each switched-capacitor section can be thought of as a resistor, we note that the relationship between the inputs and output is

$$v_{out} = \frac{v_1}{j\omega\left(\frac{C_F}{C_1}\frac{1}{f_{clk}}\right)} + \frac{v_2}{j\omega\left(\frac{C_F}{C_2}\frac{1}{f_{clk}}\right)} - \frac{v_3}{j\omega\left(\frac{C_F}{C_3}\frac{1}{f_{clk}}\right)} \tag{25.23}$$

Figure 25.25b shows how redundant switches can be combined to reduce the number of devices used. Used alone, the basic integrator has the practical problem of integrating not only the input signal but also the offset voltage of the op-amp. In many applications, a reset switch or resistor is placed across the feedback capacitor (Fig. 25.25b). An example of a lossy integrator circuit useful in first-order filter design is shown in Fig. 25.26. The transfer function of this circuit is given by

$$\frac{v_{out}}{v_{in}} = \frac{R_4}{R_3}\left(\frac{1+j\omega R_3 C_1}{1+j\omega R_4 C_2}\right) = \frac{C_3}{C_4}\left(\frac{1+j\omega\left(\frac{C_1}{C_3}\cdot\frac{1}{f_{clk}}\right)}{1+j\omega\left(\frac{C_2}{C_4}\cdot\frac{1}{f_{clk}}\right)}\right) \tag{25.24}$$

IPR2023-00698
Theta EX2019

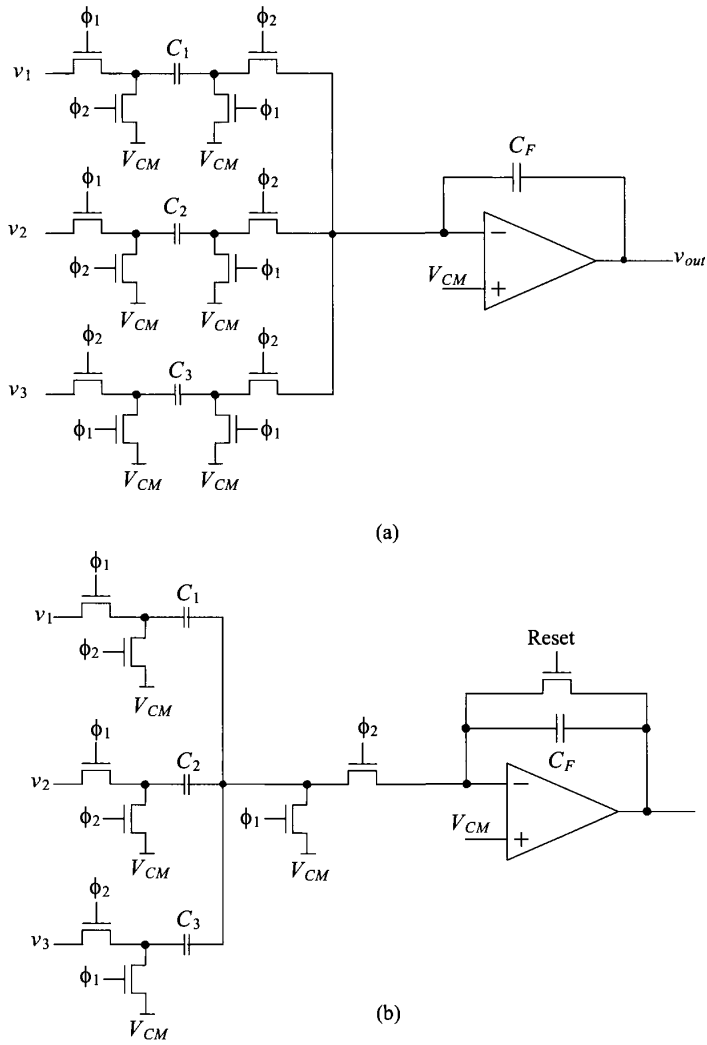


Figure 25.25 (a) Switched-capacitor implementation of a summing integrator and (b) practical implementation of the circuit combining switches and adding reset.

IPR2023-00698
Theta EX2019

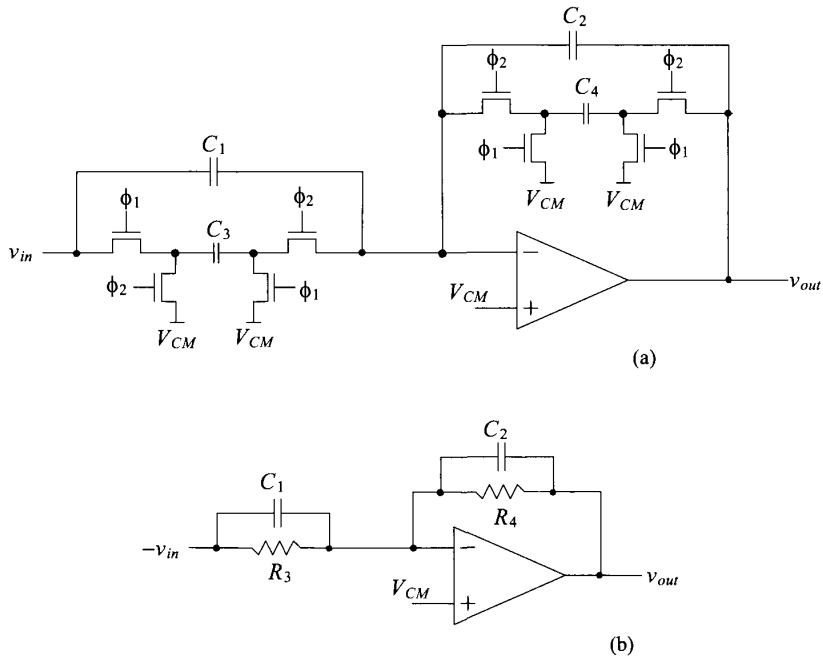


Figure 25.26 Lossy integrator (a) switched-capacitor implementation and (b) continuous time circuit.

For low-frequency input signals (low frequencies compared to the pole and zero given in Eq. (25.24), the gain of the lossy integrator is simply

$$\frac{v_{out}}{v_{in}} = \frac{C_3}{C_4} \tag{25.25}$$

which is again a precise number due to the ratio of the capacitors. Also note that the switched-capacitor resistor in the feedback loop is stray insensitive. The left side of C_4 is always connected to V_{CM}, while the right side is either connected to V_{CM} or to the output of the op-amp.

Note that we cannot eliminate the capacitor across the switched-capacitor resistor in the feedback path. If we were to do so, then the op-amp would be operating open-loop when ϕ_2 goes low. The outputs of the op-amp would then rail up at *VDD* or down at ground.

Example 25.2

Design a switched-capacitor filter with the transfer characteristics shown in Fig. 25.27.

IPR2023-00698
Theta EX2019

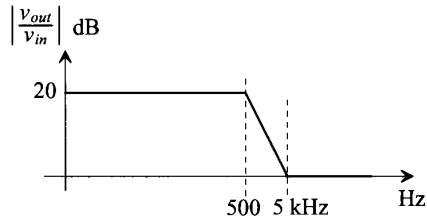


Figure 25.27 Filter characteristics for Ex. 25.2.

We can see that this transfer function has a pole at 500 Hz and a zero at 5 kHz. The lossy integrator of Fig. 25.26 will be used to realize this filter. The low-frequency gain of this circuit is 10 (20 dB). Using Eq. (25.25), we have

$$\frac{C_3}{C_4} = 10$$

while the pole and zero locations are given by

$$f_p = \frac{1}{2\pi\left(\frac{C_2}{C_4} \cdot \frac{1}{f_{clk}}\right)} = 500 \text{ and } f_z = \frac{1}{2\pi\left(\frac{C_1}{C_3} \cdot \frac{1}{f_{clk}}\right)} = 5 \text{ kHz}$$

If we set f_{clk} to 100 kHz and C_4 to 100 fF, then $C_3 = 1.0$ pF, $C_2 = 3.2$ pF, and $C_1 = 3.2$ pF. ■

Exact Frequency Response of a Switched-Capacitor Integrator

We will now develop an exact relationship between the switching frequency, f_{clk}, and the signal frequency ω. Referring to Fig. 25.28, we can write the output of the integrator as the sum of the previous output voltage, $v_{out(n)}$, at a time nT and the contribution from the current sample as

$$v_{out(n+1)} = v_{out(n)} + \frac{C_I}{C_F} \cdot v_{in(n)} \tag{25.26}$$

Since a delay in the time domain of T corresponds to a phase shift of ωT in the frequency domain, we can take the Fourier transform of this equation and get

$$e^{j\omega T} v_{out}(j\omega) = v_{out}(j\omega) + \frac{C_I}{C_F} \cdot v_{in}(j\omega) \tag{25.27}$$

Solving this equation for v_{out}/v_{in} gives

$$\frac{v_{out}}{v_{in}}(j\omega) = \frac{C_I}{C_F}\left(\frac{1}{e^{j\omega T} - 1}\right) = \frac{C_I}{C_F}\left(\frac{e^{-j\omega T/2}}{e^{j\omega T/2} - e^{-j\omega T/2}}\right) = \frac{C_I}{C_F}\left[\frac{1}{z - 1}\right] \tag{25.28}$$

where $z = e^{j\omega T}$. Remembering $f_{clk} = 1/T$ and $\omega = 2\pi f$, we get

$$\frac{v_{out}}{v_{in}}(j\omega) = \frac{1}{j\omega\left(\frac{C_F}{C_I} \cdot \frac{1}{f_{clk}}\right)}\left(\frac{\frac{\pi \cdot f}{f_{clk}}}{\sin\frac{\pi \cdot f}{f_{clk}}} \cdot e^{-j\pi \cdot f/f_{clk}}\right) \tag{25.29}$$

IPR2023-00698
Theta EX2019

Ideally, the term on the right in parentheses is unity. This occurs when f is much less than f_{clk}. This equation describes how the magnitude and phase of the integrator are affected by finite f_{clk}.

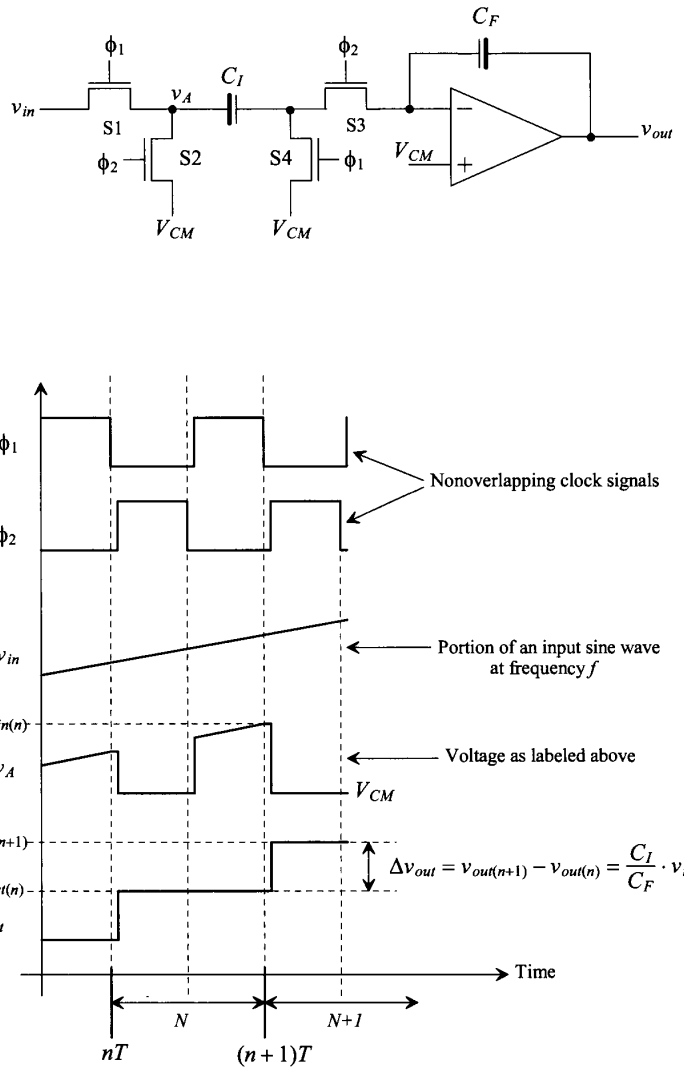


Figure 25.28 Switched-capacitor integrator used to determine the relationship between input frequency and switch clock frequency.

IPR2023-00698
Theta EX2019

Capacitor Layout

An important step in the implementation of any switched-capacitor design is the layout of the capacitors. Normally, a unit-size capacitor is laid out and then replicated to the desired capacitance (as discussed in Ch. 6). For Ex. 25.2, the unit size capacitance would nominally be 100-fF (using the 1 μm process) or 10 *fF* (in the 50 nm process), as in Fig. 25.29. Note that here, in Fig. 25.29a, we are assuming that a circle can be accurately reproduced on the reticle and patterned on the wafer. In practice, effects such as the finite e-beam size (and the granularity of the grid) used to make the reticle can make this assumption questionable. Figure 25.29b shows a layout where we've tried to minimize the number of 90° corners in an effort to avoid pattern errors when etching poly2.

Again, the absolute value of the capacitors isn't important; rather, the important value is the ratio. A total of 32 of these unit-size capacitors (using a unit-size cell of 100 *fF*) would be used to achieve the larger nominally 3.2 pF capacitors in Ex. 25.2 (see Fig. 25.30). This approach (using unit elements to make the large capacitors) eliminates errors due to uneven patterning of poly to a first order. A p+ guard ring can be placed around the capacitor to help reduce coupled substrate noise. Substrate noise can also be reduced by laying the capacitor out over an n-well that is tied to *VDD*. Injected minority carriers are collected either by the p+ or the n-well (or a combination of both). If matching of the capacitors is critical, schemes that use a common-centroid layout can be used. Also, as was discussed in Ch. 20, dummy poly strips or capacitors can be placed around the array of capacitors so that the edge differences from underetching poly are eliminated. In both cases, what we are doing is trying to ensure that all capacitors see the same adjacent structures.

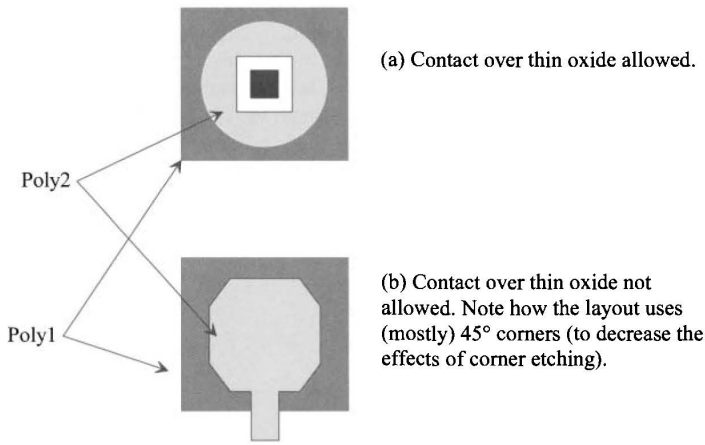


Figure 25.29 Layout of a unit cell capacitor.

IPR2023-00698
Theta EX2019

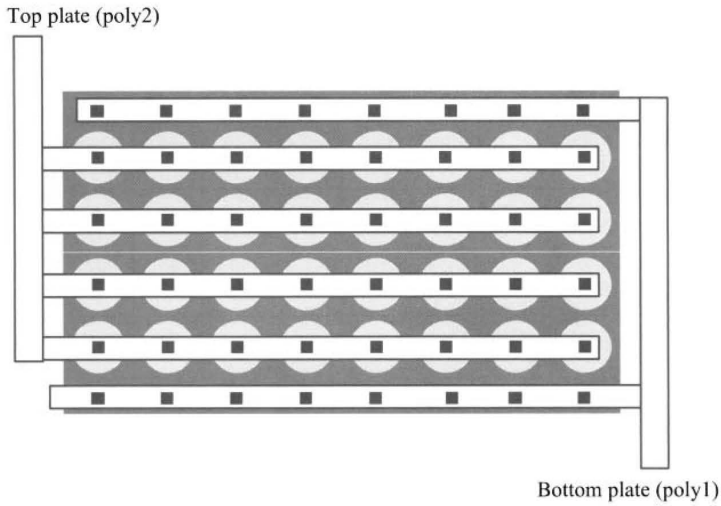


Figure 25.30 Layout of a 3.2 pF capacitor using a 100 fF unit cell.

Op-Amp Settling Time

Figure 25.31 shows an op-amp configuration in which the op-amp can source or sink current to a switched capacitor and a feedback capacitor. The time it takes to charge and discharge these capacitors is important because it directly affects the maximum switched capacitor clocking frequency, f_{clk}. The slew-rate limitations of the op-amp have been discussed in detail already. Let's now consider the limitations due to op-amp finite bandwidth. The closed-loop gain of the op-amp is given by (see Eq. [24.5])

$$A_{CL} = \frac{A_{OL}}{1 + A_{OL} \cdot \beta} \tag{25.30}$$

while the open-loop gain of an op-amp is given, with units of A/A, V/V, V/A, or A/V, by

$$A_{OL} = \frac{A_{OL}(0)}{1 + j\frac{f}{f_{3dB}}} \tag{25.31}$$

Combining these equations and assuming $1 >> 1/[\beta \cdot A_{OL}(0)]$, we get

$$A_{CL} = \frac{\frac{1}{\beta}}{1 + j\frac{f}{f_{un} \cdot \beta}} \tag{25.32}$$

where $f_{3dB} \cdot A_{OL}(0) = f_{un}$, where f_{un} may have units of Hz, Hz/Ω, or Hz·Ω ($f_u\beta$ is in Hz). The closed-loop gain reduces to a simple single-pole transfer function (see Eqs. [24.30]–[24.34]). The low-frequency gain of the circuit is $1/\beta$, while the product of f_{un} and β gives the circuit time constant of

IPR2023-00698
Theta EX2019

$$\tau = \frac{1}{2\pi f_{un} \cdot \beta} \tag{25.33}$$

keeping in mind the unity-gain frequency of the op-amp, f_{un}, is a strong function of the load capacitance. For a step-input to the op-amp, a common occurrence in switched-capacitor circuits (Fig. 25.28), the output voltage of the op-amp, again neglecting slew-rate limitations, is given by

$$v_{out} = V_{outfinal}(1 - e^{-t/\tau}) \tag{25.34}$$

For the output voltage of the op-amp to settle to less than 1% of its final value requires 5τ. A "rule-of-thumb" estimate for the settling time is simply $1/(f_u \cdot \beta)$.

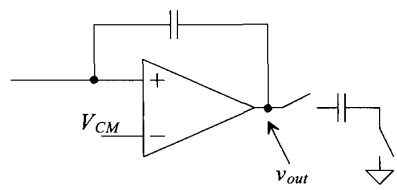


Figure 25.31 Charging and discharging a switched capacitor.

25.4 Circuits

This section presents several examples of dynamic analog circuits.

Reducing Offset Voltage of an Op-Amp

As seen in Fig. 24.4, the op-amp's offset voltage can be modeled by adding a DC voltage in series with the noninverting input of the op-amp, Fig. 25.32a. The basic idea behind eliminating the offset voltage is shown in Fig. 25.32b. A capacitor is charged to a voltage equal and opposite to the comparator offset voltage. The voltage across the capacitor is then added in series with the noninverting op-amp input to subtract away the op-amp's offset.

The dynamic analog circuit shown in Fig. 25.33 is used to implement this subtraction. For this method to be effective, *the op-amp must be stable* in the unity gain configuration. The clock signals ϕ_1 and ϕ_2 are the nonoverlapping clock signals discussed earlier (see Fig. 25.28). The nonoverlapping clocks keep switches S1, S2, and S3 from being on at the same time as switches S4 and S5. Let's consider the case shown in Fig. 25.33b where ϕ_1 is high and ϕ_2 is low. The op-amp, via the negative feedback, tries to force its inverting input to V_{CM}. However, because of the offset, the inverting input is actually charged to $V_{OS} + V_{CM}$. Note that under these conditions the op-amp is removed from the inputs. The voltage across the capacitor is then V_{OS}. When ϕ_2 is high and ϕ_1 is low, Fig. 25.33c, the op-amp functions normally, assuming the storage capacitance C is much larger than the input capacitance of the op-amp.

IPR2023-00698
Theta EX2019

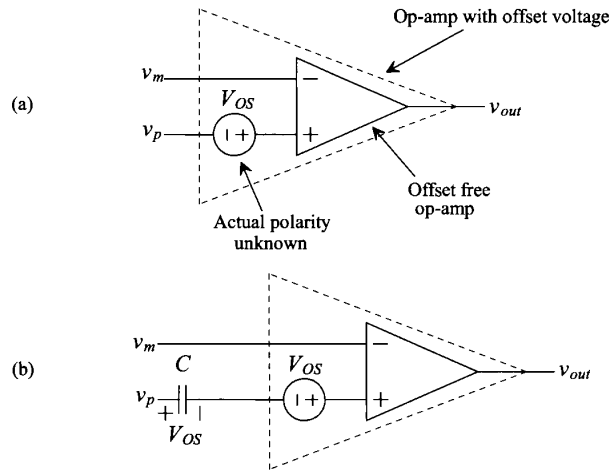


Figure 25.32 (a) Offset voltage of an op-amp modeled by a DC voltage source in series with the noninverting input of the op-amp and (b) using a capacitor to cancel the offset voltage.

Dynamic Comparator

Before we present a dynamic comparator, let's consider the *RC* switch circuit of Fig. 25.34a. When node A is connected to +1 V, node B is connected to ground. Consider what happens when the switches change positions, that is, when node A is connected to ground and the switch at node B is connected to an open. At the moment just after switching takes place, the potential at node B becomes –1 V. In other words, the voltage across the capacitor does not change instantaneously. If node A is connected back to +1 V and node B is connected back to ground a short time compared to the product of R and C, then the voltage across the capacitor remains +1 V.

A more useful circuit for CMOS is shown in Fig. 25.34b. If the switch across C_B is connected to ground when node A is connected to $V1$, then V_B, when the switches change positions, is given by

$$V_B = (V2 - V1) \cdot \frac{C_A}{C_A + C_B} \tag{25.35}$$

Figure 25.35 shows a dynamic comparator based on the inverter. When ϕ_1 is high, the voltage on the v_m input is connected to node A, while the voltage on node B is set via S3 so that the input and output voltages of the inverter are equal. (The inverter is operating as a linear amplifier where both M1 and M2 are in the saturation regions.) When ϕ_2 goes high (ϕ_1 is low since the clocks are nonoverlapping), the v_p input is connected to node A. If C_A is much larger than the input capacitance of the inverter (C_B), then the voltage change on the input of the inverter (V_B) is

IPR2023-00698
Theta EX2019

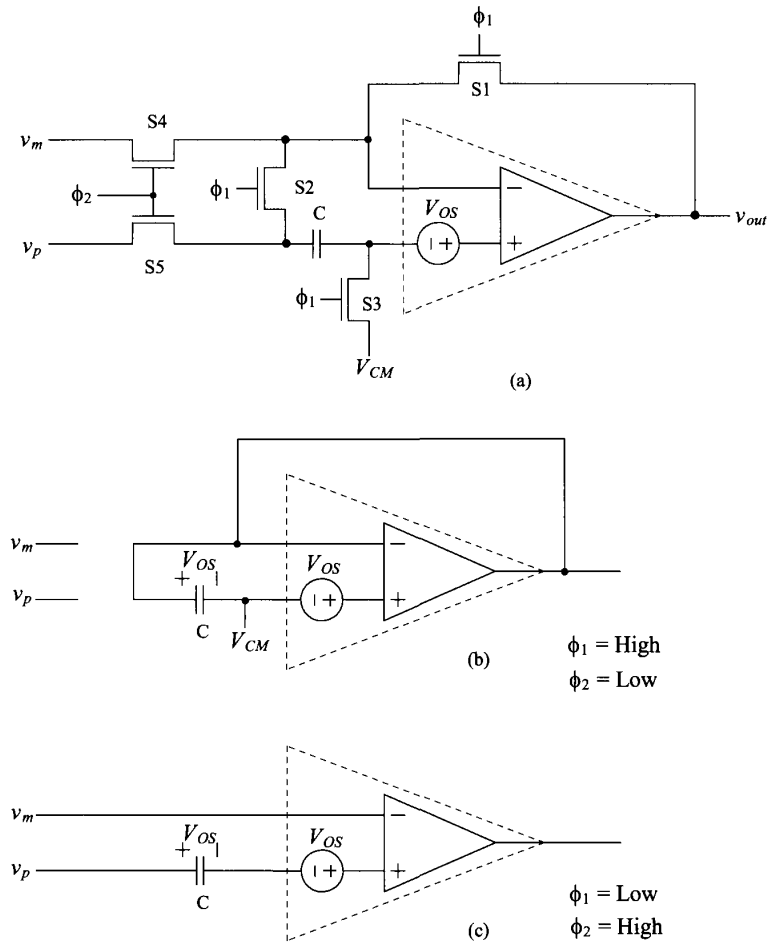


Figure 25.33 Dynamic reduction of the offset voltage.

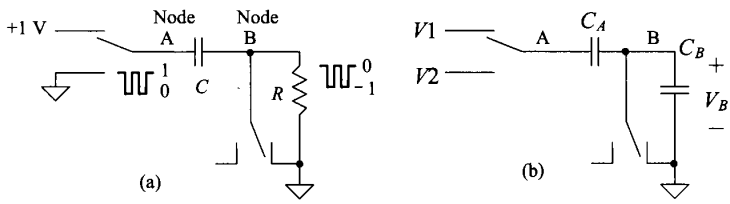


Figure 25.34 Circuits used to illustrate switching in dynamic circuits.

IPR2023-00698
Theta EX2019

$$v_{in} = v_p - v_m \tag{25.36}$$

Provided the gain of the inverter is large, this change causes the inverter output to rail, that is, go to either *VDD* or ground. The output is then latched and available during ϕ_1. The gain of the comparator can be increased by using additional inverter stages.

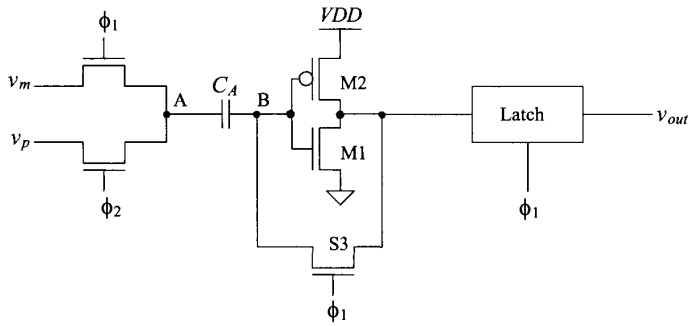


Figure 25.35 A dynamic comparator.

Another high-performance dynamic comparator configuration is based on the sense amplifier in Figs. 16.26 or 16.32 (or Fig. 16.35 so the final outputs only change on the clock's rising edge) that use positive feedback. The offset voltage of the overall comparator is reduced, using either input offset storage (IOS) or output offset storage (OOS) around the comparator preamp. Figure 25.36 shows the two types of offset cancellation techniques. In the IOS configuration in (a), the preamp must be stable in the unity feedback configuration. In the OOS configuration in (b) the MOSFETs in the preamp must remain in saturation when the offset voltage is stored on the capacitors. If a differential amplifier is used as the preamp, this condition is usually easily met.

The size of the storage capacitors is based on three important considerations: (1) preamp or latch input capacitance, (2) charge injection, and (3) kT/C noise. For the IOS scheme, the input storage capacitance must be much larger than the input capacitance of the preamp, so that the storage capacitors don't attenuate the input signals. For example, if the storage capacitors have the same capacitance value as the input capacitance of the preamp, then one-half of the input signals reaches the preamp. For the OOS scheme, the storage capacitors should be much larger than the input capacitance of the dynamic latch.

Dynamic Current Mirrors

Using dynamic techniques can reduce the effects of threshold voltage mismatches in current mirrors. Consider the circuit of Fig. 25.37. When ϕ_1 is high and ϕ_2 is low (again, these clock signals are nonoverlapping), switches S1 and S3 are on, while switch S2 is off. A current I_{ref} flows through M1, setting its gate-source voltage. This information [the gate-source voltage (actually the charge) of M1] is stored on *C*. When S2 closes with S1 and S3 off, a current I_{out} equal to I_{ref}, neglecting channel length modulation, flows. This circuit behaves like a current source when ϕ_2 is high and as an open when ϕ_2 is low. The circuit shown in Fig. 25.38 shows a dynamic current mirror that operates continuously.

IPR2023-00698
Theta EX2019

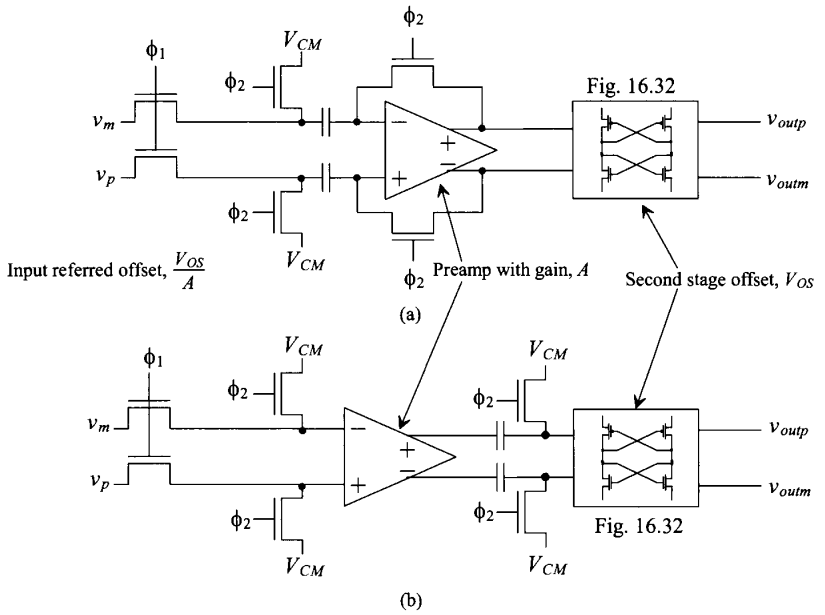


Figure 25.36 (a) Input offset storage (IOS) and (b) output offset storage (OOS).

When ϕ_1 is high, M2 sinks current, and when ϕ_2 is high, M1 sinks current. These circuits are useful in eliminating the mismatch effects and, thus, differences in the output currents, resulting from threshold voltage and transconductance parameter differences between devices. Since a single-reference current can be used to program the current in a string of current mirrors, only the finite output resistance of the mirrors causes current differences.

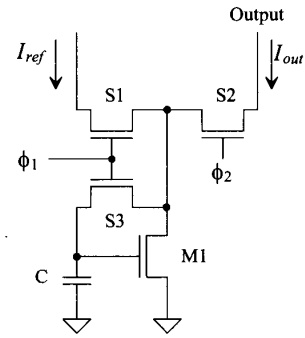


Figure 25.37 Dynamic biasing of a current mirror.

IPR2023-00698
Theta EX2019

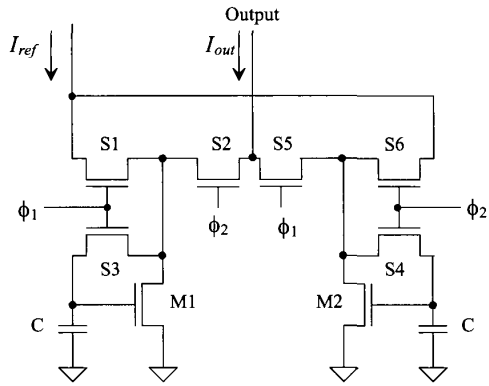


Figure 25.38 Dynamic current mirror that operates during both clock phases.

Dynamic Amplifiers

Figure 25.39 shows a dynamic amplifier. The circuit amplifies when ϕ is low and dynamically biases M1 and M2, and therefore does not amplify, when ϕ is high. If C1 and C2 are large compared to the input capacitance of M1 and M2, then the input AC signal, v_{in}, is applied to both gates. This biasing scheme makes the amplifier less sensitive to threshold and power supply variations. Other dynamic amplifier configurations exist, which have differential inputs and operate over both clock cycles.

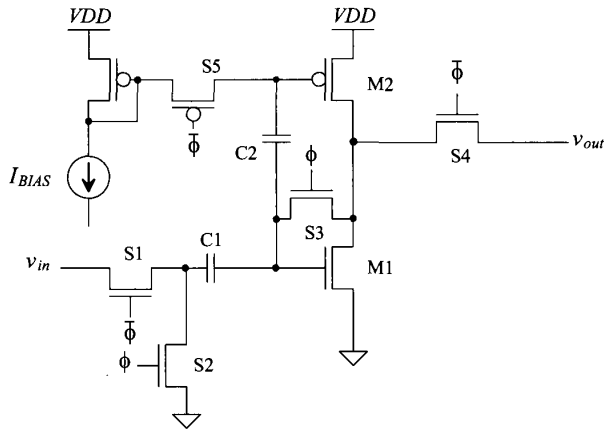


Figure 25.39 Dynamic amplifier used to reduce biasing sensitivity to power supply and threshold voltage.

IPR2023-00698
Theta EX2019

ADDITIONAL READING

[1] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, 2nd ed., Oxford University Press, 2002. ISBN 0-19-511644-5.

[2] D. Johns and K. Martin, *Analog Integrated Circuit Design*, John Wiley and Sons, New York, 1997.

[3] B. Razavi and B. A. Wooley, "Design Techniques for High-Speed, High-Resolution Comparators," *IEEE Journal of Solid-State Circuits,* vol. 25, no. 12, pp. 1916–1926, December 1992.

[4] C. Eichenberger and W. Guggenbuhl, "On Charge Injection in Analog MOS Switches and Dummy Switch Compensation Techniques," *IEEE Transactions on Circuits and Systems*, vol. 37, no. 2, pp. 256–264, February 1990.

[5] R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, McGraw-Hill Publishing Co., 1990.

[6] E. J. Kennedy, *Operational Amplifier Circuits: Theory and Applications*, Holt, Rinehart and Winston, New York, 1988.

[7] J. Shieh, M. Patil, and B. Sheu, "Measurement and Analysis of Charge Injection in MOS Analog Switches," *IEEE Journal of Solid State Circuits*, vol. 22, no. 2, pp. 277–281, April 1987.

[8] G. Wegmann, E. Vittoz, and F. Rahali, "Charge Injection in Analog MOS Switches," *IEEE Journal of Solid State Circuits*, vol. 22, no. 6, pp. 1091–1097, December 1987.

[9] R. Castello and P. R. Gray, "A High-Performance Micropower Switched -Capacitor Filter," *IEEE Journal of Solid-State Circuits,* vol. SC-20, no. 6, pp. 1122–1132, December 1987.

[10] A. G. Dingwall and V. Zazzu, "An 8-MHz Subranging 8-bit A/D Converter," *IEEE Journal of Solid-State Circuits,* vol. SC-20, no. 6, pp. 1138–1143, December 1985.

[11] A. B. Grebene, *Bipolar and MOS Integrated Circuit Design*, John Wiley and Sons, New York, 1984.

[12] P. W. Li, M. J. Chin, P. R. Gray, and R. Castello, "A Ratio-Independent Algorithmic Analog-to-Digital Conversion Technique," *IEEE Journal of Solid-State Circuits,* vol. SC-19, no. 6, pp. 828–836, December 1984.

[13] S. Masuda, Y. Kitamura, S. Ohya, and M. Kikuchi, "CMOS Sampled Differential Push-Pull Cascode Operational Amplifier," *IEEE International Symposium on Circuits and Systems,* vol. 3, pp. 1211–1214, 1983.

[14] R. Gregorian, K. W. Martin, and G. Temes, "Switched-Capacitor Circuit Design," *Proceedings of the IEEE,* vol. 71, no. 8, pp. 941–966, August 1983.

[15] D. J. Allstot and W. C. Black, "Technology Design Considerations for Monolithic MOS Switched-Capacitor Filtering Systems," *Proceedings of the IEEE*, vol. 71, no. 8, pp. 967–986, August 1983.

IPR2023-00698
Theta EX2019

[16] K. Martin, "Improved Circuits for the Realization of Switched-Capacitor Filters," *IEEE Transactions on Circuits and Systems,* vol. CAS-25, no. 4, pp. 237–244, April 1980.

[17] R. W. Broderson, P. R. Gray, and D. A. Hodges, "MOS Switched-Capacitor Filters," *Proceedings of the IEEE,* vol. 67, no. 1, January 1979.

[18] D. J. Allstot, R. W. Broderson, and P. R. Gray, "MOS Switched-Capacitor Ladder Filters," *IEEE Journal of Solid-State Circuits,* vol. SC-13, no. 6, pp. 806–814, December 1978.

[19] J. McCreary and P. R. Gray, "All MOS Charge Redistribution Analog-to-Digital Conversion Techniques - Part 1," *IEEE Journal of Solid State Circuits*, vol. 10, pp. 371–379, December 1975.

PROBLEMS

In the following problems, where appropriate, use the short-channel CMOS process with a scale factor of 50 nm and a *VDD* of 1 V.

25.1 Using SPICE simulations, show the effects of clock feedthrough on the voltage across the load capacitor for the switch circuits shown in Fig. 25.40. How does this voltage change if the capacitor value is increased to 100fF?

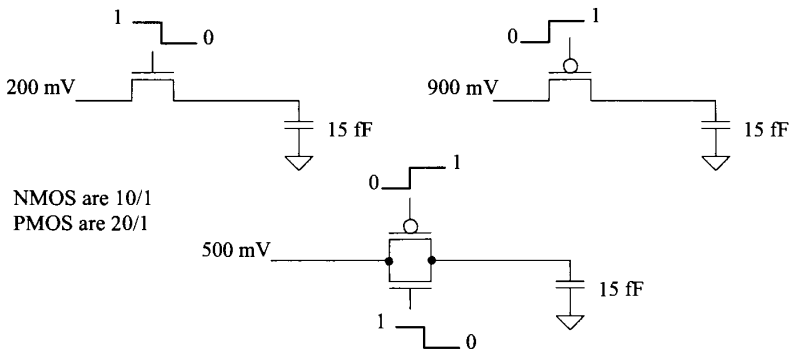


Figure 25.40 Circuits used in Problem 25.1 to show clock feedthrough.

25.2 Repeat problem 25.1 if dummy switches are used. Show schematics of how the dummy switches are added to the schematics.

25.3 Using a voltage-controlled voltage source for the op-amp (see Fig. 20.19 for example) with an open-loop gain of 10^6, use SPICE to show how the track-and-hold seen Fig. 25.8 operates with a sinewave input. What happens if the input sinewave's amplitude is above $VDD-V_{THN}$? Use a 100 MHz clock (strobe) pulse with a 50% duty cycle and an input sinewave frequency of 5 MHz. Note that the input sinewave should be centered around V_{CM} (= 500 mV).

IPR2023-00698
Theta EX2019

25.4 Using the topology seen in Fig. 25.13a and the SPICE op-amp model in Fig. 25.18, show how both V_{in} and V_{CM} can be varied while the op-amp's inputs are equal, that is, $v_p \approx v_m$. Is the output common-mode voltage always at V_{CM} using the simple SPICE model for the fully-differential output op-amp in Fig. 25.18? Why or why not? Give an example supporting your answer.

25.5 Suppose, in Fig. 25.19, that instead of the two input sine waves being connected to ground they are tied (together) to a common mode signal (say a noise voltage). Show that a common mode signal (like a sine wave) won't change the circuits' output signals. The amplitude of the common-mode signal shouldn't be so large that the NMOS switches shut off.

25.6 Show that the switched-capacitor circuits shown in Fig. 25.41 behave like resistors, for $f << f_{clk}$, with the resistor values shown.

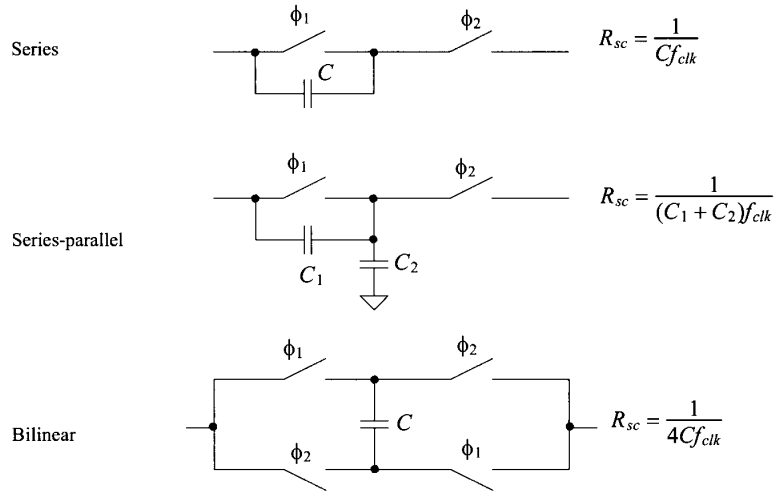


Figure 25.41 Alternative forms of switched-capacitor resistors.

25.7 Simulate the operation of the switched-capacitor resistor seen in Fig. 25.21. Plot the mean of the current flowing in the voltage sources v_1 or v_2 to show that the circuit actually behaves like a resistor. Comment on the selection of the bottom plate of the capacitor shown in Fig. 25.21a.

25.8 Comment on the selection of the bottom plate of C_F shown in Fig. 25.23.

25.9 Sketch the schematic, similar in form to Fig. 25.23, of the fully-differential switched-capacitor integrator made using a differential input/output op-amp. What is the transfer function of this topology?

IPR2023-00698
Theta EX2019

25.10 Repeat Ex. 25.2 if the low-frequency gain is 40 dB and the zero is located at 50 kHz.

25.11 Using the results given in Eq. (25.29), plot the magnitude of v_{out}/v_{in} against f/f_{clk}. Comment on the resulting plot.

25.12 An important consideration in SC circuits is the slew-rate requirements of the op-amps used. In the derivation in Fig. 25.28, we assumed that a voltage source was connected to the input of the circuit. In reality, the input of the circuit is provided by an op-amp. When ϕ_1 goes high, in this figure, the capacitor C_I is charged to the input voltage v_{in} (= v_A). If C_I is 5 pF and f_{clk} is 100 kHz, estimate the minimum slew-rate requirements for the op-amp providing v_{in}.

25.13 Suppose that the op-amp in problem 25.12 is used with a feedback factor of 0.5. Estimate the minimum unity gain frequency, f_{un}, that the op-amp must possess.

25.14 Simulate the operation of the dynamic comparator shown in Fig. 25.35.

IPR2023-00698
Theta EX2019

Chapter
26

Operational Amplifiers II

In the last chapter we saw that MOSFET switches cause charge injection and clock feedthrough in the circuits where they are used. To reduce the effects of these problems (and others), fully-differential op-amp topologies are used. As discussed in Sec. 25.2, the fully-differential output op-amp requires the design of a common-mode feedback (CMFB). The CMFB circuit keeps the op-amp's outputs balanced around a known voltage (generally the common-mode voltage of $V_{CM} = VDD/2$).

In this chapter we discuss the design of fully-differential output op-amps and CMFB circuits. Our discussion is centered around practical design where power, speed, offsets, and gain are (as usual) of importance. Throughout the chapter, the 50 nm process (with a *VDD* of 1 V) is used to illustrate the design techniques.

26.1 Biasing for Power and Speed

The biasing circuits we developed earlier were used for general analog design. In this chapter we want to design circuits that are used for very high speed with the least amount of power dissipation possible.

For high-speed design, we must use the minimum channel length (L = 1). However, using minimum channel lengths results in large mismatches between devices and low MOSFET output resistance (hence, why we used L = 2 for the designs presented earlier). The results are low gain and large input-referred offset voltages. Further, for low power design we want to use the lowest biasing currents possible. This (low biasing currents) is in direct conflict with high-speed design. As discussed in Ch. 9, the device speed figure-of-merit, FOM, was the transition frequency, f_T. Low biasing current is the same as low overdrive voltage. As seen in Eq. (9.55), using a low value of overdrive voltage results in slower circuits. Of course, if the overdrive voltage is too high, the MOSFET enters the triode region too soon. For general analog design, we set the overdrive voltage to 5% of *VDD*. For high-speed design, we might set the overdrive voltage to 10% of *VDD* or larger. *To minimize power and maximize speed, we will use minimum size devices.* For the NMOS, we'll use a 10/1. To match the drive, we'll use 20/1 for the PMOS devices. The question we need to answer having made these selections is: "Will these devices have the strength to drive a load capacitance quickly?"

IPR2023-00698
Theta EX2019

26.1.1 Device Characteristics

Figure 26.1 shows the IV characteristics of our selected devices (an NMOS of 10/1 and a PMOS of 20/1). If we set the overdrive voltages to roughly 10% of *VDD* (here 100 mV), then knowing, from Table 9.2, the threshold voltages are 280 mV (typical), we can use gate-source voltages of, nominally, 400 mV with a corresponding drain current of 20 μA. If we want to increase the speed, then we must use a higher overdrive voltage (say bias the devices up at 50 μA).

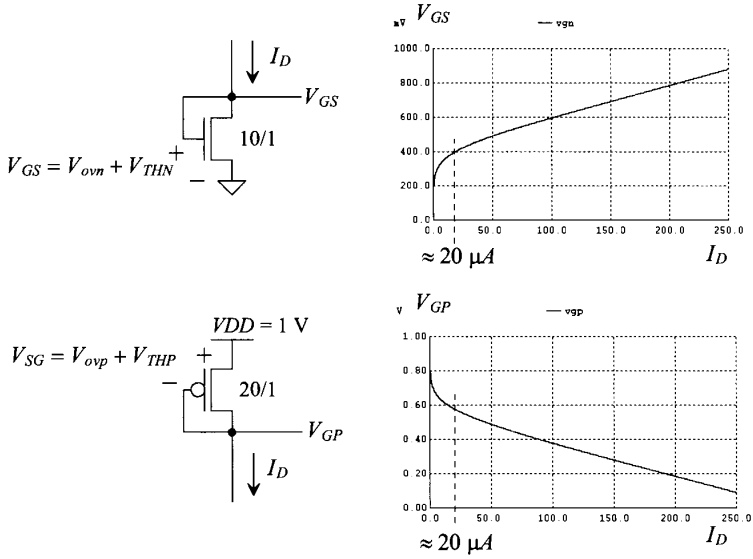


Figure 26.1 Gate-source voltages plotted against drain currents.

Next we need to determine if these devices will have the drive current needed to charge a specific size capacitor in the time required. Looking at Table 8.1 as a guide for selecting capacitor sizes, we see that using a 100 fF capacitor results in an RMS noise (because of the MOSFET switch thermal noise) of 200 μV. From Fig. 8.33 (the assumed PDF for thermal noise), the peak-to-peak value of this noise is (roughly) 1.2 mV. In this chapter *we'll use a load capacitance of 250 fF* for a peak-to-peak noise of 750 μV. Looking at the simulation results in Fig. 26.1, we see that a reasonable pulsed drain current estimate for the devices is 100 μA. The rate we can charge the load capacitor is then

$$\frac{dV_{out}}{dt} = \frac{I}{C_L} = \frac{100\,\mu A}{250\,fF} = 400\ mV/ns \qquad (26.1)$$

If our clock frequency, f_{clk}, is 100 MHz, then half a clock cycle is 5 ns and the device sizes and bias conditions will likely be adequate (remembering that *VDD* is 1 V and V_{CM} is 500 mV). However, if our clock frequency approaches 1 GHz (or the load capacitance increases), then we must increase the widths of the devices (to get more drive) and the overdrive voltages (to get more speed).

IPR2023-00698
Theta EX2019

26.1.2 Biasing Circuit

When designing the bias circuits earlier, Fig. 20.47 for example, our goals were to provide biasing for general analog design. Here, in this chapter, our goal is to design a bias circuit for our op-amp designs with minimum power dissipation. When designing biasing circuits fully-differential topologies have some advantages over the single output op-amps presented in Ch. 24. For example, our earlier bias circuit provided biasing, V_{pcas} and V_{ncas}, for the floating current sources used in the class AB output stages (see Fig. 24.29 for example). These bias voltages won't be needed here.

Towards understanding this last statement, consider the two-stage op-amp seen in Fig. 26.2. Consider what happens if the noninverting op-amp input, v_p, increases relative to the inverting input, v_m. The drain voltage of M1R drops while the drain voltage of M1L rises. The decrease in the gate voltage of M6R causes it to turn on and the output to go high. At the same time, the increase in the drain voltage of M1L causes M4R to shut off and thus so does M7R (giving class AB operation). In simple terms, we can yank the gates of M4 or M6 down independent of the diff-amp's tail current.

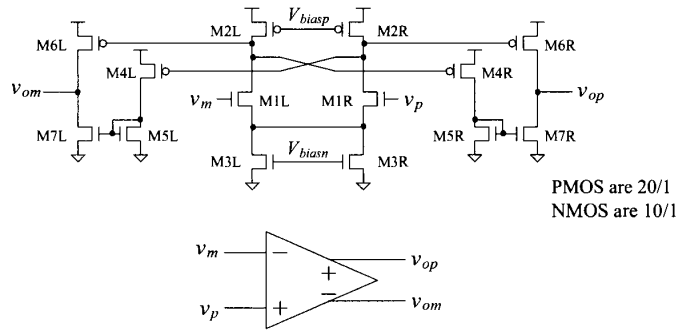


Figure 26.2 A two-stage fully-differential op-amp. Compensation and CMFB are not shown. Output stage operates class AB. See discussion in the next section concerning the output voltage of the diff-amp.

Layout of Differential Op-Amps

One of the common things we do in this chapter is draw schematics that are symmetrical. If we were to draw a line down the middle of the op-amp in Fig. 26.2, separating the left and right sides of the schematic, we could fold the left side of the schematic directly over onto the right side of the schematic and see a perfect match. For example, instead of drawing the diff-amp's tail current source (M3) as a single MOSFET with twice the width, we've drawn it as two MOSFETs in parallel, each having the same width. *Drawing schematics in this fashion is useful when doing layout.* We can fold the schematic in half and lay out like devices, e.g., M6L and M6R, directly next to each other. Further, then the outputs (and inputs) of the op-amp are laid out right next to each other.

IPR2023-00698
Theta EX2019

Self-Biased Reference

Figure 26.3 shows the bias circuit we'll use in this chapter (see also Fig. 20.22 and the associated discussion). We used the length of 2 MOSFETs for the added amplifier to minimize power dissipation and boost the amplifier's gain. The current pulled from *VDD* is approximately 50 μA. Figure 26.4 shows the simulation results of how the current varies with changes in *VDD*. Notice how V_{biasn} is close to the 400 mV, and V_{biasp} is close to *VDD* – 400 mV (of course the absolute value of V_{biasp} varies with *VDD* but the ideal value of V_{SG} for the PMOS devices will be 400 mV). Note that, as discussed in Sec. 20.1.4, it is important to ensure that the reference is stable over all possible operating conditions and loads (connected to V_{biasn} and V_{biasp}). Notice that we aren't discussing how the reference currents vary with process shifts in the MOSFETs and the resistor. This is an important practical concern.

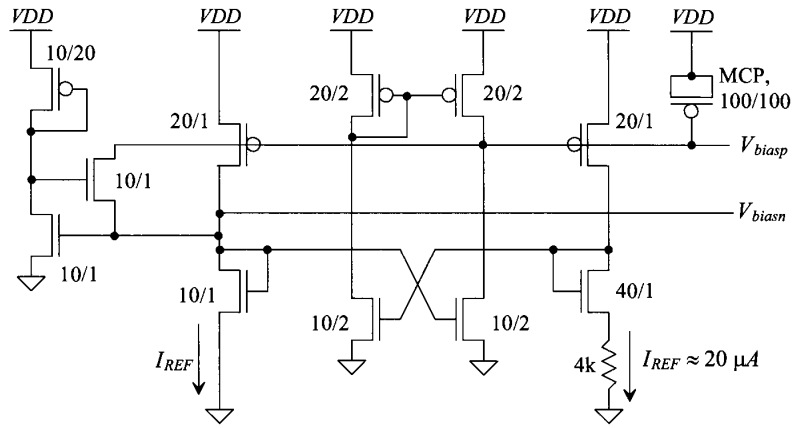


Figure 26.3 Biasing circuit used in this chapter. This bias circuit pulls approximately 50 microamps.

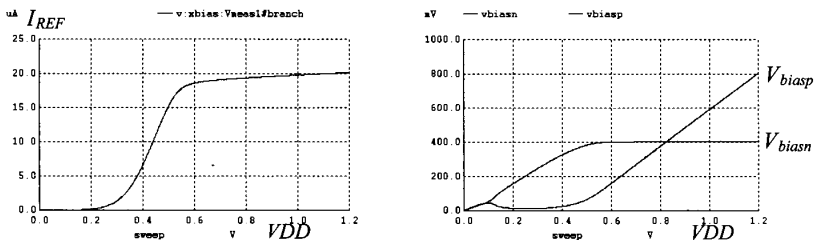


Figure 26.4 Simulating how the reference current changes with VDD.

IPR2023-00698
Theta EX2019

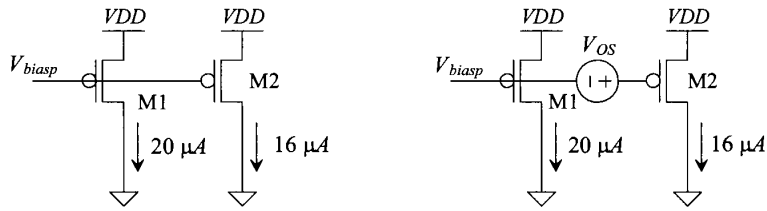
26.2 Basic Concepts

Before we discuss the design of op-amps, let's look at some basic concepts that will be useful when evaluating a specific op-amp topology.

Modeling Offset

In a real circuit, especially an analog circuit designed with minimum length devices, mismatches can be a significant factor in the selection of an op-amp topology. In a SPICE simulation, all of the MOSFETs are perfectly matched. We have to come up with a method, in SPICE, to determine an op-amp's sensitivity to offsets.

Consider the circuit in Fig. 26.5. In (a) the measured currents should be equal. Both MOSFETs V_{SG}s and V_{SD}s are equal (and they are the same size). However, for whatever reason (e.g., threshold voltage mismatch), there exists a mismatch between the two devices that causes a difference in their drain currents. We can model this mismatch in a SPICE simulation, as seen in (b), by adding a DC voltage source in series with the gate of M2. For a general design, we can insert these offset voltages at various points in the circuit and verify that the circuit still functions correctly. The next question that needs answering is: "What value of V_{OS} should be used?" While no absolute answer can be given here, **we'll use a V_{OS} of 50 mV** (5% of *VDD* or roughly 20% of the threshold voltage). The reader might feel that this value is way too high. However, if we can design circuits that function properly over the process, voltage, and temperature (PVT) variations with an offset this large, it is likely the op-amp will be difficult to "break."



(a) M1 and M2 are mismatched.

(b) M1 and M2 are perfectly matched (as in a SPICE simulation).

Figure 26.5 How we add an offset into the circuit to model mismatch.

A Diff-Amp

Figure 26.6a shows a basic diff-amp without an offset. This offset-free diff-amp has simulated output voltages of 700 mV. The diff-amp in (b) is simulated with an input-referred offset voltage of 50 mV. The polarity of the offset doesn't matter because it simply swaps the two output voltage values. In other words, we could have put the same polarity offset on the other input of the diff-amp and swapped the voltages on the output of the op-amp. We put the offset voltage on the op-amp's input because the input has the greatest effect on the diff-amp's output voltages. Note: we can think of the 50 mV as an input signal causing an output signal difference, from the ideal 700 mV, of 900 mV −550 mV or 350 mV (assuming linear operation). This indicates the gain of the diff-amp is only 7!

IPR2023-00698
Theta EX2019

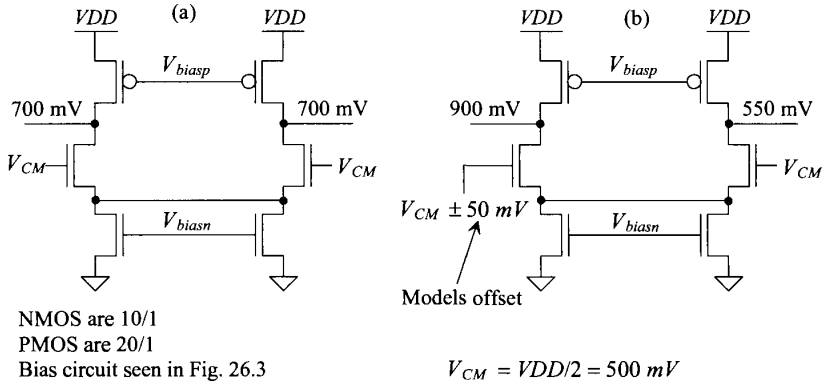


Figure 26.6 Comparing the diff-amp's output voltages with and without an offset.

The simulation results seen in Fig. 26.6 give some practical information that we need to consider. To begin, remember from Ch. 24 that we frequently use the diff-amp's output to bias the next stage (see Figs. 22.8 and 24.2). With a *VDD* of 1 V and a V_{SG} of nominally 400 mV, V_{biasp} is roughly 600 mV, as seen in Fig. 26.4. However, as seen in Fig. 26.6a, the diff-amp's output is 700 mV. The result, when the diff-amp is used with a second-stage and feedback, is an additional input-referred offset. It would be nice to know that the outputs of the diff-amp, in the ideal case, go to a known value (preferably a voltage that can be used to bias the next stage). The diff-amp in Fig. 26.6 is an example of two current sources, the PMOS biased with V_{biasp} and the NMOS biased with V_{biasn}, fighting each other for control of the output voltage. In general, we want to avoid this situation.

A Single Bias Input Diff-Amp

Figure 26.7 shows a diff-amp that generates its own bias reference for the PMOS devices. Notice that the two gate-drain-connected PMOS devices behave simply like a MOSFET with twice the width of the other two PMOS devices. Similarly, the four NMOS devices that are biased from V_{biasn} behave like a MOSFET with four times the width of the other NMOS devices. When the diff-amp's inputs are equal, the same current, *I*, flows in all of the MOSFETs. If the + diff-amp input is raised significantly above the – input, all of the bias tail current flows in the left two NMOS devices of the diff-amp (each will conduct 2*I*). The current flowing in the gate-drain-connected PMOS device, in either case, is the same, 2*I*, keeping the PMOS's gate voltage constant. The outputs of this diff-amp can be used to bias the next stage. Offsets, however, cause the diff-amp's outputs to vary from their ideal values. One drawback to using this diff-amp is that it dissipates twice the power of the diff-amp in Fig. 26.6. Another drawback is the larger input capacitance.

The Diff-Amp's Tail Current Source

Notice that we are not using a cascode tail current source to bias the diff-amps in this chapter, as we did in earlier chapters. We're avoiding the cascode structure because we'd need an additional bias voltage (resulting in more power dissipation). In a

IPR2023-00698
Theta EX2019

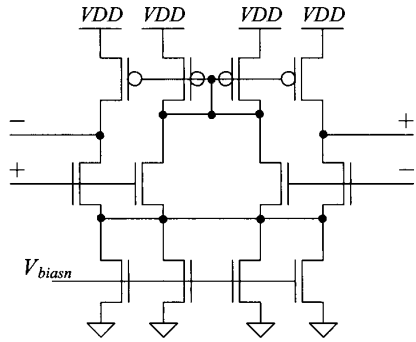


Figure 26.7 A fully-differential diff-amp that generates its own bias for the PMOS.

fully-differential op-amp topology (both the inputs and the outputs of the op-amp, with feedback, are double-ended signals swinging around V_{CM}), the input common-mode voltage of the op-amp is constant (= V_{CM}). The common-mode rejection ratio isn't as important when the input common-mode voltage of the op-amp doesn't vary.

Using a CMFB Amplifier

Another possible way to set the diff-amp's output voltages to a known value is seen in Fig. 26.8. An amplifier (called a common-mode feedback amplifier or CMFB amplifier) is used to amplify the difference between the average of the diff-amp's outputs and V_{biasp}. If the gain of the CMFB amplifier is large, then the average of the two outputs will be very close to V_{biasp}. Note that the CMFB amplifier's output signal, V_{CMFB} , is common, through M1L and M1R, to both outputs. Any variation in V_{CMFB} affects each output by the same amount. This is important because all we want the CMFB amplifier to do is make

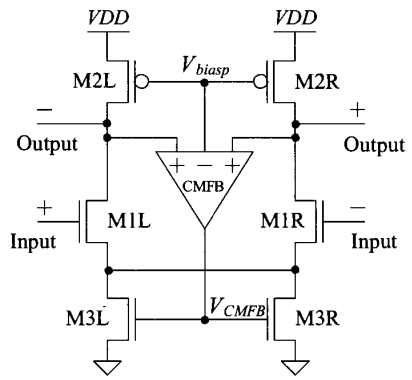


Figure 26.8 Using a common-mode feedback (CMFB) amplifier to set the output voltages.

IPR2023-00698
Theta EX2019

sure that the diff-amp's outputs vary around V_{biasp}. The CMFB amplifier shouldn't affect the differential amplification in the diff-amp. When the diff-amp's outputs are equal (the inputs to the diff-amp are equal and neglecting offsets), they should be V_{biasp}.

Figure 26.9 shows one possible implementation of a CMFB amplifier. When the + inputs (noticing the PMOS connected to these inputs have half the width of the other PMOS) are equal to V_{biasp}, the currents that flow in M1 and M2 are 20 μA. This provides the proper mirroring action to the diff-amp (M3L, R) in Fig. 26.8. If the average of the + inputs moves above the – input, the current flowing in M2 decreases and the current in the 200/1 PMOS increases. This causes V_{CMFB} to increase. The result is that M3L and M3R turn on more and pull both of the outputs in Fig. 26.8 down (until their average is equal to V_{biasp}). Several concerns exist with this topology. To begin, whenever we employ feedback, we must be concerned with stability. To stabilize the CMFB loop, we can add capacitors to the outputs of the diff-amp (inputs to the CMFB amplifier). We may be able to stabilize the CMFB loop by adding capacitors to the CMFB amplifier's outputs. However, since we will need compensation capacitors for compensating the differential action of the op-amp, we might as well use these capacitors for both loops. The other concern with the CMFB amplifier in Fig. 26.9 is the input common-mode range. When the CMFB amplifier's + inputs move significantly away from the – input, we don't get the correct balancing. This keeps large signal swings on the diff-amp's outputs in Fig. 26.8 from being balanced around V_{biasp}.

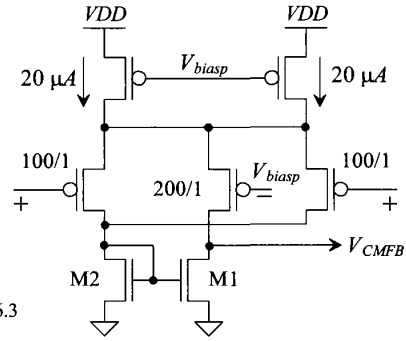
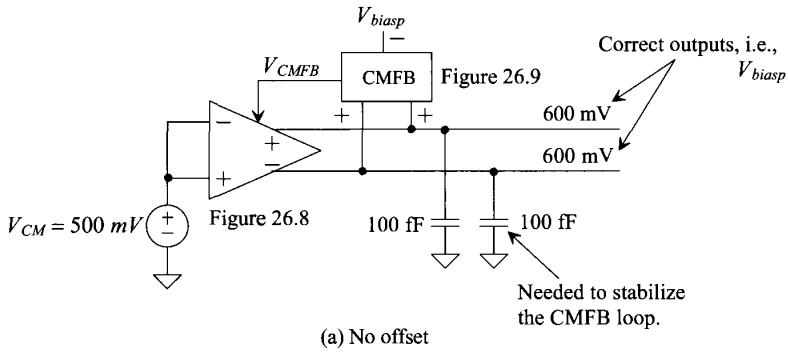


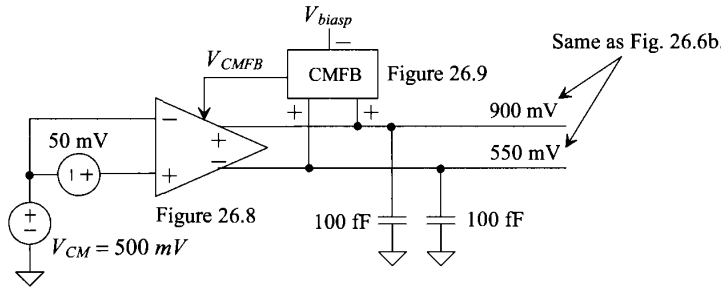
Figure 26.9 Implementation of the CMFB amplifier in Fig. 26.8.

Figure 26.10 shows simulation results using the CMFB amplifier in Fig. 26.9 with the amplifier in Fig. 26.8. The addition of the capacitors stabilizes the CMFB loop *(we need to discuss this further)*. In (a) we see the correct balancing. The CMFB loop adjusts the diff-amp's output voltages to V_{biasp}. In (b) we re-simulate with an offset and see that the CMFB circuit isn't balancing the outputs. If it were, we would see one output above V_{biasp} by some amount and the other output below V_{biasp} by the same amount. With these output voltages (the inputs to the CMFB amplifier), one of the PMOS devices on the input of the CMFB is off and we don't get proper amplifier action. *We must use a CMFB circuit that can balance the outputs over the entire range of diff-amp output voltages.*

IPR2023-00698
Theta EX2019



(a) No offset



(b) With a 50 mV offset. Note how the CMFB isn't doing anything.

Figure 26.10 Simulating the operation of the CMFB circuit in Fig. 26.9.

Compensating the CMFB Loop

Consider the schematic seen in Fig. 26.11. The diff-amp behaves like an operational transconductance that can be compensated, as discussed in Sec. 24.3 (see Eq. [24.44]). The CMFB loop can be compensated in a similar fashion. The AC common-mode signal is represented in this schematic as v_c. If the gain of the CMFB amplifier is A_c, then following the procedure leading to Eq. (24.44), we can write the unity-gain frequency of the CMFB loop as

$$f_{un,cm} = \frac{A_{cm} \cdot g_{mn}}{2\pi C_L} \tag{26.2}$$

If we want to compensate the CMFB loop with the same load capacitance used to compensate the differential forward signal path, then we must ensure that the gain of the CMFB amplifier is less than or equal to unity,

$$A_{cm} \leq 1 \tag{26.3}$$

IPR2023-00698
Theta EX2019

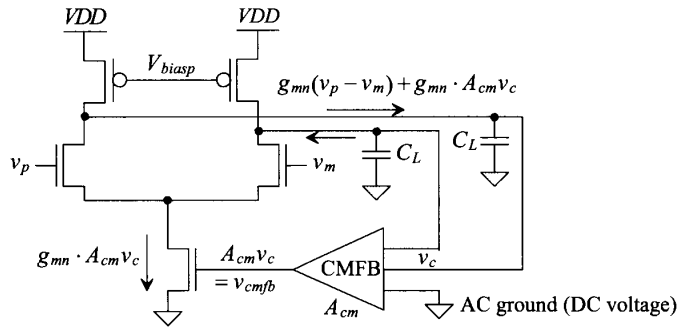


Figure 26.11 Schematic view of differential and CM feedback.

Reviewing the CMFB amplifier in Fig. 26.9 that has a current mirror active load, we see that A_{cm} is greater than 1. Removing the capacitors in Fig. 26.10 and resimulating will show that the CMFB loop is unstable. The added capacitors are relatively large and will overcompensate the differential signal path of the diff-amp. What we need to do is reduce the gain of the CMFB amplifier or reduce the CMFB loop's forward gain.

Towards reducing the gain, examine the CMFB amplifier seen in Fig. 26.12. This is the same amplifier topology seen in Fig. 26.9 except that here we've used a diode-connected load instead of a current mirror load (to reduce the gain). The schematic is drawn symmetrical around its center for ease of layout, as discussed earlier. Note that we used the common-mode voltage in this schematic, V_{CM}, as the voltage that the outputs of the amplifier will swing around (the more general case) rather than V_{biasp} as used in Fig. 26.10. Using this CMFB amplifier in the circuits of Fig. 26.10 (where the diff-amp of Fig. 26.8 has a low gain) won't precisely balance the outputs. The loop gain around the CMFB loop isn't large enough for proper operation. As we saw in Ch. 24, low first-stage gain can be remedied by using a telescopic input (cascode-load diff-amp) or a folded-cascode OTA. The other problem with this CMFB amplifier, again, is the limited allowable swing on the + inputs. As we've already seen, the input common-mode range limits the range of output voltages this CMFB can balance properly.

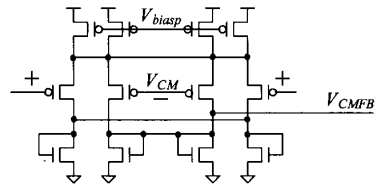


Figure 26.12 A CMFB amplifier with a gain of nominally unity.

IPR2023-00698
Theta EX2019

Towards reducing the CMFB loop's forward gain, consider breaking the diff-amp's tail current up into parts, as seen in Fig. 26.13. The CMFB signal is applied to only one gate of the tail current. When compared to the topology in Fig. 26.8, the forward gain of the CMFB loop is halved. Further reduction can be implemented by adjusting the sizes of the transistors to further reduce the strength of the V_{CMFB} signal. *This is a common practical way of making CMFB loops stable.*

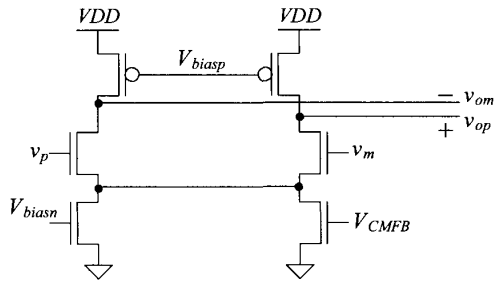
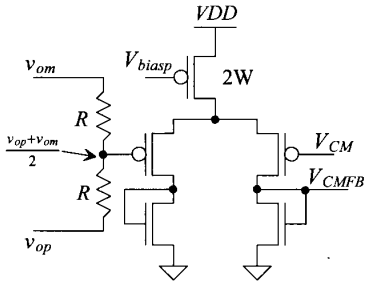


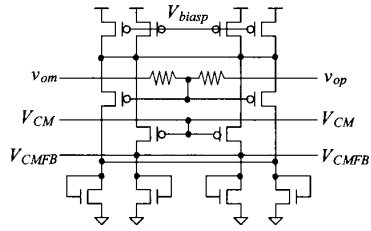
Figure 26.13 Reducing the forward gain of the CMFB loop.

Extending the CMFB Amplifier Input Range

The problem with the previous CMFB amplifier topologies based on a diff-amp is the diff-amp's limited input range. It's desirable to have a CMFB amplifier that functions over the entire range of possible amplifier output voltages. Towards this goal, consider the conceptual schematic in Fig. 26.14a. The resistors average the two outputs. This average is compared with the common-mode voltage (or a bias voltage as used in Fig. 26.8). Figure 26.14b shows the practical implementation of the amplifier for symmetry (at the cost of extra power dissipation). The practical problem with this topology is the



(a) Using resistors to average differential output signals.



(b) Symmetrical implementation of the CMFB circuit in (a).

Figure 26.14 Increasing CMFB amplifier input range.

IPR2023-00698
Theta EX2019

loading by the resistors. If we were to connect this CMFB amplifier in the circuit configuration of Fig. 26.8, the resistors, unless they are huge (>100k) would load the differential amplifier and lower its gain. This topology is used, most often, on the output of an op-amp that has output buffers (and can thus drive resistive loads).

When using resistors for averaging in high-speed applications, we may have some parasitic effects that should be considered. The output signals have to charge, through the averaging resistors, the input capacitance of the MOSFET, as seen in Fig. 26.15. To ensure that the balancing action works at high speeds, capacitors (shown dashed in the figure) can be added, shunting the resistors. These capacitors can be very important if the size of the resistors is increased to reduce their loading on the output of the amplifier.

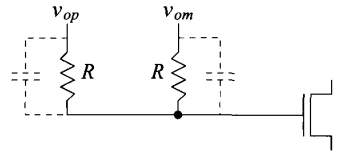


Figure 26.15 Adding parasitic capacitances across the resistors to compensate for the input capacitance of the MOSFET.

Dynamic CMFB

Figure 26.16 shows a switched-capacitor (SC) implementation of a CMFB circuit. The clocks, as in all SC circuits, are nonoverlapping (never high at the same time) clock signals, as seen in Fig. 25.28. The SC resistors are formed with the C_1 capacitors. The C_2 capacitors are used for the high-speed averaging just discussed (the dashed capacitors in Fig. 26.15). The SC resistors, as we'll see in a moment, perform both the averaging and the differencing needed in a CMFB amplifier. If the ϕ_2 controlled switches connected to amplifier's outputs, v_{op} and v_{om}, are transmission gates, the circuit can provide balancing from *VDD* to ground.

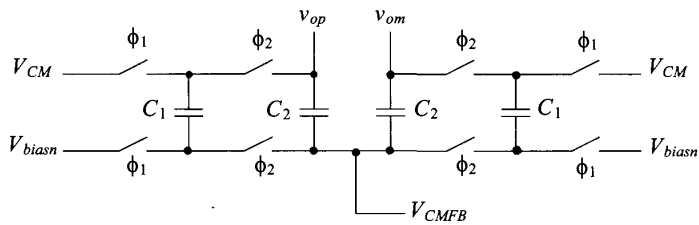


Figure 26.16 A switched-capacitor CMFB circuit.

IPR2023-00698
Theta EX2019

To describe the operation of this circuit, consider the case when ϕ_1 is high. During this time, the total charge stored on both C_1 capacitors is

$$q_1 = 2 \cdot (V_{biasn} - V_{CM}) \cdot C_1 \tag{26.4}$$

When the ϕ_1 switches shut off, the ϕ_2 switches turn on. The total charge on both capacitors is then

$$q_2 = (V_{CMFB} - v_{op}) \cdot C_1 + (V_{CMFB} - v_{om}) \cdot C_1. \tag{26.5}$$

The change in V_{CMFB} is proportional to difference in q_1 and q_2 or

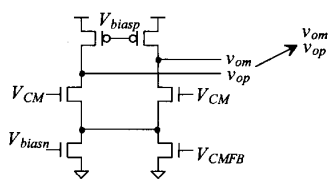
$$\Delta V_{CMFB} \cdot 2(C_1 + C_2) \propto (q_1 - q_2) \tag{26.6}$$

This equation is important because it shows the CMFB voltage will continue to change until the two charges, q_1 and q_2, are equal. Note that if v_{op} and v_{om} are balanced around V_{CM}, their net contributions, when ϕ_2 goes high, to V_{CMFB} are zero. Looking at the difference in the charges, we get

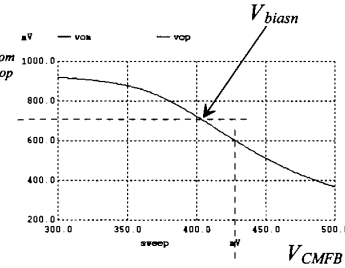
$$q_1 - q_2 = 2C_1\left(V_{biasn} - V_{CMFB} + \frac{v_{op} + v_{om}}{2} - V_{CM}\right) \tag{26.7}$$

This equation is quite interesting. Ideally, V_{CMFB} is equal to V_{biasn}, and the average of the outputs is equal to the common-mode voltage. If the actual value of V_{CMFB}, for balanced outputs, is 10 mV offset from V_{biasn}, then the average of the outputs will be 10 mV offset from V_{CM}. The offsets can occur because of improper device sizing (under ideal conditions the currents don't sum correctly) or mismatches.

As an example of where this offset can come from, consider the amplifier seen in Fig. 26.17 (also in Figs. 26.6 and 26.13). As seen in Fig. 26.6a, when V_{CMFB} is V_{biasn} or roughly 400 mV, the outputs of the diff-amp are 700 mV. As seen in Fig. 26.10 and the associated discussions, it can be useful, for next stage biasing, if the outputs are set to V_{biasp} (600 mV). From Fig. 26.17, the value of V_{CMFB} at this output voltage is roughly 425 mV or a 25 mV offset.



NMOS 10/1
PMOS 20/1
Bias circuit from Fig. 26.3



Value of CM feedback voltage when the outputs are 600 mV.

Figure 26.17 Plotting the output voltages as a function of the CM feedback voltage.

IPR2023-00698
Theta EX2019

26.3 Basic Op-Amp Design

Reviewing the data in Table 9.2, we see that the open circuit gains are 25 (NMOS) and 50 (PMOS). In this chapter we both increased the biasing current and reduced the channel length from the values used in Table 9.2. Each of these changes has the effect of reducing the MOSFET's open circuit gain. If, for example, the open circuit gains are both now 10, then a common-source amplifier with current source load will have a gain of 5. A cascode amplifier will have a gain of 25, and a two-stage op-amp using a cascoded first-stage a gain of only 125. In other words, we adjusted our biasing for high-speed operation but we are going to face some issues with getting large open-loop gain.

This is a good time to remember the useful simulation netlists that we've developed. Figure 26.18 shows the IV curves, output resistance, and transconductance for a 10/1 NMOS and a 20/1 PMOS based on Figs. 9.31 to 9.33. Notice in (a) that the NMOS's drain current at a V_{GS} of 400 mV and a V_{DS} of 100 mV is approximately 13 μA. The PMOS's drain current under the same conditions, (b), is closer to 9 μA. In (c) and

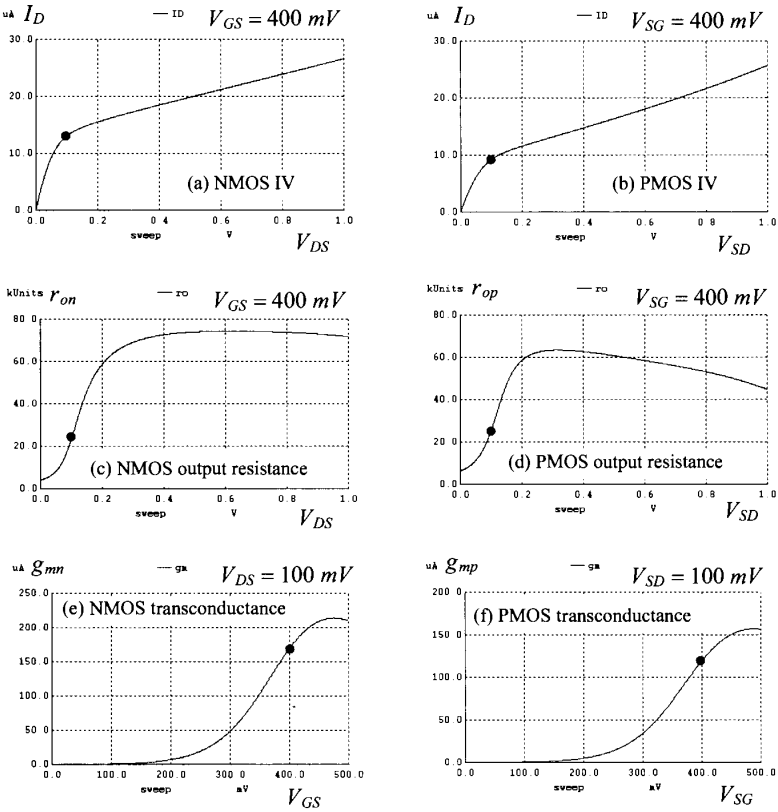


Figure 26.18 Characteristics of NMOS (10/1) and PMOS (20/1) devices.

IPR2023-00698
Theta EX2019

(d), at a V_{DS} of 100 mV, the output resistance is only 25k. Using the transconductances in (e) and (f), the open circuit gains are, roughly, 4.375 (NMOS) and 3.125 (PMOS). In the actual circuits, the biasing points will vary (but in any case the gain of single stages will be low).

The Differential Amplifier

Figure 26.19 shows a cascode load diff-amp based on the topology seen in Fig. 26.7. Seen in the figure are typical values for the voltages in the circuit assuming gate-source voltages of 400 mV and, for the bottom two rows of NMOS devices, drain-source voltages of 100 mV. Notice how we used V_{biasn} to bias the second row of PMOS devices, which puts 200 mV across the drain-source voltages of the PMOS devices. This results in larger gain and reduced output swing. Diff-amp output swing is not an issue if this amplifier is used as the first-stage in a two-stage op-amp.

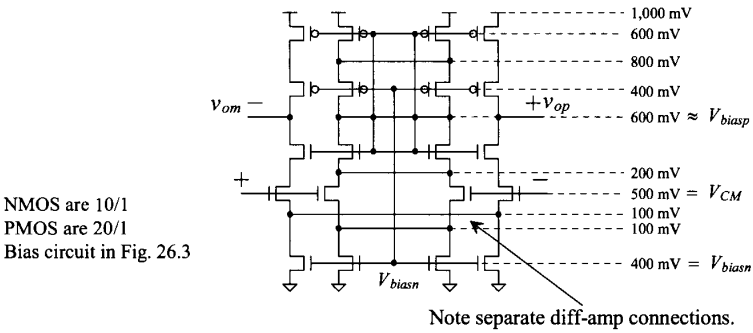


Figure 26.19 Fully-differential cascode diff-amp.

Figure 26.20 shows how the output voltages of the diff-amp in Fig. 26.19 vary with changes in V_{CM}. A 200 mV change in the common-mode voltage results in a, roughly, 50 mV change in the diff-amp's common-mode output voltages. Looking at Fig. 26.1, we see that the change in the drain current will be, again roughly, $\pm$ 5 μA around the quiescent value set with V_{CM} equal to 500 mV.

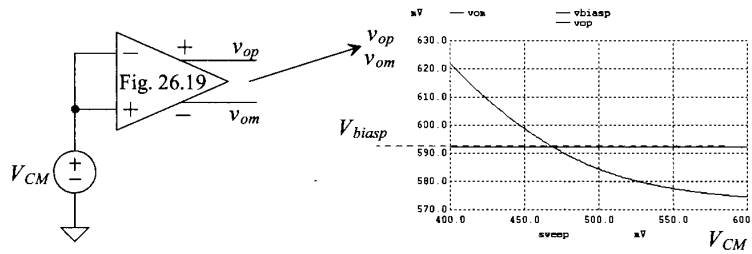


Figure 26.20 Varying the common-mode voltage and looking at the output.

IPR2023-00698
Theta EX2019

Figure 26.21 shows the DC characteristics of the diff-amp. The gain is approximately 40. An important concern is how the mismatches in the MOSFETs used in the diff-amp affect the operation and biasing of the amplifier. Before discussing this issue, let's add the second stage and CMFB circuitry to form an op-amp. Note that, with a diff-amp gain of 40, and a second stage gain of 10, our op-amp's open loop gain will only have a value in the hundreds.

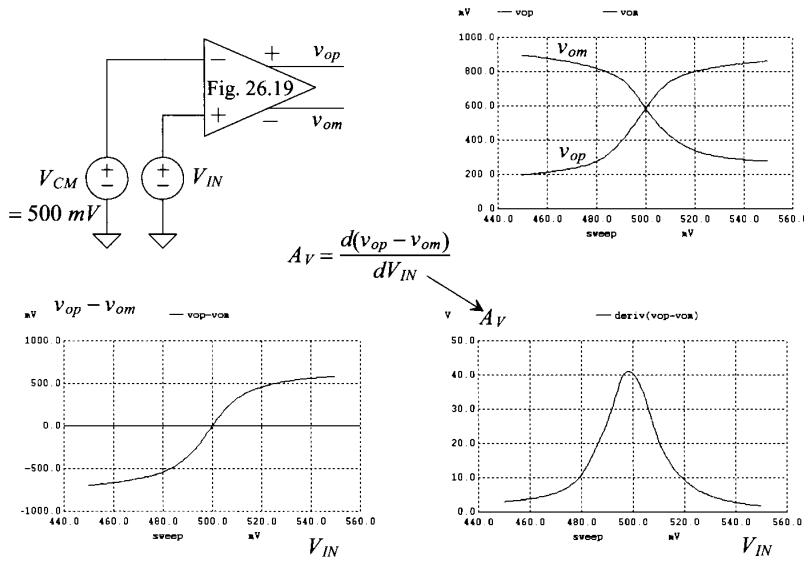


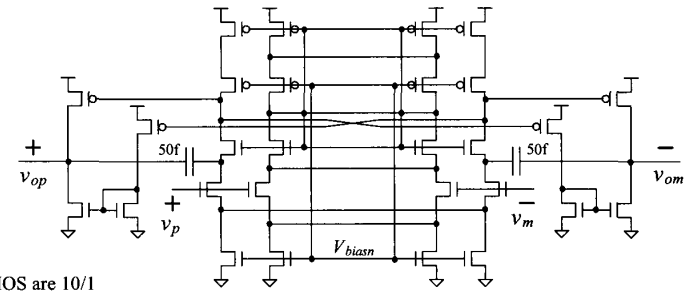
Figure 26.21 DC behavior and gain of the diff-amp in Fig. 26.19.

Adding a Second Stage (Making an Op-Amp)

Figure 26.22 shows a two-stage op-amp without CMFB circuit. The second stage of the op-amp operates class AB, as seen in Fig. 26.2, and the associated discussion. We've spent a considerable amount of time discussing how the output voltages of the diff-amp are approximately V_{biasp}. It should be clear after studying the op-amp in Fig. 26.22 why this is important. This voltage sets the quiescent current flowing in the output stages. There are eight vertical branches in this op-amp, so we can estimate the current pulled from *VDD* under quiescent conditions as 160 μA.

We used 50 fF capacitors for compensation in this op-amp. We can estimate the slew-rate limitations caused by the diff-amp driving the compensation capacitor, Eq. (22.34), as $20\ \mu A/50\ fF = 400\ mV/ns$. Using a class AB output stage, we don't have slew-rate limitations associated with driving a load capacitance from a constant current source. As discussed at the beginning of the chapter, the output MOSFET's drain currents can be pulsed to a value greater than 100 μA. When the op-amp is driving a 250 fF capacitive load, the speed limitations associated with charging the load capacitance are similar to the limitations we get when the diff-amp drives the compensation capacitor.

IPR2023-00698
Theta EX2019



NMOS are 10/1
PMOS are 20/1
Bias circuit in Fig. 26.3

Figure 26.22 Basic two-stage op-amp without CMFB.

Figure 26.23 shows the DC characteristics of the op-amp in Fig. 26.22 where, once again, we've held the inverting op-amp input at the common-mode voltage and swept the voltage on the noninverting input. Notice how the outputs swing all the way from ground to *VDD* (= 1V). Further notice how the differential output voltage swings from −1 to +1 V (a doubling in the output swing as discussed in Sec. 25.2.) The op-amp's DC gain is approximately 500 without a DC load. Notice how the two op-amp outputs cross at 400 mV (not at V_{CM} where they should). We'll discuss the CMFB in a moment.

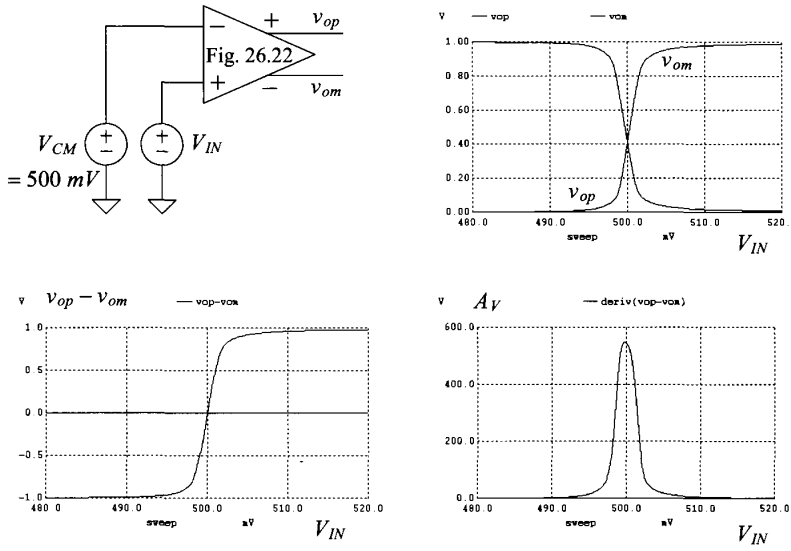


Figure 26.23 DC behavior and gain of the op-amp in Fig. 26.22.

IPR2023-00698
Theta EX2019

Step Response

To determine both the stability of the op-amp in Fig. 26.22 and the settling time under certain loading conditions, consider the configuration seen in Fig. 26.24. We used relatively small resistors, 20k, in this circuit to reduce the RC time constant associated with the outputs of the op-amp charging the input capacitance of the op-amp. If, for example, the input capacitance of the op-amp is 25 fF (from parasitics and the MOSFETs used on the op-amp's input), then the RC time associated with charging this capacitance through a 20k resistor is 0.5 ns. As seen in Fig. 26.24, the settling time is approximately 2.5 ns so this RC time can have a significant effect on the settling time and stability of the circuit (important).

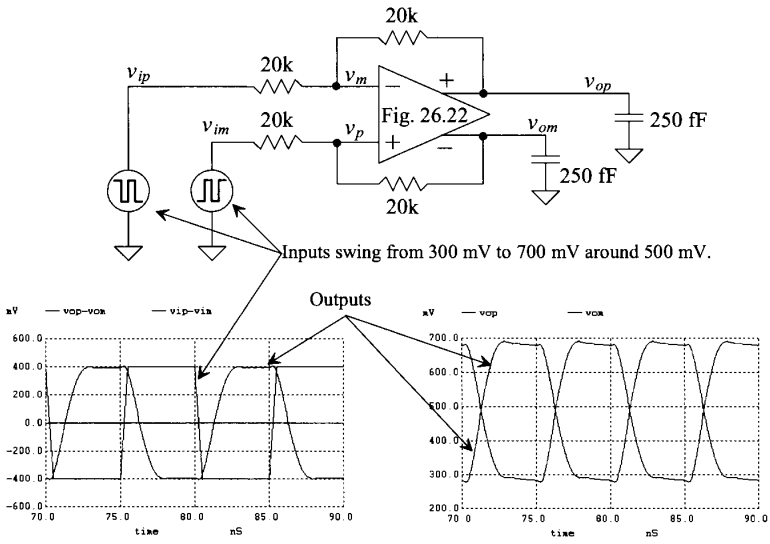


Figure 26.24 Step response of the op-amp in Fig. 26.22 driving 250 fF load capacitors and 20k feedback resistors.

Notice how, in Fig. 26.24, we used input signals that don't swing rail-to-rail. Since we don't have a CMFB circuit in the op-amp, the exact common-mode output voltage is an unknown. It may be 400 mV or it may be 600 mV. We get some help in setting the circuit's output common-mode level by using input signals with common-mode voltages of 500 mV and DC feedback. Looking at the simulation results in Fig. 26.24, we might get a false sense of not needing a CMFB circuit. To illustrate this is indeed a false sense, consider the sample-and-hold seen in Fig. 26.25 (see also Fig. 25.19). When the ϕ_1 switches are closed, the op-amp's inputs and outputs should be held to $V_{CM} \pm V_{OS}$ (the op-amp is placed in the follower configuration). As seen in the figure, the outputs during this time are driven close to 400 mV. On closer inspection, we see the output common-mode level is wandering downwards until eventually the op-amp shuts off.

IPR2023-00698
Theta EX2019

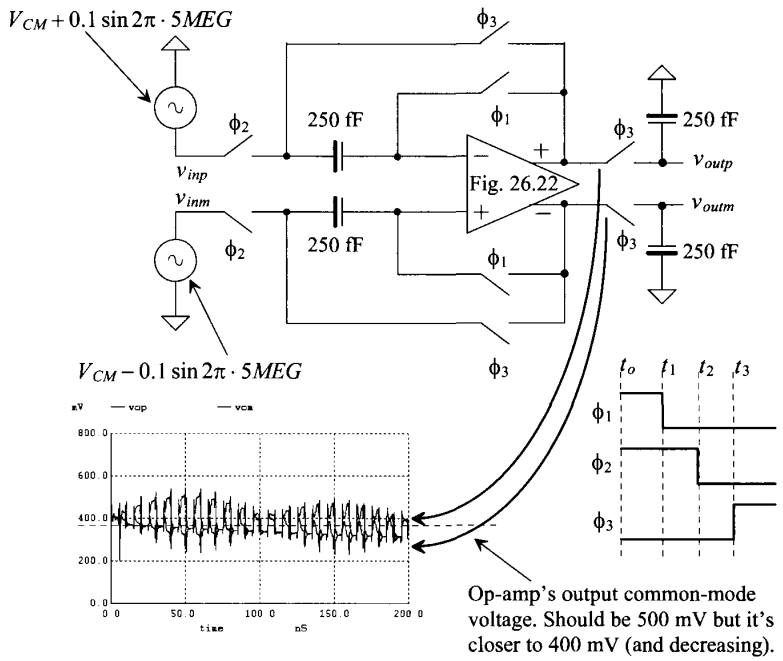


Figure 26.25 A sample-and-hold circuit. Notice how the output common-mode voltage is wandering.

Adding CMFB

Figure 26.26 shows how we can modify the basic op-amp to allow for a CMFB signal input. If the outputs of the op-amp (their average or common-mode voltage) are too high,

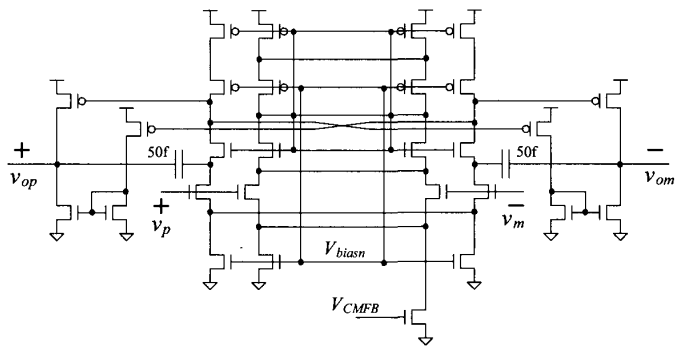


Figure 26.26 Modifying the op-amp for a CMFB input signal.

IPR2023-00698
Theta EX2019

V_{CMFB} goes up. This causes the output voltage of the diff-amp to go up (increasing V_{CMFB} causes the bias current in the center MOSFETs of the diff-pair to increase). An increase in the diff-amp's output voltage lowers the quiescent current flowing in the output buffers, causing the output voltage to move downwards. This is a good time to remember one of the fundamentals from Ch. 20, namely, if two MOSFET gate-source voltages are equal and they have the same drain current, then their drain-source voltages must be equal. For the output buffer in Fig. 26.26, this means that as we reduce the drain currents flowing in the output buffer (by driving V_{CMFB} high), the gate-source voltages of the NMOS devices decrease. Because of the fundamental concept just mentioned, this causes the output voltages of the op-amp to decrease. Figure 26.27 shows the op-amp's output voltage change with V_{CMFB}. For stability concerns, it's of interest to determine the gain from the CMFB input to the outputs. From the simulation, the gain is approximately 25 (about 10 times less than the differential gain). The lower forward CMFB gain allows us to use a diff-amp with a current mirror load for the CMFB amplifier (discussed next).

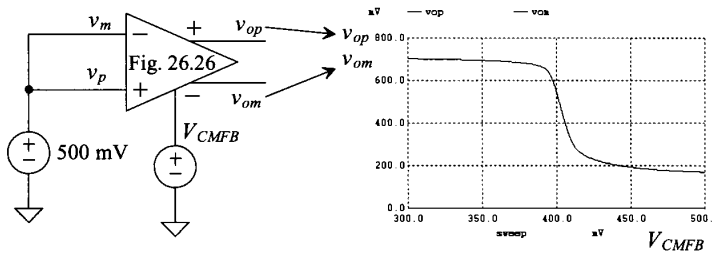


Figure 26.27 The CMFB input to output relationship. The gain is approximately 25 (considerably less than the forward differential gain).

CMFB Amplifier

Now that we know how our CMFB signal, V_{CMFB}, in the op-amp of Fig. 26.26 affects the output voltages and that the gain this signal sees is approximately 25 ($v_{op,n}/V_{CMFB}$), we need to discuss the CMFB amplifier. Consider the two diff-amps seen in Fig. 26.28. In the top diff-amp, gate-drain-connected loads are used. As seen in the simulations, the gain from the averaged input, V_{CMA}, to the diff-amp's output, V_{CMFB} is considerably less than 1 (and so the CMFB is guaranteed to be stable). Unfortunately, the CMFB amplifier's output voltage isn't high enough. (As seen in Fig. 26.27, we need approximately 400 mV.) The diff-amp's tail current can be increased in size (use at least two PMOS for the tail current) to increase the output voltage. However, the low gain, say around 0.3, combined with the CMFB gain through the op-amp, again around 25, means that the CMFB loop's overall gain is only around 7. This isn't large enough to precisely balance the outputs. Using the diff-amp with current mirror load, Fig. 26.28b, gives a gain of approximately 8. This combined with the op-amp's CMFB gain gives an overall CMFB loop gain of 200. This is less than the op-amp's differential gain (so we can use the same compensation capacitors to stabilize the CMFB loop), Fig. 26.23, and large enough to precisely balance the op-amp's outputs.

IPR2023-00698
Theta EX2019

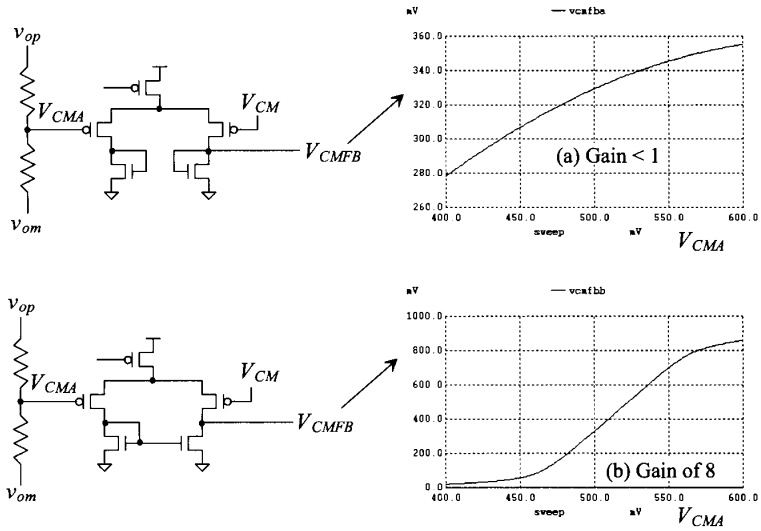


Figure 26.28 Gains of CMFB amplifiers.

The Two-Stage Op-Amp with CMFB

Figure 26.29 shows the complete schematic of the op-amp. Notice that we've doubled the width of the MOSFETs in the output buffer. The added loading from the CMFB averaging resistors will lower the open-loop gain. To compensate for this reduction, the

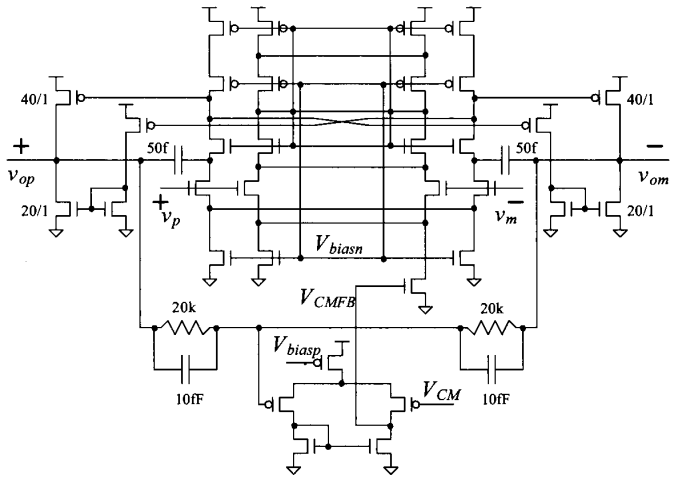


Figure 26.29 Complete schematic of op-amp with CMFB.

IPR2023-00698
Theta EX2019

drive strength of the MOSFETs on the output of the op-amp was increased. Figure 26.30 shows the DC behavior of the op-amp when placed in the configuration seen in Fig. 26.23. Notice that the outputs cross at the ideal common-mode voltage of 500 mV. The gain is reduced because of the loading by the 20k resistors in the CMFB circuit (and so the gain can be increased by increasing the values of the resistors in the CMFB circuit). However, **we see a problem** with these simulation results. The outputs are only swinging from 200 to 800 mV (not from 0 to *VDD* as possible with a push-pull output stage). Further, we can estimate the current pulled from *VDD* for the op-amp in Fig. 26.29 by counting the number of branches in the op-amp (noting that the output stage counts twice because we've doubled the widths of these devices). Including the CMFB amplifier, there are 11 branches. If the bias current through each branch is 20 μA and the current pulled by the bias circuit is 50 μA, then we would expect the op-amp to pull 270 μA (perhaps even a little less because many of the MOSFETs are biased near the triode region). When we look at the simulation results that generated Fig. 26.30, we see that the current is closer to 500 μA (way off indicating, again, that we've got a problem).

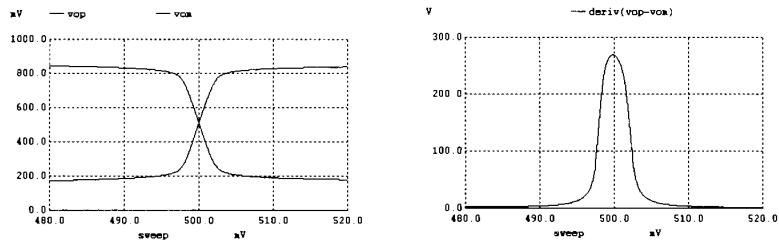


Figure 26.30 Simulating the operation of the op-amp in Fig. 26.29.

Origin of the Problem

The origin of the problem (that is, the op-amp drawing too much current and the output swing not reaching the rails) can be traced to the output buffer. We spent a considerable amount of time discussing how we want to set the outputs of the diff-amp to V_{biasp}. When we look at the simulations, we see the diff-amp's outputs are considerably less than the ideal 600 mV of V_{biasp}. To understand why, let's look at the output buffer in Fig. 26.31. Since the problems appear with the addition of the CMFB circuit, we assume that the two gates of the PMOS devices are moving at the same potential (tied together), that is, that the inputs to the buffer are moving with the diff-amp's output common-mode level. When the output of the diff-amp is V_{biasp} (600 mV), 20 μA flows in all of the MOSFETs. (To keep things simpler, we don't include the doubling in the widths used in the op-amp output devices.) The gate potentials of the NMOS are, roughly, V_{biasn} (400 mV). Because of the symmetry of the circuit, this means that the output is also at 400 mV. To increase the output voltage to 500 mV, we must drop the potential on the gates of the PMOS until the gate-drain-connected NMOS has a V_{GS} of 500 mV (again because of the symmetry). This increases the current (significantly) flowing in the output buffer, lowers its gain, and reduces the linear output swing. The overdrive voltages essentially change from the desired 100 mV when 20 μA of current flows to 200 mV when the output is driven to 500 mV.

IPR2023-00698
Theta EX2019

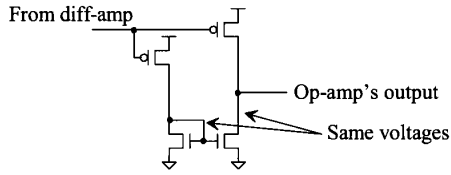


Figure 26.31 Output buffer used in the op-amp of Fig. 26.29.

Figure 26.32 shows one solution to this problem. We've added a device to cascode the output buffer's NMOS (the one connected to the output terminal). The speed shouldn't be affected by the addition of the device (which operates near or in the triode region). The added device allows the op-amp's output to swing more freely (but the current still won't be precisely set). The added device won't affect the output swing range of the op-amp. Figure 26.33 shows the op-amp with modified output buffer.

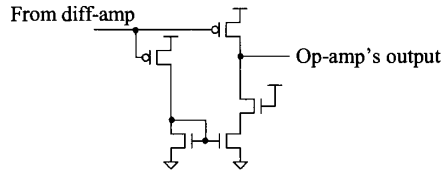


Figure 26.32 Adding a device to allow the output voltage to swing.

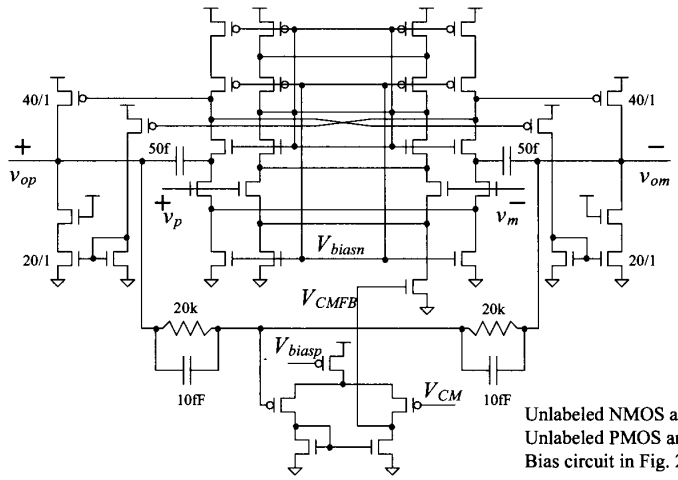


Figure 26.33 Op-amp with modified output buffer.

IPR2023-00698
Theta EX2019

Simulation Results

Figures 26.34 to 26.36 show simulation results based on the circuit topologies in Figs. 26.23 to 26.25, respectively, using the op-amp in Fig. 26.33. In Fig. 26.34 we see that the outputs now swing close to the power supply rails (unlike what we saw in Fig. 26.30). Also, the gain is higher. Figure 26.35 shows the step response of the amplifier in the topology seen in Fig. 26.24. Finally, Fig. 26.36 shows the outputs of the sample-and-hold in Fig. 26.25 using the op-amp in Fig. 26.33.

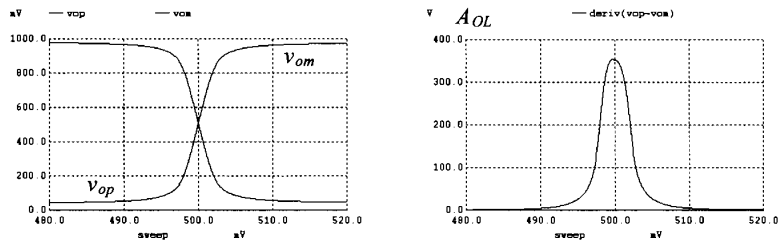


Figure 26.34 Resimulating the op-amp in Fig. 26.33 in the configuration seen in Fig. 26.23.

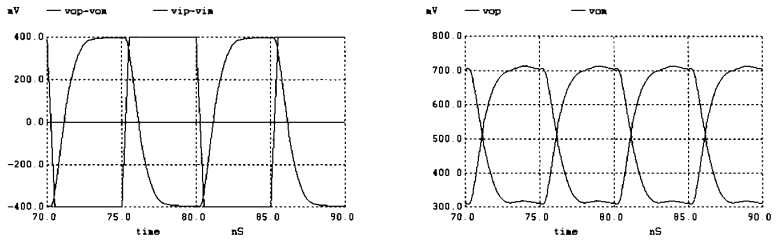


Figure 26.35 Regenerating the simulation results using the topology in Fig. 26.24 with the op-amp in Fig. 26.33.

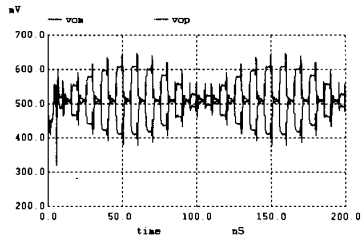


Figure 26.36 Using the op-amp in Fig. 26.33 in the sample-and-hold circuit of Fig. 26.25. Figure shows the outputs of the op-amp.

IPR2023-00698
Theta EX2019

Using MOSFETs Operating in the Triode Region

For the output buffer in Fig. 26.32, we added a transistor in the drain portion of the circuit. We might wonder if we can accomplish better control of the current flowing in the output buffer by adding circuitry to the source side of the buffer. Towards answering this question, consider the portion of the output buffer seen in Fig. 26.37. If both outputs are at V_{CM}, the gate-source voltages of M2 and M3 are the same and so are the gate-source voltages of M1 and M4. If one output goes high and the other output goes low (and the outputs are centered around V_{CM}), then the net drain current of each M3 is constant. If one output goes below the threshold voltage of an NMOS device, then the balancing stops working. The practical problem of using triode-operating MOSFETs (in a CMFB circuit or any type of amplifying configuration) is that the gain through a MOSFET operating in the triode region is low (so it's difficult to provide control).

A portion of the output buffer.

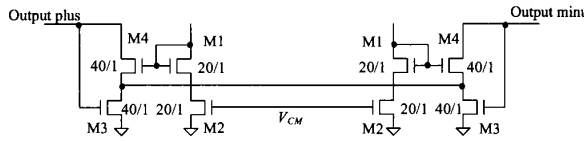


Figure 26.37 Using triode-operating MOSFETs to balance the outputs (bad).

Start-up Problems

Examine the circuit in Fig. 26.38 using the op-amp in Fig. 26.33. Because the op-amp's inputs are at 0 V, the diff-amps on the input of the op-amp are off. This causes the gates of the PMOS devices in the output buffer to be pulled to *VDD*. The circuit remains in this state and doesn't move the outputs or the inputs up to V_{CM}. To avoid this start-up problem, we need to ensure that there is some DC path to *VDD* or V_{CM} to "start-up" the op-amp.

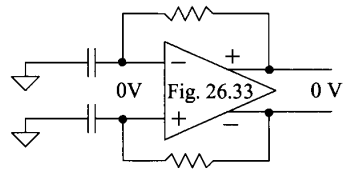


Figure 26.38 The op-amp in Fig. 26.33 won't turn on in this topology.

Lowering Input Capacitance

Consider connecting the gates of the NMOS diff-amp used for biasing to V_{CM} (see the bold line in Fig. 26.39). This reduces the input capacitance of the op-amp and shouldn't affect the normal operation as long as the common-mode voltage of the op-amp (the gates

IPR2023-00698
Theta EX2019

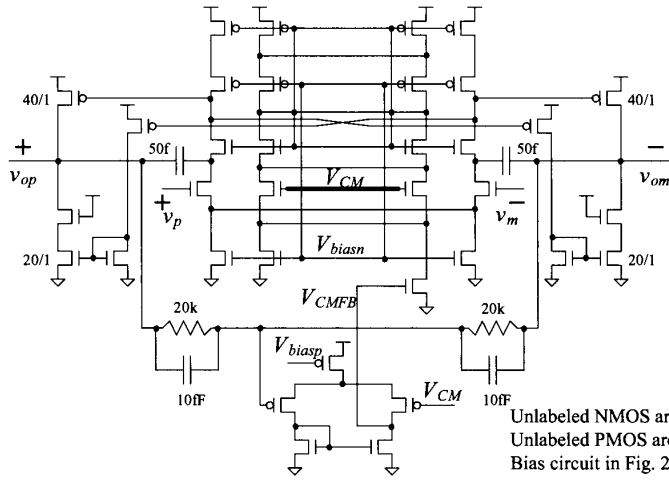


Figure 26.39 Connecting the bias circuit diff-amp's inputs to the common-mode voltage.

of the other diff-amp) is V_{CM}. Simulating the operation of the op-amp in Fig. 26.39 to generate the data in Figs. 26.34 to 26.36 (in the same topologies), we see no change in the simulation data (the netlists used to verify this statement can be found at cmosedu.com).

Looking at Fig. 26.39, we might now wonder why we need two biasing branches down the middle of the op-amp. Further, looking at the voltages in Fig. 26.19, we might wonder if we can make the op-amp more tolerant to changes in *VDD*. The way the op-amp is biased now the open-loop gain drops significantly if *VDD* drops to 900 mV.

Making the Op-Amp More Practical

As just mentioned, having two identical branches down the middle of our op-amp wastes power (although it is useful for a symmetrical layout). We can cut one of the branches out of the design and reduce the power dissipated by the op-amp. We might further wonder if using V_{biasn} for biasing the cascode PMOS current sources is such a good idea. Reviewing Fig. 26.4, we see that V_{biasn} essentially stays at 400 mV after the reference turns on (*VDD* gets above a certain value). In a practical op-amp, we want this voltage to decrease as *VDD* drops in an effort to keep the PMOS devices operating in the saturation region.

Examine the op-amp in Fig. 26.40. We've added a wide-swing bias circuit, see Fig. 20.38, and cut out one of the biasing branches used in the op-amp seen in Fig. 26.39. The power dissipation remains essentially the same as in the previous op-amp topologies. Note how we apply the CMFB to one side of the bias circuit (not to the 10/3 wide-swing bias branch). Further notice how we've reduced the gain of the CMFB loop by using two 10/2 MOSFETs. It may be a good idea to reduce the strength of the CMFB loop even further by increasing the lengths of the NMOS devices. For example, we might change these devices from 10/2 to 10/4. The ability to drive V_{CMFB} considerably above V_{biasn} eliminates the concern that the NMOS device can sink the current needed to bias the circuit at the correct point (and why the CMFB loop can so easily become unstable).

IPR2023-00698
Theta EX2019

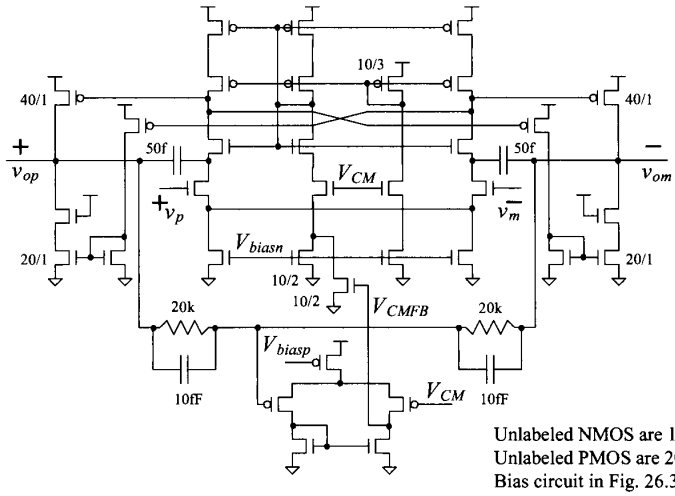


Figure 26.40 Making the op-amp more practical.

A more practical problem is the change in the MOSFET's drain currents with drain-source voltages. For example, looking at Figs. 26.18a and b, we see that it's possible for the NMOS's drain current to be twice the PMOS's drain current for the same gate-source voltages. This can cause op-amp failure in the CMFB circuit. If, for example, the current flowing in the 10/2 MOSFET connected to V_{biasn} is larger than the current sourced by the PMOS, the voltage V_{CMFB} goes to zero and the CMFB loop doesn't work properly. A good "rule-of-thumb" is for the fixed current flowing in the CMFB-controlled bias circuit to be 25–50% of the total expected current.

Increasing the Op-Amp's Open-Loop Gain

The op-amp that we've developed in this section has an open-loop gain in the hundreds. As we'll see in Ch. 29 (Eq. [29.59]), the open-loop gain of the op-amps used in a data converter has a direct effect on the maximum attainable resolution. Towards increasing the gain of the op-amp, let's use the gain-enhancement (GE) techniques presented in Sec. 24.4.

Figure 26.41 shows how we should *not* implement GE. An amplifier with fully-differential outputs, like the amplifier in Fig. 26.7, regulates the drains of the top PMOS devices. We know that this is bad because we would need a CMFB circuit to balance the outputs of the added amplifier.

Figure 26.42 shows how we can add GE to the op-amp developed in this section. Notice that GE is implemented with amplifiers having single-ended outputs. When designing the added amplifiers, as discussed in Sec. 24.4, the bandwidth isn't important for high-speed operation. The lengths of the MOSFET in the added amplifiers can be increased to reduce power and boost gain. An important concern is the added amplifier's input common-mode range.

IPR2023-00698
Theta EX2019

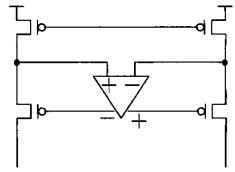


Figure 26.41 How not to implement GE in an op-amp (unless the added amplifier employs CMFB).

The practical problem with the topology seen in Fig. 26.42 is the implementation of the CMFB. With the GE added to the op-amp we now have four additional feedback loops. Variations in V_{CMFB} affect all GE loops in addition to the CMFB through the op-amp. Making these loops stable becomes extremely challenging. What we need is to implement the CMFB without including the diff-amp and GE amplifiers. We can only do this by controlling the output buffer common-mode level, Fig. 26.43. This circuit includes the output buffer we used in Fig. 26.37. Now, however, we add an amplifier in series with the triode-operating MOSFETs to boost the CMFB gain (so that the outputs can be balanced around V_{CM}). The problem with this approach is that the CMFB isn't compensated using the same capacitors as the differential forward signal path (i.e., no Miller effect). We must be concerned with CMFB stability.

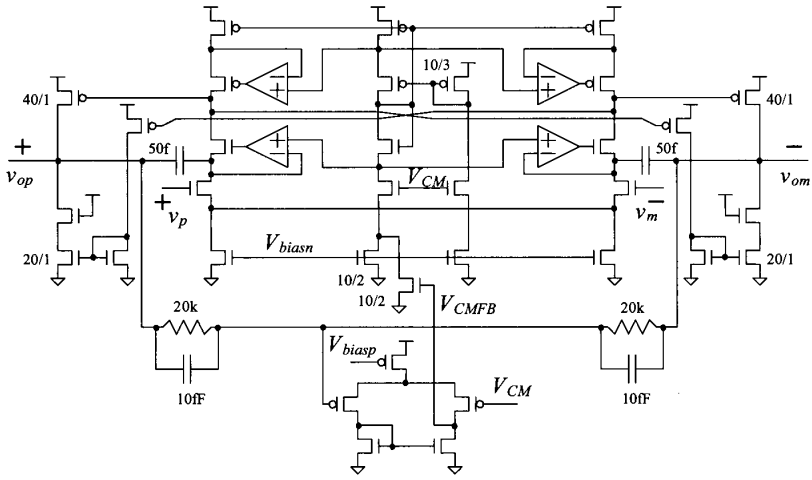


Figure 26.42 Adding gain-enhancement to the op-amp.

Let's discuss the op-amp design in Fig. 26.43. We begin by showing the DC behavior of the op-amp (to show that the outputs are indeed balanced). Figures 26.44a and (b) show the DC behavior and gain of the op-amp in Fig. 26.43 in the topology seen in Fig. 26.24. Again note that by increasing the values of the 20k resistors used to average

IPR2023-00698
Theta EX2019

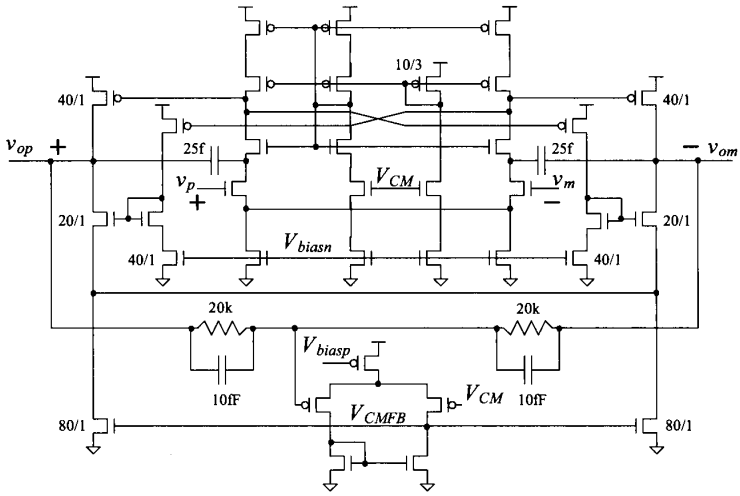
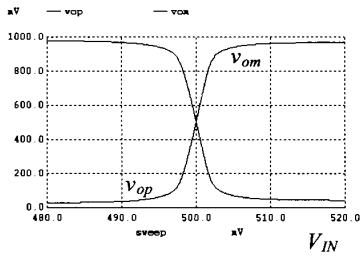
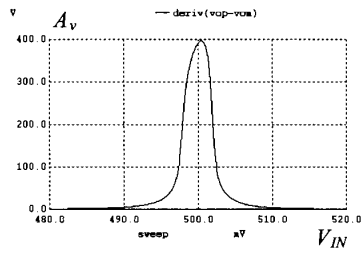


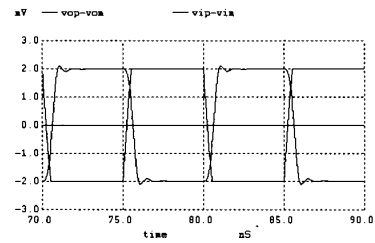
Figure 26.43 Providing CMFB through just the output buffer. Using an amplifier with triode-operating MOSFETs for CMFB (good).



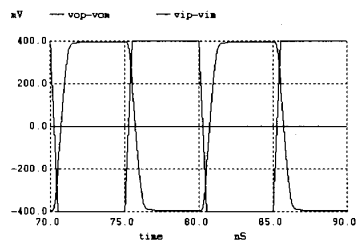
(a) DC behavior



(b) Op-amp gain



(c) Small-signal step response



(d) Large-signal step response

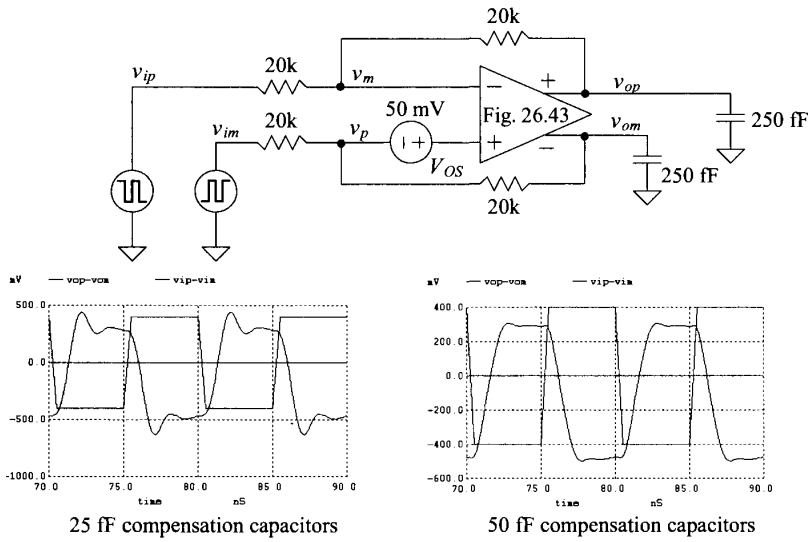
Figure 26.44 Behavior of the op-amp in Fig. 26.43 (see text).

IPR2023-00698
Theta EX2019

the outputs we can increase the DC gain of the op-amp. Because the forward differential path compensation (now) doesn't include the CMFB path, we've reduced the compensation capacitors from 50 fF to 25 fF (there are practical issues with this change that we'll discuss next). Note that the widths of the added triode-operating MOSFETs are increased to ensure that they operate well within the triode region and don't significantly affect the output drive capability of the op-amp. Figures 26.44c and (d) show the small- and large-signal step responses (see Fig. 26.24). The settling times are approximately 1–2 ns. The current drawn from *VDD* (including the the bias circuit current) is approximately 250 μA (200 μA for the op-amp alone).

Offsets

We said in Sec. 26.2 that we want to add offsets of 50 mV in series with the gates of our MOSFETs to see if the op-amp "breaks." Consider the addition of an input-referred offset voltage in the schematic seen in Fig. 26.45. This offset is used to model the overall offset voltage of the op-amp (see Fig. 24.4). If we look at the effects of the offset on the circuitry in Fig. 26.43, we see that one of the NMOS devices in the diff-pair will have a higher overdrive voltage (larger g_m) than the other NMOS device. Since the unity-gain frequency of an op-amp is given by $g_m/2\pi C_c$, the effect is a shift in the unity-gain frequency (and potential instability). As seen in Fig. 26.45, using the 25 fF compensation capacitors destabilizes the op-amp (as indicated by the ringing). Increasing the compensation capacitor's value (back) to 50 fF makes the step response cleaner. We could argue that 50 mV is an unrealistically high offset voltage, so it's better to leave the capacitors at 25 fF (and this may be the case). However, in a practical CMOS process, the op-amp's characteristics shift with the process shifts (and temperature). It's better to overcompensate than to have an unstable op-amp.



25 fF compensation capacitors

50 fF compensation capacitors

Figure 26.45 How an offset can affect the step response (compensation).

IPR2023-00698
Theta EX2019

Op-Amp Offset Effects on Outputs

Notice, in Fig. 26.45, that an offset shifts the differential output signal by twice the offset voltage. The individual op-amp output voltages remain centered around V_{CM}. To describe this in more detail, consider the test setup seen in Fig. 26.46. We know the CMFB forces

$$\frac{v_{op}+v_{om}}{2}=V_{CM} \text{ or } v_{op}=2V_{CM}-v_{om} \tag{26.8}$$

Further, assuming large op-amp open-loop gain,

$$v_p+V_{OS}=v_{pos}\approx v_m \tag{26.9}$$

Equating currents, we can write

$$\frac{V_{CMI}-v_p}{R_{in}}=\frac{v_p-v_{om}}{R_f} \text{ or } \frac{V_{CMI}-v_m+V_{OS}}{R_{in}}=\frac{v_m-V_{OS}-v_{om}}{R_f} \tag{26.10}$$

and

$$\frac{V_{CMI}-v_m}{R_{in}}=\frac{v_m-v_{op}}{R_f} \tag{26.11}$$

Subtracting Eq. (26.11) from Eq. (26.10), we get

$$\frac{V_{OS}}{R_{in}}=\frac{-V_{OS}+v_{op}-v_{om}}{R_f} \tag{26.12}$$

or

$$v_{op}-v_{om}=V_{OS}\cdot\left(\frac{R_f}{R_{in}}+1\right) \tag{26.13}$$

The voltages on the outputs of the op-amp are then

$$v_{op}=\frac{V_{OS}}{2}\cdot\left(1+\frac{R_f}{R_{in}}\right)+V_{CM} \text{ and } v_{om}=V_{CM}-\frac{V_{OS}}{2}\cdot\left(1+\frac{R_f}{R_{in}}\right) \tag{26.14}$$

The op-amp's input voltages are

$$v_m\approx v_p+V_{OS}=\frac{V_{OS}}{2}+V_{CM}\cdot\left(\frac{R_{in}}{R_{in}+R_f}\right)+V_{CMI}\cdot\left(\frac{R_f}{R_{in}+R_f}\right) \tag{26.15}$$

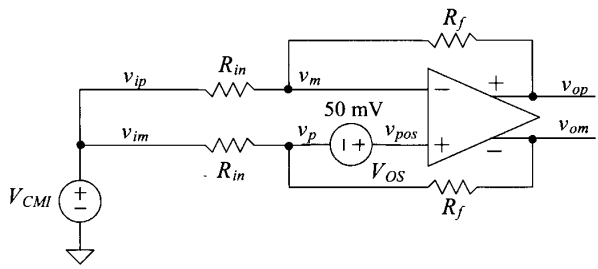


Figure 26.46 How an offset voltage causes an imbalance in the outputs.

IPR2023-00698
Theta EX2019

If the input signal's common-mode voltage, V_{CMI}, is the same as the op-amp's common-voltage, V_{CM}, and V_{OS} is zero, then both op-amp inputs are held at V_{CM}. If $V_{CM} = V_{CMI}$ but the offset isn't zero, then $v_m = V_{OS}/2 + V_{CM}$ and $v_p = V_{CM} - V_{OS}/2$. Neglecting the offset voltage, if the input signal's common mode voltage isn't V_{CM}, then the op-amp's input common-mode voltage ($v_p = v_m$) will be different from the ideal value of V_{CM}. This can shut the op-amp off and cause undesirable behavior.

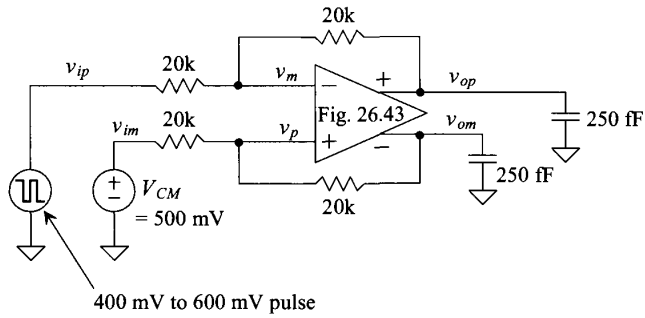


Figure 26.47 Problems with single-ended input signals.

Single-Ended to Differential Conversion

Even if the input signals are referenced around V_{CM} and the op-amp is offset free, we can still have problems. Consider the circuit in Fig. 26.47. In this circuit the input signal is single-ended. The other input to the op-amp is tied to V_{CM}. Even though the input is balanced around 500 mV, it is not a truly differential signal. When the input signal is up at 600 mV, the effective input common-mode voltage, V_{CMI}, is 550 mV (the average of the two inputs). When the input signal is at 400 mV, V_{CMI} is 450 mV. Figure 26.48 shows the simulation results using the circuit in Fig. 26.47. These results are very interesting. Consider what's happening between 70 and 75 ns. During this time, v_{ip} is 600 mV and v_{im} is 500 mV (and so $v_{ip} - v_{im} = 100\ mV$). The op-amp's inputs move above the ideal 500 mV, as seen in the figure. This has the effect of increasing the transconductance of the diff-amp (the diff-amp's tail current is operating near/in the triode region so the tail

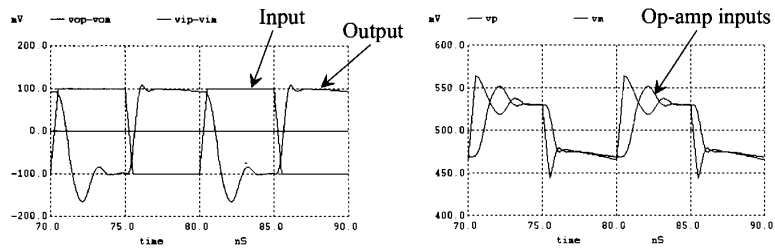


Figure 26.48 Simulating the operation of the circuit in Fig. 26.47.

IPR2023-00698
Theta EX2019

current increases as the input common-mode voltage goes up). The increase in the diff-amp's g_m causes the unity gain frequency to increase and, with a fixed compensation capacitor, the phase margin to decrease (causing overshoot and ringing in the circuit's step response). Next consider what happens between 75 and 80 ns. The input signal switches to $400\ mV - 500\ mV = -100\ mV$ and the op-amp's inputs drop down to around 475 mV. This causes the diff-amp's g_m to decrease and the op-amp to slow down (become more stable). However, at such a low input common-mode voltage the gain of the op-amp will drop and the outputs may wander. (See Sec. 30.3.1 for design information).

CMFB Settling Time

If we look at the voltage, V_{CMFB}, in the op-amps in Fig. 26.40 or 26.43, we may see that they aren't settling as quickly as the differential mode signals. Is this bad? As seen in Eq. (26.13), variations in V_{CMI} don't (ideally) affect the differential output signal. However, if the variations in V_{CMFB} and thus V_{CMI} are large, one of the output signals can saturate at close to *VDD* or ground (the op-amp's gain drops), affecting the differential signal path.

CMFB in the Output Buffer (Fig. 26.43) or the Diff-Amp (Fig. 26.40)?

In this section we've presented the design of op-amps using continuous-time CMFB. Various design trade-offs and topologies were presented. At this point a good question is: "Which CMFB scheme is better?" Controlling the output common-mode level through the output buffer (Fig. 26.43) is simple, easy to ensure stability, and fairly robust. However, consider what happens if the op-amp in Fig. 26.43 is used in the topology in Fig. 26.46 (without an offset). If V_{CMI} is ground, then, according to Eq. (26.15), the input voltages to the op-amp will move to 250 mV while the outputs of the op-amp remain at 500 mV. Thinking about this for a moment and neglecting the fact that an op-amp input common-mode voltage of 250 mV will shut the op-amp off, we see that the outputs of the op-amp must source a current back through the feedback resistors to the inputs. Further, anytime $V_{CMI} < V_{CM}$, the PMOS in the output buffer must source a DC current back to the inputs. Similarly, if $V_{CMI} > V_{CM}$, the NMOS in the output buffer must sink a current from the inputs of the op-amp. When using CMFB in the output buffers, we have no way of increasing the quiescent or DC current flowing in the output buffer to source/sink current from the input source (or to a DC load connected to ground). The output voltage of the diff-amp in the op-amp of Fig. 26.43 biases, or sets, the quiescent current in the output buffer. The CMFB scheme used in the output buffer simply adjusts the drive strength of the NMOS (in the output buffer) to set the op-amp's output common-mode level.

Using the CMFB scheme in Fig. 26.40, we can adjust the output voltage of the diff-amp (and ultimately the op-amp's output voltage) and thus the bias current in the output buffer (both the NMOS and the PMOS connected to the output buffer). However, if the output buffer must source/sink a significant amount of current, we can have problems with this topology too. For example, if the diff-amp's output voltage increases (turning off the PMOS in the output buffer and attempting to pull the op-amp's outputs down) while the NMOS devices connected to the outputs are sinking significant current, then it's possible that the outputs will get pulled upwards and the CMFB will fail (likely causing the gain of the op-amp to drop and the outputs to have limited swing).

To ensure the most robust op-amp design (biasing tolerant to offsets and easy to compensate the CMFB loops), CMFB circuits can be placed around both op-amp stages. This method is used in our last op-amp design example discussed next.

IPR2023-00698
Theta EX2019

26.4 Op-Amp Design Using Switched-Capacitor CMFB

In this section we turn our attention towards op-amp designs for switched-capacitor (SC) circuits. We'll develop an op-amp design based on the topologies discussed in the last section. In this section, however, we'll use SC CMFB instead of continuous-time CMFB. Again, we'll use the nanometer CMOS process with minimum lengths and (roughly) 100 mV overdrive voltages (for drain currents of 20 μA, see Fig. 26.18). We'll stick with two-stage designs because of the low open-circuit gains present when doing high-speed design.

Clock Signals

The SPICE listing for the clock signals used in the simulations in this section is seen below. Note that the 100 MHz phi1 and phi2 clocks are used with the NMOS transistors (phi1 and phi2 are not high at the same times), while the complements of these clock signals are used with the PMOS switches (not low at the same times).

```
*Clock Signals
Vphi1  phi1   0 DC 0  Pulse  0 1 0  200p  200p 4n 10n
Vphi1b phi1b  0 DC 0  Pulse  1 0 0  200p  200p 4n 10n
Vphi2  phi2   0 DC 0  Pulse  0 1 5n 200p 200p 4n 10n
Vphi2b phi2b  0 DC 0  Pulse  1 0 5n 200p 200p 4n 10n
R1     phi1   0      1MEG
R2     phi1b  0      1MEG
R3     phi2   0      1MEG
R4     phi2b  0      1MEG
```

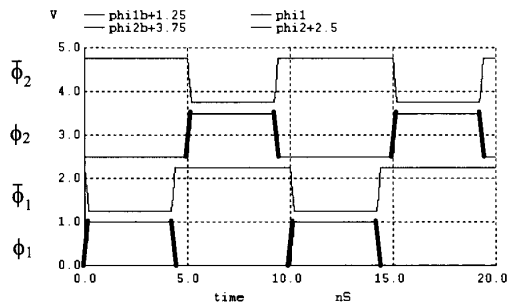


Figure 26.49 Generating nonoverlapping clocks for SC circuits.

Switched-Capacitor CMFB

Figure 26.50 shows the SC CMFB circuit from Fig. 26.16 along with a symbolic representation. We've selected 10 fF capacitors for the high-speed averaging capacitors. The kT/C noise associated with these capacitors is a common-mode signal and so it shouldn't affect the differential-mode signal. We don't want these capacitors to be too large because they will load the output of the amplifier. We made the switched-capacitors associated with the differencing and averaging 50 fF so that the changes in V_{CMFB} during one clock cycle won't be too large due to differences between the ideal V_{CMFB} and the actual V_{CMFB}. When the average of the outputs, v_{op} and v_{om}, is above their ideal value, the

IPR2023-00698
Theta EX2019

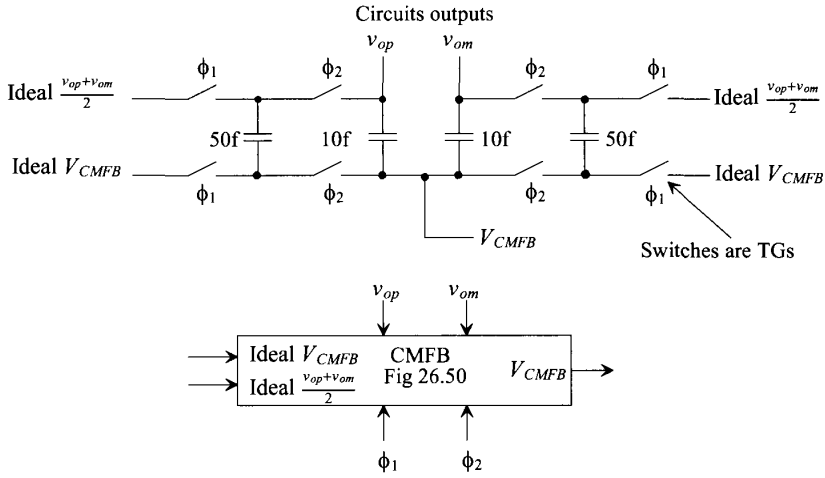


Figure 26.50 A switched-capacitor CMFB circuit (see Fig. 26.16). Switches implemented with transmission gates.

CMFB circuit's outputs moves upwards. It's important that this causes the average value of the outputs to move downwards. In other words, we've got to ensure that negative feedback is used in the CMFB loop.

Figure 26.51 shows some simulation results using the SC CMFB in Fig. 26.50. The ideal V_{CMFB} is 400 mV. The ideal average output, $\frac{v_{op}+v_{om}}{2}$, is 500 mV (V_{CM}). Prior to 150 ns, both v_{op} and v_{om} are 500 mV. At 150 ns, these two voltages jump (in the simulation) to 700 mV (they jump 200 mV away from the ideal average output of 500 mV). The voltage, V_{CMFB}, also jumps at this time. The key point to notice is that V_{CMFB} moves in the same direction as do changes in the output common-mode voltage. Knowing this is important when we are ensuring that our CMFB loop employs negative feedback.

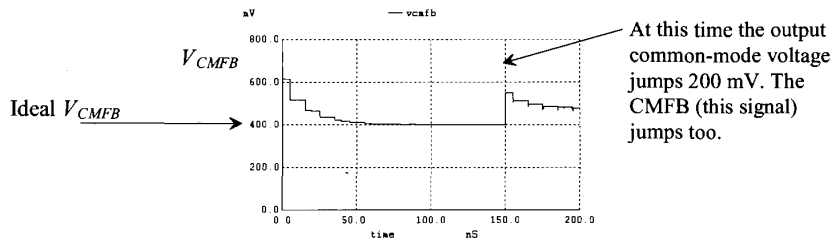


Figure 26.51 How the common-mode feedback signal increases to pull the outputs down.

IPR2023-00698
Theta EX2019

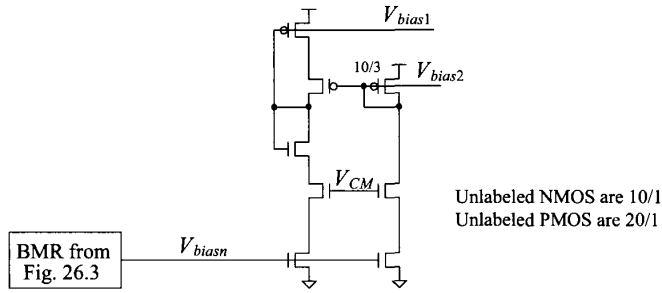


Figure 26.52 Biasing circuit for the op-amp developed in this section.

The Op-Amp's First Stage

As just mentioned, the output of the SC CMFB circuit, V_{CMFB} , moves in the same direction as movement in the average of the amplifier's outputs, v_{op} and v_{om} (the output common-mode level). Keeping this in mind, consider the bias circuit and diff-amp schematics seen in Figs. 26.52 and 26.53. We've separated the bias circuit out from the diff-amp, see Fig. 26.43, for a couple of reasons. To begin, an increase in the CMFB signal in Fig. 26.43 caused the outputs of the diff-amp to increase. As just mentioned, we want the diff-amp's outputs to move in the opposite direction of V_{CMFB}. Moving the CMFB-controlled MOSFET into the tail current of the diff-amp provides this control. Next, our bias circuit may now be shared with several op-amps. The benefit of sharing the bias circuit is a reduction in power dissipation.

Before adding the CMFB circuit, let's look at a SC circuit using an op-amp. Reviewing the sample and hold in Fig. 26.25, we see that when the ϕ_1 switches are closed, the op-amp's inputs and outputs are at (ideally) V_{CM}. The op-amp, during this time, is in

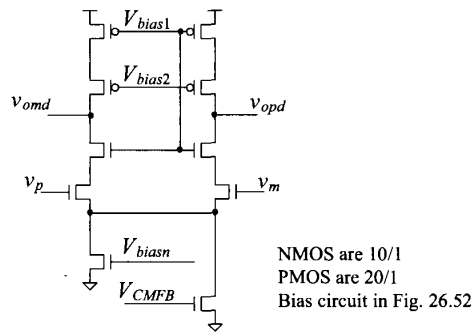


Figure 26.53 Diff-amp used with the bias circuit of Fig. 26.52.

IPR2023-00698
Theta EX2019

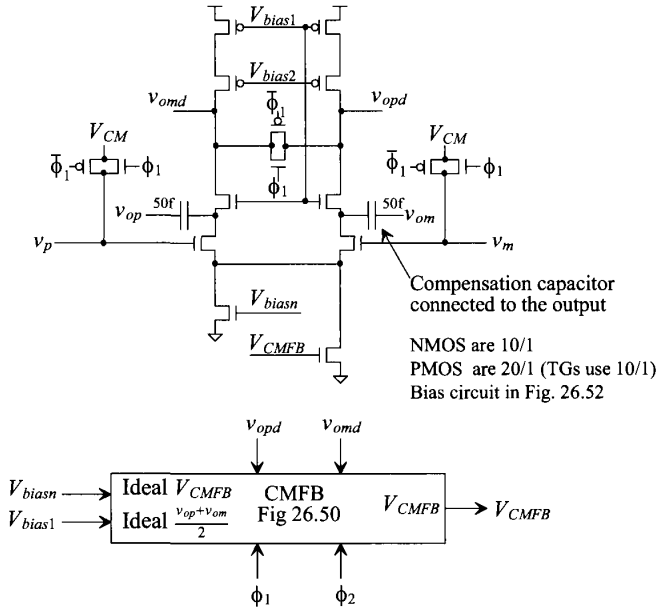


Figure 26.54 First-stage diff-amp with SC CMFB.

the unity-follower configuration. Instead of placing the op-amp in the follower configuration during this time, let's: 1) short the inputs to op-amp (Fig. 26.54) to V_{CM}, 2) short the outputs of the diff-amp (the inputs to the output buffer) together, 3) use a SC CMFB circuit around the diff-amp to ensure a balance condition (and to set the diff-amp's output voltage to V_{bias1}), 4) short the outputs of the op-amp together (so now both the output buffers inputs are shorted and outputs are shorted), and 5) use a SC CMFB around the output buffer to ensure balanced outputs. In other words, during ϕ_1, the differential inputs are shorted together. At the same time, the outputs of the op-amp are shorted together.

Figure 26.55 shows the simulated output of the diff-amp in Fig. 26.54. The SC CMFB circuit sets the outputs of the diff-amp at V_{bias1} or roughly 600 mV. By shorting the two diff-amp outputs together and connecting the inputs to V_{CM} , we are ensuring that even with horrible mismatch, the outputs of the diff-amp will be equal and can be used to bias the output buffer. If, for example, the diff-pair shows a 50 mV mismatch (V_{OS}) then, when ϕ_1 is high, the drain currents in each side of diff-amp are set differently (making this scheme very tolerant to offsets). When ϕ_1 goes low and the op-amp is placed in a feedback configuration, each of the diff-amp's inputs will move (in opposite directions) by $V_{OS}/2$ (and so this topology doesn't store the op-amp's offset voltage).

Note that an important concern, as discussed in the last section, is the input common-mode voltage of the diff-amp. If, for example, the input common-mode voltage

IPR2023-00698
Theta EX2019

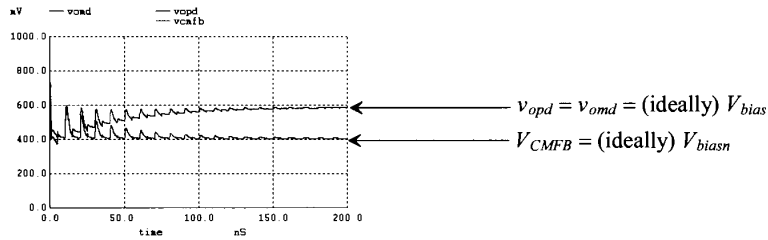


Figure 26.55 Simulating the operation of the circuit in Fig. 26.54.

drops from 500 mV, in Fig. 26.54, to 400 mV the NMOS devices will start shutting off. The op-amp will not settle or behave properly.

The Output Buffer

Figure 26.56 shows a schematic of the output buffer and SC CMFB stage. The nodes v_{omd} and v_{opd} are connected to the diff-amp outputs in Fig. 26.54. When ϕ_1 is high, the outputs of the buffer are shorted together. The SC CMFB is used to set the outputs to the common-mode voltage during this time. Figure 26.57 shows the simulation results for both stages of the op-amp (the op-amp is made with Figs. 26.54 and 26.56) with the inputs and outputs floating. When ϕ_1 is high, the inputs to the op-amp are connected to V_{CM}.

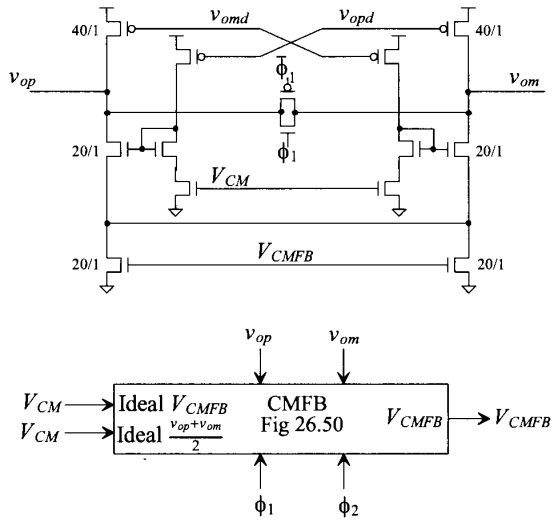


Figure 26.56 Output buffer and CMFB.

IPR2023-00698
Theta EX2019

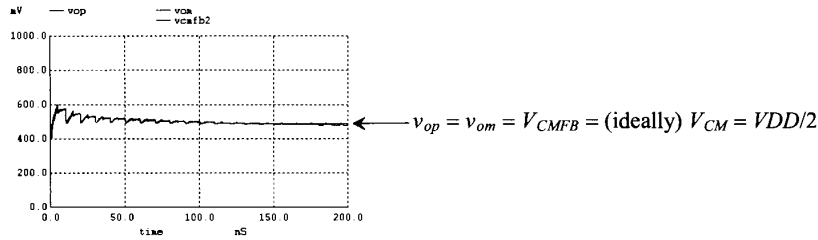


Figure 26.57 Simulating the operation of the op-amp (Figs. 26.54 and 26.56).

Is shorting the outputs of the buffer (or diff-amp) necessary? If the op-amp is operating open-loop (when ϕ_1 goes high), we don't want the outputs to float. By shorting the diff-amp and output buffer output terminals together, we ensure that they are set to a known value. If we were to use this op-amp in a SC integrator, we wouldn't short the inputs of the diff-amp to V_{cm}, the outputs of the diff-amp together, or the output buffer outputs together since there is always feedback around the op-amp. In other words, the same op-amp topology can be used in a SC integrator but without the three TGs seen in Figs. 26.54 and 26.56.

An Application of the Op-Amp

An application of the op-amp we've just developed is seen in Fig. 26.58. This is the sample and hold developed earlier (Figs. 25.15 and 26.25) except that now, during ϕ_1 (high) the inputs to the op-amp are shorted to V_{CM} (Fig. 26.54) and the outputs are shorted together (Fig. 26.56). When ϕ_2 goes high (ϕ_1 is low since the two clocks are nonoverlapping), the op-amp moves into the follower configuration and holds its outputs at the value of the inputs when the ϕ_1 switches shut off.

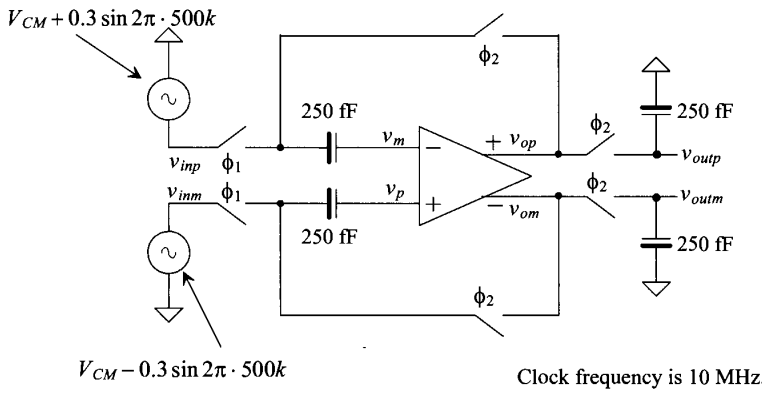
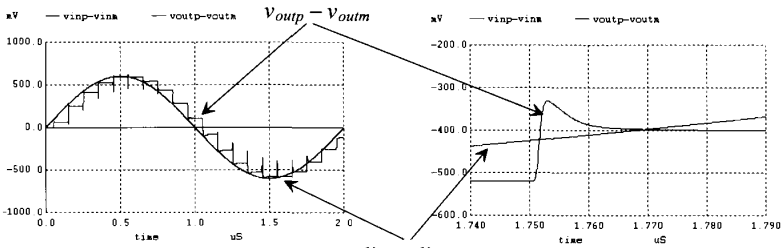
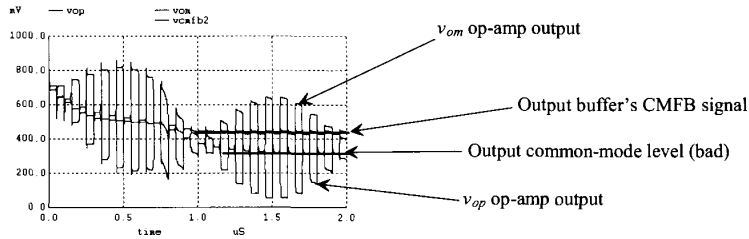


Figure 26.58 Simulating the operation of the op-amp formed with the diff-amp in Fig. 26.54 and buffer in Fig. 26.56.

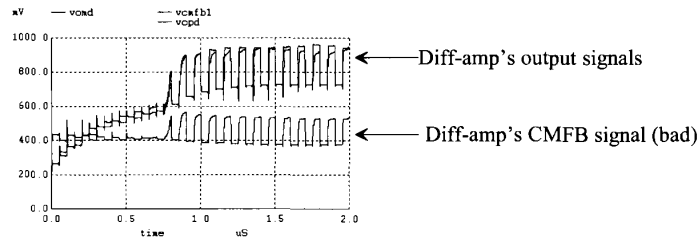
IPR2023-00698
Theta EX2019



(a) Circuits inputs and outputs (b) Slow settling time



(c) Op-amp's output and buffer's CMFB signal



(d) Diff-amp's outputs and CMFB signal

Figure 26.59 Simulating the operation of the circuit in Fig. 26.58 with the op-amp made with the diff-amp in Fig. 26.54 and the output buffer in Fig. 26.56.

Simulation Results

The simulation results are seen in Fig. 26.59. To ensure that we see the full settling behavior of the op-amp and any other issues or concerns, we use a slow clock (10 MHz) for the initial simulations. As we gain confidence that our op-amp is settling correctly and stable, we can increase the clock frequency and the input signal frequency to see how the sample and hold performs. The main thing that's different with this design compared to previous designs is that the op-amp isn't operating with feedback all of the time. When the op-amp does get put into a feedback configuration (ϕ_2 goes high), there will be a start-up time (e.g., the inputs of the op-amp will move away from V_{CM} because of the op-amp's offset).

IPR2023-00698
Theta EX2019

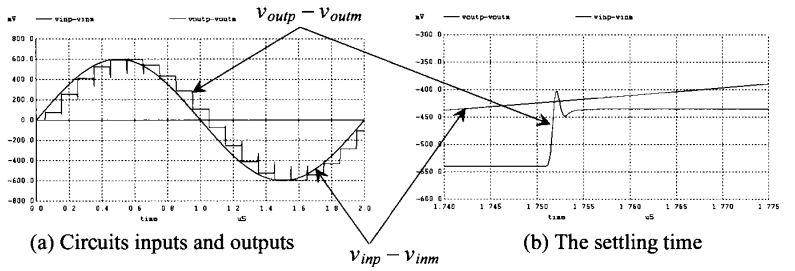
Returning to the simulation results in Fig. 26.59, we see, in (b), that the settling time is quite long (approximately 10 ns). Letting the simulation run for a longer period of time, we see that the settling time increases and, ultimately, that the op-amp stops functioning correctly. In (c) we see the problem: the outputs of the op-amp aren't balanced around V_{CM}. Further, movement downwards causes the common-mode voltage on the input of the diff-amp to drop. The result is that the input diff-amp starts to shut off. The voltage across the diff-amp's tail current source drops and makes it difficult to control the diff-amp's output common-mode level. As seen in (d), the diff-amp's CMFB signal is increasing in an attempt to increase the current flowing in the diff-amp and pull the diff-amp's output voltages downwards. However, because the common-mode voltage on the input of the diff-amp has decreased, the voltage across the tail currents is small (tens of mV). For example, if the input common-mode voltage drops from 500 mV (V_{CM}) to 450 mV and the V_{GS} of the input diff-amp is 400 mV, then only 50 mV is left to drop across the tail current sources (one biased with V_{biasn} and the other with V_{CMFB}). We've discussed this problem earlier.

We can do one of two things to make this decrease in input common-mode voltage less of a problem: 1) reduce the diff-amp's bias current or 2) increase the widths of the input diff-pair. In both cases we need to increase the voltage dropped across the tail current sources by decreasing the V_{GS} of the NMOS diff-pair (the inputs to the diff-amp are, ideally, at V_{CM}). If we decrease the bias current, the g_m of the diff-pair decreases and so does the speed of the op-amp (remembering $f_{un} = g_m/2\pi C_c$). By increasing the width of the diff-pair, we increase g_m and thus the unity-gain frequency of the op-amp. The issue with increasing the width of the diff-pair is that the overdrive voltage for these MOSFETs decreases and thus so does their f_T. If we keep the increase in the width to a modest level, the parasitic pole associated with these diff-pair MOSFETs shouldn't, in any significant way, affect the stability of the op-amp.

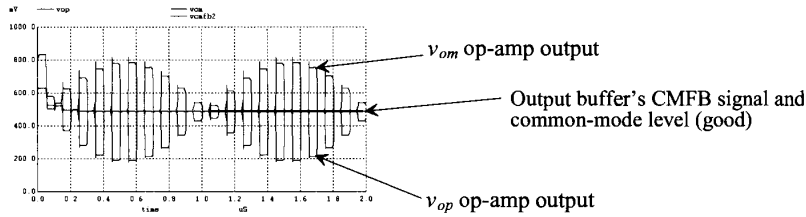
Figure 26.60 shows the operation of the op-amp made with the diff-amp in Fig. 26.54 and the output buffer in Fig. 26.56 if we increase the widths of the diff-pair from 10 to 40. We also changed the MOSFETs in the biasing circuit Fig. 26.52 with gates connected to V_{CM} to have widths of 40 to maintain symmetry (important for reducing the offset voltage). To further increase the speed we also reduced C_c by 4 (from 50 to 12.5 fF) and increased the widths of the devices in the output buffer by 4 (to move f_2 up and away from f_{un}). The settling time, as seen in (b), is < 3 ns. Note that increasing the g_m of the output buffer by using wider devices (more current) increases f_2 allowing a corresponding increase in f_{un}. This is an **important** technique for increasing the speed of an op-amp.

Reviewing the op-amp designs in the last section, we see that their performance may also be improved by a modest increase in the widths of the diff-pair and output buffer. However, as just mentioned, we have to be careful to maintain symmetry. For example, if we take the op-amp in Fig. 26.43, increase the widths of the first stage diff-pair from 10 to 40, and simulate, we see that the outputs don't swing all the way up (or down) to the correct values. Further investigation reveals that the diff-pair moved into the triode region. To move the MOSFETs back into the saturation region, we must also increase the widths of the MOSFETs in the bias circuit (with gates at V_{CM}) from 10 to 40. This adjusts the bias voltages for the NMOS/PMOS diff-amp loads. By providing the CMFB back through the diff-amp (instead of through the output buffer), we can adjust the diff-amp's output voltage. Thus the op-amp is more tolerant to issues in the bias circuit.

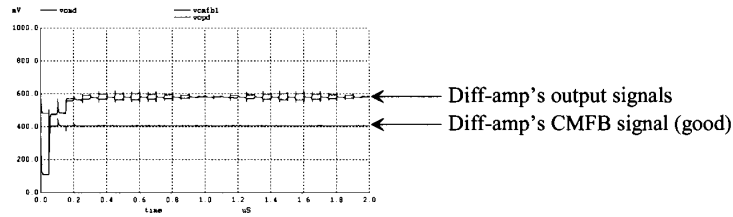
IPR2023-00698
Theta EX2019



(a) Circuits inputs and outputs (b) The settling time



(c) Op-amp's output and buffer's CMFB signal



(d) Diff-amp's outputs and CMFB signal

Figure 26.60 Regenerating the data in Fig. 26.59 after increasing the widths of the input diff-pair and output buffer by 4 while reducing the compensation capacitors by 4 (from 50 to 12.5 fF).

A Final Note Concerning Biasing

In (most) of the designs presented in this chapter, we biased the gates of the NMOS cascode devices in the diff-amp's load at the same (ideally) voltage as the diff-amp's output. For example, in Fig. 26.53, the gate of the NMOS device is tied to V_{bias1}, while its drain (the output of the diff-amp) is also (ideally) at V_{bias1}. This selection is fine for learning design. It minimizes power by avoiding an extra bias reference circuit. However, if *VDD* starts to drop the drain-source voltages across the diff-pair and tail current sources can drop to the point where the op-amp shuts off. Again, increasing the widths of the diff-pair helps with this concern. Also, *VDD* noise can be more of a concern since we usually want the NMOS devices' gates referenced to ground (V_{bias1} is referenced to *VDD*).

A more general, and better, solution is to bias the gates of the NMOS with a gate-drain-connected NMOS, as seen in Figs. 22.30, 26.62, and 26.63. This allows the gate voltage of the NMOS to move to a higher voltage than its drain (allowing the

IPR2023-00698
Theta EX2019

op-amp to function with lower *VDD*). For example, the diff-amp in Fig. 26.53 has a DC output voltage of roughly 600 mV (= V_{bias1}). A more appropriate voltage for the gate of the NMOS device is 800 mV. Remembering that the gate-source voltages of the NMOS diff-pair (Fig. 26.18) are roughly 400 mV puts the drains of the diff-pair at 400 mV. The result: *VDD* is more evenly divided across the devices in the diff-amp.

ADDITIONAL READING

[1] V. S. L. Cheung, H. C. Luong, M. Chan, and W. H. Ki, "A 1-V 3.5mW CMOS Switched-Opamp Quadrature IF Circuitry for Bluetooth Receivers," *Symp. VLSI Circuits Dig.* 16, pp. 140–143, June 2002.

[2] M. Keskin, U. Moon, and G. C. Temes, "A 1-V 10-MHz clock-rate 13-bit CMOS DS modulator using unity-gain-reset opamps," *IEEE Journal of Solid-State Circuits,* vol. 37, pp. 817–824, July 2002.

[3] V. S. L. Cheung, H. C. Luong, and W. H. Ki, "A 1V CMOS switched-opamp switched-capacitor pseudo-2-path filter," *IEEE International Solid-State Circuits Conference,* vol. 35, pp. 154–155, February 2000.

[4] I. E. Opris, L. D. Lewicki, and B. C. Wong, "A single-ended 12-bit 20 Msample/s self-calibrating pipeline A/D converter," *IEEE Journal of Solid-State Circuits,* vol. 33, pp. 1898–1903, December 1998.

[5] D. C. Thelen Jr. and D. D. Chu, "A low noise readout detector circuit for nanoampere sensor applications," *IEEE Journal of Solid-State Circuits,* vol. 32, pp. 337–348, March 1997.

[6] P. D. Walker and M. M. Green, "A tuneable pulse-shaping filter for use in a nuclear spectrometer system," *IEEE Journal of Solid-State Circuits,* vol. 31, pp. 850–855, June 1996.

[7] G. Caiulo, F. Maloberti, G. Palmisano, and S. Portaluri, "Video CMOS power buffer with extended linearity," *IEEE Journal of Solid-State Circuits,* vol. 28, pp. 845–848, July 1993.

[8] K. Nakamura and L. R. Carley, "An enhanced fully differential folded-cascode op amp," *IEEE Journal of Solid-State Circuits,* vol. 27, pp. 563–568, April 1992.

[9] R. Castello, G. Nicollini, and P. Monguzzi, "A high-linearity 50-Ω CMOS differential driver for ISDN applications," *IEEE Journal of Solid-State Circuits,* vol. 26, pp. 1809–1816, December 1991.

[10] S. M. Mallya and J. H. Nevin, "Design procedures for a fully differential folded-cascode CMOS operational amplifier," *IEEE Journal of Solid-State Circuits,* vol. 24, pp. 1737–1740, December 1989.

[11] M. Banu, J. M. Khoury, and Y. Tsividis, "Fully Differential Operational Amplifiers with Accurate Output Balancing," *IEEE Journal of Solid State Circuits,* vol. 23, No. 6, pp. 1410–1414, December 1988.

[12] R. Castello and P. R. Gray, "A high-performance micropower switched-capacitor filter," *IEEE Journal of Solid-State Circuits,* vol. 20, pp. 1122–1132, Dec. 1985.

IPR2023-00698
Theta EX2019

[13] T. C. Choi, R. T. Kaneshiro, R. Broderson, and P. R. Gray, "High-Frequency CMOS Switched Capacitor Filters for Communication Applications," *IEEE Journal of Solid State Circuits,* vol. SC-18, pp. 652–664, December 1983.

[14] D. Senderowicz, S. F. Dreyer, J. H. Huggins, C. F. Rahim, and C. A. Laber, "A family of differential NMOS analog circuits for a PCM codec filter chip," *IEEE Journal of Solid-State Circuits,* vol. 17, pp. 1014–1023, December 1982.

PROBLEMS

Unless otherwise indicated, use the 50 nm CMOS process from the examples in this chapter, the biasing circuit in Fig. 26.3, and 10/1 NMOS and 20/1 PMOS.

26.1 Using simulations, determine the transition frequencies, f_T , for the NMOS and PMOS devices seen in Fig. 26.1 at the nominal operating conditions indicated in the figure. Show that by increasing the MOSFET's overdrive voltage, the f_T s of the MOSFETs increases.

26.2 Simulate the operation of the two-stage op-amp in Fig. 26.2. Show that the quiescent current in the output buffers is considerably below the desired 20 μA. (Note that the inputs of the op-amp should be held at V_{CM} in the simulation to keep the diff-amp conducting current.) Does this affect the speed of the output buffer? Why or why not?

26.3 Plot reference current against resistor value for the BMR seen in Fig. 26.3. Use simulations to determine the I_{REF} for each value of resistance.

26.4 Comment on the benefits and/or concerns with the CMFB input connections seen in Fig. 26.61. Use simulations to support your answers. Note the similarity to the way CMFB was implemented in Fig. 26.17.

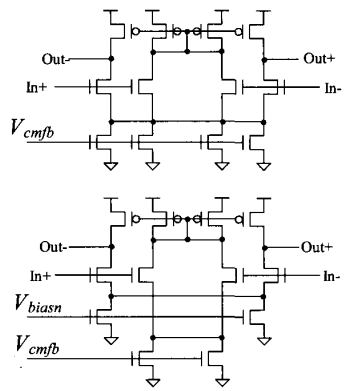


Figure 26.61 Circuits for Problem 26.4.

26.5 Verify, using simulations, that the circuit in Fig. 26.9 does indeed amplify the difference between V_{biasp} and the average on the + inputs of the amplifier.

IPR2023-00698
Theta EX2019

Comment on the operation of the circuit, making sure it is clear that the limitations, uses, and operation of the amplifier are understood.

26.6 Simulate the operation of the CMFB circuit in Fig. 26.16.

26.7 Is *VDD* divided evenly amongst the drain-source voltages of the MOSFETs in Fig. 26.19? Suggest a method to better divide *VDD* amongst the MOSFETs used in the diff-amp.

26.8 Suggest a simple method to speed up the step response of the op-amp in Fig. 26.22 in the circuit of Fig. 26.24. Verify the validity of your suggestion using SPICE simulations.

26.9 Suggest an alternative method to the one seen in Fig. 26.32 for controlling the output buffer's current. Using the modification, regenerate the results seen in Fig. 26.36.

26.10 Is the CMFB loop stable in the op-amp of Fig. 26.40? Use the large-signal test circuit seen in Fig. 26.24 (at a slower frequency) and simulations to look at the stability of this loop. Suggest, and verify with simulations, methods to improve the stability of the CMFB loop.

26.11 Repeat Problem 26.10 for the op-amp in Fig. 26.43.

26.12 As discussed at the end of the chapter and in Fig. 26.59 and the associated discussion, the input common-mode voltage which drops below V_{CM}, can shut off the op-amp's input diff-amp and cause problems. The diff-amp's NMOS devices have a nominal V_{GS} of 400 mV, leaving only 100 mV across the diff-amp's tail current. Show, using the op-amp in Fig. 26.39 in the configuration seen in Fig. 26.24 with an input common-mode voltage less than 500 mV (the input pulse waveforms average to a voltage less than V_{CM}), the resulting problems. Show that increasing the widths of the NMOS diff-pair (and the mirrored NMOS in the bias circuit with gates tied to V_{CM}) from 10 to 30 helps to increase the operating range (four MOSFET widths are increased from 10 to 30). What happens if only the diff-pair widths (two MOSFETs) are increased?

26.13 Repeat Problem 26.12 for the op-amp in Fig. 26.40. What happens if the input common-mode voltage becomes greater than 500 mV?

26.14 Repeat Problem 26.12 for the op-amp in Fig. 26.43. What happens if the input common-mode voltage becomes greater than 500 mV? Why does the op-amp in Fig. 26.40 perform better with variations in the input common-mode voltage than the op-amp in Fig. 26.43?

26.15 Design and simulate the operation of an op-amp using gain-enhancement (Fig. 26.42) and with an open-loop DC gain greater than 2,000, based on the topologies seen in Figs. 26.40 or 26.43. Simulate the operation of your design and generate outputs like those seen in Fig. 26.44.

26.16 Figure 26.62 shows an op-amp based on the topology seen in Fig. 26.40 but biased for lower *VDD* operation. Select the size of the added gate-drain connected MOSFET to allow for proper operation. Simulate the operation of the design showing the DC gain and large signal step responses (as in Fig. 26.34 and 26.35).

IPR2023-00698
Theta EX2019

26.17 Repeat Problem 26.16 for the op-amp in Fig. 26.43.

26.18 Using the diff-amp in Fig. 26.63 in the configuration seen in Fig. 26.54 (with SC CMFB) and the output buffer in Fig. 26.56 (again with SC CMFB), regenerate the waveforms seen in Fig. 26.60.

26.19 A switched-capacitor integrator is an example of a circuit that uses an op-amp that can't have the outputs of its diff-amp or output buffer shorted when ϕ_1 goes high. Using the diff-amp in Fig. 26.63 (with SC CMFB) and the output buffer in Fig. 26.56 without the switches (again with SC CMFB), demonstrate the operation of a SC integrator.

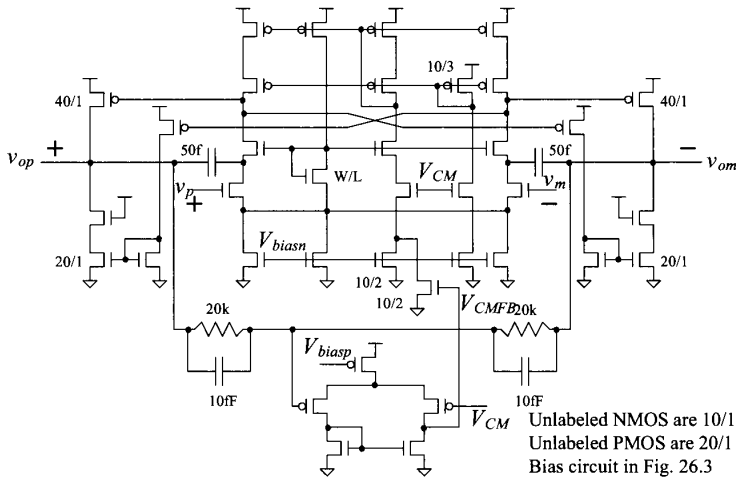


Figure 26.62 An op-amp with better biasing (but more power) for lower VDD operation.

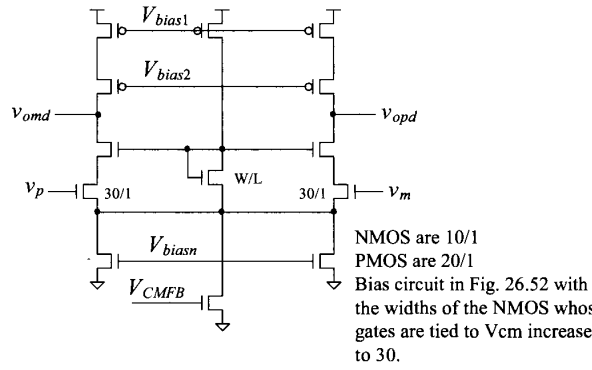


Figure 26.63 Modifying the diff-amp seen in Fig. 26.53 for wider swing operation.

IPR2023-00698
Theta EX2019

Chapter

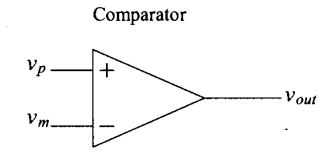
27

Nonlinear Analog Circuits

In this book we've studied digital circuits (two discrete amplitude levels) and linear analog circuits (circuits whose input signals are linearly related to the circuits' output signals). In this chapter, we discuss several circuits that are not purely analog or digital. We term these circuits *nonlinear analog circuits* (the inputs are not linearly related to the outputs). In particular, we discuss voltage comparator analysis and design, adaptive biasing, and analog multiplier design.

27.1 Basic CMOS Comparator Design

The schematic symbol and basic operation of a voltage comparator are shown in Fig. 27.1. The comparator can be thought of as a decision-making circuit. If the +, v_p, input of the comparator is at a greater potential than the –, v_m, input, the output of the comparator is a logic 1, whereas if the + input is at a potential less than the – input, the output of the comparator is at a logic 0. Although the basic op-amps developed in Ch. 24 can be used as a voltage comparator, in some less demanding low-frequency or speed applications, we do not consider the op-amp as a comparator. Instead, in this chapter, we discuss practical comparator design and analysis where propagation delay and sensitivity are important.



$v_p > v_m$ then $v_{out} = VDD$ = logic 1

$v_p < v_m$ then $v_{out} = 0$ = logic 0

Figure 27.1 Comparator operation.

IPR2023-00698
Theta EX2019

A block diagram of a high-performance comparator is shown in Fig. 27.2. The comparator consists of three stages: the input preamplifier, a positive feedback or decision stage, and an output buffer. The pre-amp stage (or stages) amplifies the input signal to improve the comparator sensitivity (i.e., increases the minimum input signal with which the comparator can make a decision) and isolates the input of the comparator from switching noise (often called kickback noise) coming from the positive feedback stage (*this is important*, see the discussion in Sec. 16.2.1). The positive feedback stage determines which of the input signals is larger. The output buffer amplifies this information and outputs a digital signal. Designing a comparator can begin with considering input common-mode range, power dissipation, propagation delay, and comparator gain.

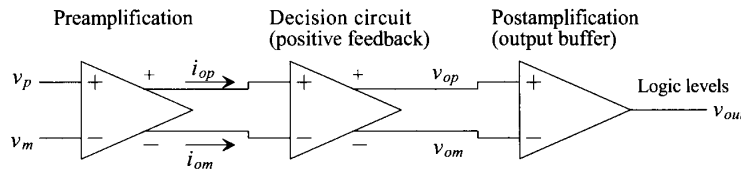


Figure 27.2 Block diagram of a voltage comparator.

Preamplification

For the preamplification (pre-amp) stage, we chose the circuit of Fig. 27.3. For this first section we'll use the long-channel CMOS process to illustrate the design procedures (since long-channel MOSFETs follow the square-law equations). This circuit is a differential amplifier with active loads. The sizes of M1 and M2 are set by considering the diff-amp transconductance, g_m, and the input capacitance. The transconductance sets the gain of the stage, while the input capacitance of the comparator is determined by the sizes of M1 and M2. Notice that there are no high-impedance nodes in this circuit, other than the input and output nodes. This is important to ensure high speed. Using the sizes given in the schematic, we can relate the input voltages to the output currents (noting i_{op} and i_{om} are the small-signal AC currents in the circuit) by

$$i_{op} = \frac{g_m}{2}(v_p - v_m) + \frac{I_{SS}}{2} = I_{SS} - i_{om} \tag{27.1}$$

Noting that if $v_p > v_m$, then i_{op} is positive i_{om} is negative ($i_{op} = -i_{om}$).

Decision Circuit

The decision circuit is the heart of the comparator and should be capable of discriminating mV-level signals. We should also be able to design the circuit with some hysteresis (see Ch. 18) for use in rejecting noise on a signal. The circuit that we use in the comparator under development is shown in Fig. 27.4. The circuit uses positive feedback from the cross-gate connection of M6 and M7 to increase the gain of the decision element.

IPR2023-00698
Theta EX2019

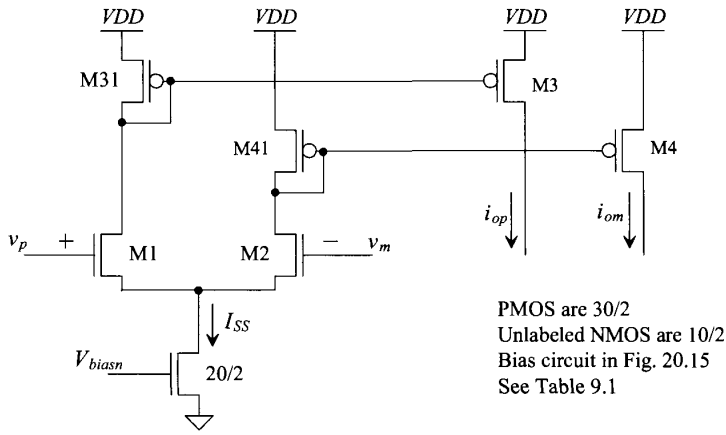


Figure 27.3 Preamplification stage of comparator.

Let's begin by assuming that i_{op} is much larger than i_{om} so that M5 and M7 are on and M6 and M8 are off. We also assume that $\beta_5 = \beta_8 = \beta_A$ and $\beta_6 = \beta_7 = \beta_B$. Under these circumstances, v_{om} is approximately 0 V and v_{op} is

$$v_{op} = \sqrt{\frac{2i_{op}}{\beta_A}} + V_{THN} \tag{27.2}$$

If we start to increase i_{om} and decrease i_{op}, switching starts to take place when the gate-source voltage of M8 is equal to V_{THN}. As we increase M8's V_{GS} beyond V_{THN} (by further increasing i_{om} with the corresponding decrease in i_{op}), M6 starts to take current away from M5. This decreases the drain-source voltage of M5/M6 and thus turns M7 off.

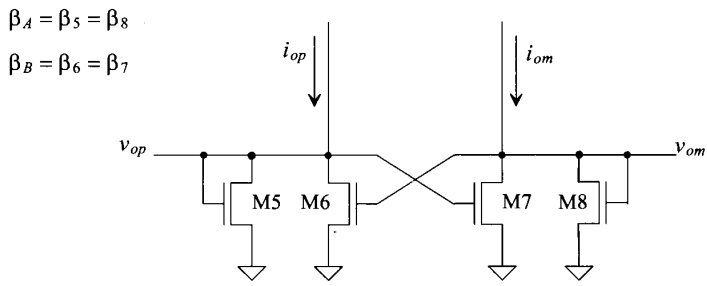


Figure 27.4 Positive feedback decision circuit.

IPR2023-00698
Theta EX2019

When M8 is just about to turn on (M8's V_{GS} is approaching V_{THN} but the drain currents of M8 and M6 are still zero), the current flowing in M7 is

$$i_{om} = \frac{\beta_B}{2}(v_{op} - V_{THN})^2 \tag{27.3}$$

and the current flowing in M5 is

$$i_{op} = \frac{\beta_A}{2}(v_{op} - V_{THN})^2 \tag{27.4}$$

Noting that the current in M7 (at the switching point) mirrors the current in M5, we can write

$$i_{op} = \frac{\beta_A}{\beta_B} \cdot i_{om} \tag{27.5}$$

If $\beta_A = \beta_B$, then switching takes place when the currents, i_{op} and i_{om}, are equal. Unequal βs cause the comparator to exhibit hysteresis. Relating these equations to Eq. (27.1) yields the switching point voltages (review Ch. 18), or

$$V_{SPH} = v_p - v_m = \frac{I_{SS}}{g_m} \cdot \frac{\frac{\beta_B}{\beta_A} - 1}{\frac{\beta_B}{\beta_A} + 1} \text{ for } \beta_B \geq \beta_A \tag{27.6}$$

and

$$V_{SPL} = -V_{SPH} \tag{27.7}$$

Consider the following example.

Example 27.1

For the circuit shown in Fig. 27.5, estimate and simulate the switching point voltages for two designs: (1) $W_5 = W_6 = W_7 = W_8 = 10$ with $L = 1$ and (2) $W_5 = W_8 = 10$ and $W_6 = W_7 = 12$ with $L = 1$.

For the first case (using the data in Table 9.1)

$$\beta_A = \beta_B = 120\,\frac{\mu A}{V^2} \cdot \frac{10}{1} = 1.2\ mA/V^2$$

so that, as seen in Eq. (27.6) $V_{SPH} = V_{SPL} = 0$. In other words, the comparator does not exhibit hysteresis. Simulation results are shown in Fig. 27.6. The input v_m is set to 2.5 V, while the v_p input is swept from 2.48 to 2.52 V. Note that the amplitudes of v_{op} and v_{om} are limited.

For the second case

$$\beta_A = 120\,\frac{\mu A}{V^2} \cdot \frac{10}{1} \text{ and } \beta_B = 120\,\frac{\mu A}{V^2} \cdot \frac{12}{1}$$

or $\beta_B = 1.2\beta_A$. Using Eqs. (27.6) and (27.7) with $I_{SS} = 40$ μA and $g_m = 150$ μA/V (again, see Table 9.1), we get

$$V_{SPH} = -V_{SPL} = \frac{I_{SS}}{g_m} \cdot \frac{\frac{\beta_B}{\beta_A} - 1}{\frac{\beta_B}{\beta_A} + 1} = \frac{40}{150} \cdot \frac{1.2 - 1}{1.2 + 1} = 24\ mV$$

IPR2023-00698
Theta EX2019

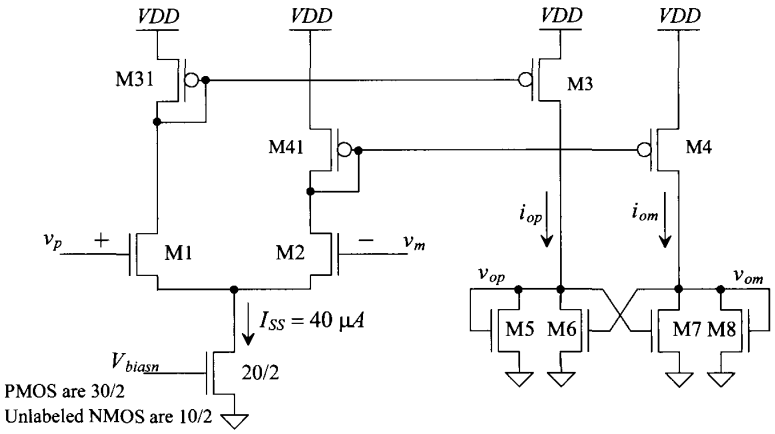
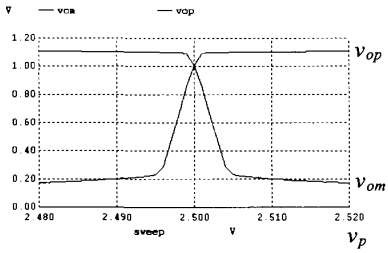
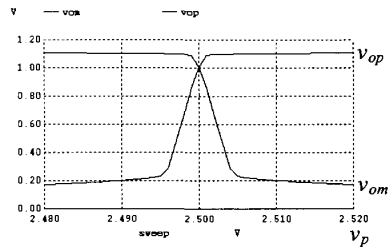


Figure 27.5 Schematic of the pre-amp and decision circuit used in Ex. 27.1.



(a) Minus input held at 2.5 V while the positive input is swept from 2.4 to 2.6 V.



(b) Minus input held at 2.5 V while the positive input is swept from 2.52 to 2.4.

Figure 27.6 The outputs of the decision circuit in Fig. 27.5 without hysteresis.

The simulation results are shown in Fig. 27.7. Figure 27.7a shows a sweep of v_p from 2.4 to 2.6 V, with v_m held at 2.5 V. Since V_{SPH} is 24 mV, the decision circuit switches states when v_p is 24 mV above v_m (i.e., when v_p = 2.524 V). The case of v_p being swept from 2.6 to 2.4 V is shown in part (b) of the figure. Switching occurs in this situation when v_p is approximately 2.476 V, or 24 mV, less than 2.5 V. ■

Output Buffer

The final component in our comparator design is the output buffer or post-amplifier. The main purpose of the output buffer is to convert the output of the decision circuit into a logic signal (i.e., 0 or *VDD*). The output buffer should accept a differential input signal

IPR2023-00698
Theta EX2019

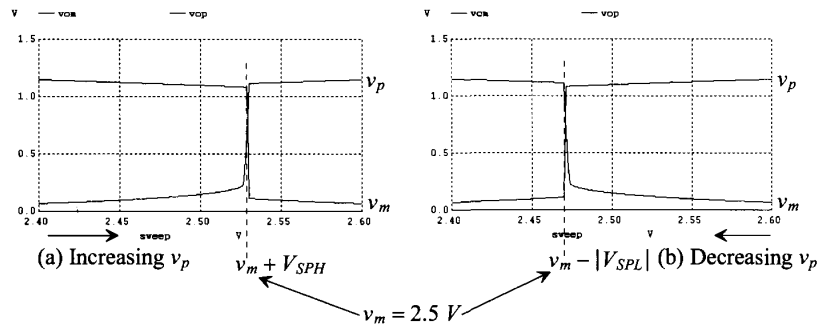


Figure 27.7 Simulating the pre-amp and decision circuit with hysteresis.

and not have slew-rate limitations. We might try to use a NAND SR latch as seen back in Fig. 16.35. However, with process shifts we might run into the situation where the switching points of the gates aren't centered in the middle of the limited swing of the decision circuit's outputs. For a general comparator design, we'll use a diff-amp to help regenerate the digital output signals. For a simple design for the output buffer, we can use the self-biased diff-amp seen back in Fig. 18.17. To move the decision circuit's output swing into the common-mode range of the diff-amp, we can add a gate-drain connected device as seen in Fig. 27.8. An inverter was added on the output of the amplifier as an additional gain stage and to isolate any load capacitance from the self-biasing differential amplifier. This comparator works very well. However, the current in the self-biased stage can be (relatively) huge. This can be a problem if power is a concern (as it is in the Flash ADC discussed in Ch. 29).

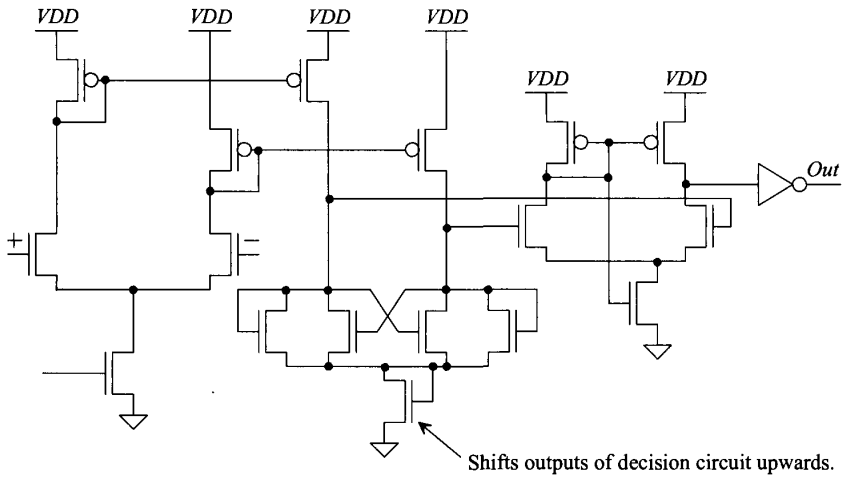


Figure 27.8 Using a self-biased diff-amp for the output buffer.

IPR2023-00698
Theta EX2019

Figure 27.9 shows a comparator design with both a PMOS and NMOS diff-amp feeding the decision circuit (for input common-mode range beyond the power supply rails). The output buffer is a PMOS diff-amp driving an inverter. While this op-amp won't be as fast as the op-amp seen in Fig. 27.8, it is a good general purpose design. The current drawn from *VDD* will be considerably less than the design in Fig. 27.8. Note that the output currents of the PMOS diff-amp, on the input of the comparator, are added to the NMOS (current) outputs to ensure that current is always sourced to the decision circuit. We might be tempted to connect the PMOS diff-amp directly to the decision circuit. However, this would result in a larger minimum input common-mode voltage.

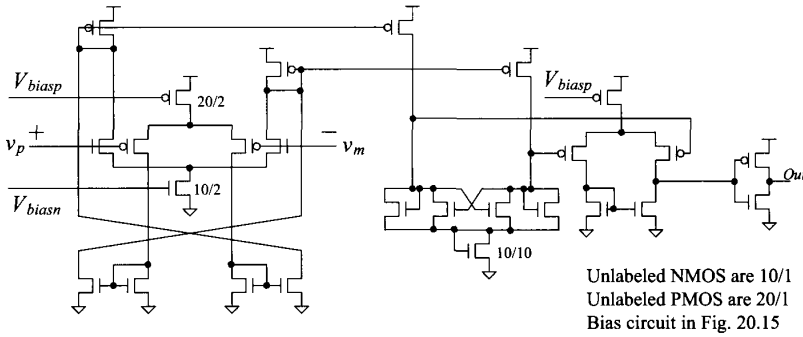


Figure 27.9 General-purpose comparator with rail-to-rail input common-mode range.

27.1.1 Characterizing the Comparator

Comparator DC Performance

Figure 27.10 shows the DC simulated performance of the comparator in Fig. 27.9. The positive comparator input is swept from 0 to *VDD* (= 5 V) while the negative input is stepped for each simulation from 0 to 5 in 500 mV increments (to show wide input common-mode range). Also seen in this figure is the comparator's current draw from *VDD*. Included in this current is the current supplied to the bias circuit.

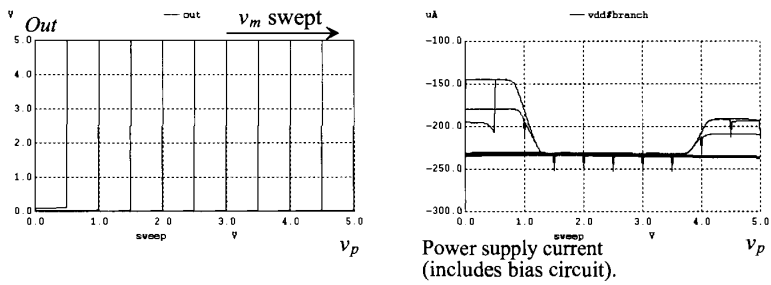


Figure 27.10 DC performance of the comparator in Fig. 27.9

IPR2023-00698
Theta EX2019

If we take the derivative of these transfer curves, the gain of the comparator and thus the smallest difference that can be discriminated between v_p and v_m becomes known. We have to be careful here with the step size used in the DC simulation. If, for example, we were to use a step size of 1 mV, then the maximum gain we get from the simulation is 1,000. Figure 27.11 shows the expanded view of the comparator with the minus input held at 2.5 V while the positive input is swept from 2.499 to 2.501 V. The offset seen in the figure is a systematic offset (as discussed earlier). The gain of the comparator is, roughly, 175,000. Notice that the comparator was designed without hysteresis. However, in a practical comparator mismatch in the decision circuit results in hysteresis (and so if hysteresis isn't desired, the layout of the decision circuit is critical and may require common-centroid techniques).

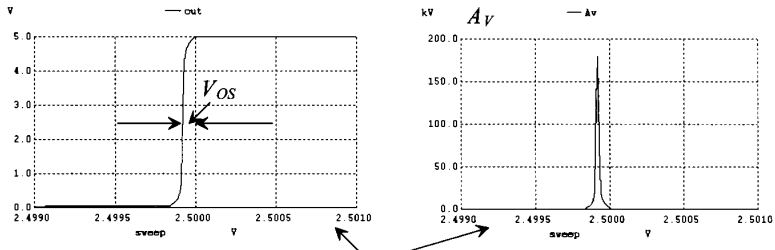


Figure 27.11 The gain of the comparator in Fig. 27.9

Transient Response

The transient response of a comparator can be significantly more difficult to characterize than the DC characteristics. Let's begin by considering v_m = 2.5 V DC, with the v_p input to the comparator a 10 ns wide pulse and an amplitude varying from 2.45 to 2.55 V. This is termed a narrow pulse with a 50 mV overdrive; we are driving the + input of the comparator 50 mV over the negative input. The simulation results for these inputs are shown in Fig. 27.12. If the pulse amplitude or width is reduced much beyond this, the comparator does not make a full transition. Notice how we only show simulation data after 120 ns. This was done to ensure the bias circuit had time to start-up and stabilize before changing the inputs of the comparator.

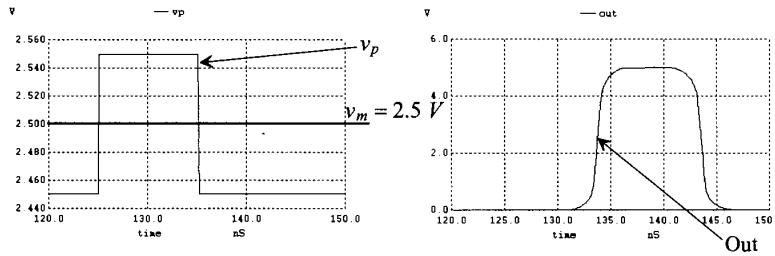


Figure 27.12 Transient response of the comparator in Fig. 27.9.

IPR2023-00698
Theta EX2019

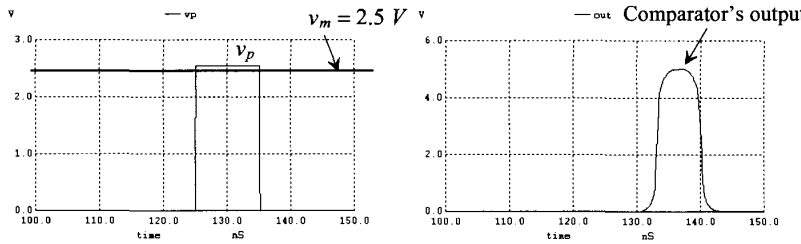


Figure 27.13 Transient response of the comparator in Fig. 27.9 with one side of the pre-amp initially shut off.

Although these results are interesting, they are not practically useful in some situations. In few cases will the comparator discriminate between signals similar to what is seen in Fig. 27.12. A better indication of comparator performance is to apply a signal that starts at a voltage where one side of the differential amplifier is cut off and to finish at a voltage slightly larger than the reference voltage (2.5 V in the above example), Fig. 27.13. In applying a 0 to 2.6 V pulse to the v_p input, the left side of the diff-amp is initially off and then on. The internal comparator nodes must be charged over a wider voltage range with a 0–2.6 V input when compared to the 2.45–2.55 V example of Fig. 27.12. In order to keep one side of the pre-amp from turning off, the circuit of Fig. 27.14 can be used. The added MOSFET ensures that the voltage between the drains of M31 and M41 don't deviate too much. This configuration is sometimes referred to as a clamped input stage. Note that the size of the added MOSFET should be as small as possible to avoid unneeded loading in the diff-amp (which slows down the comparator's performance).

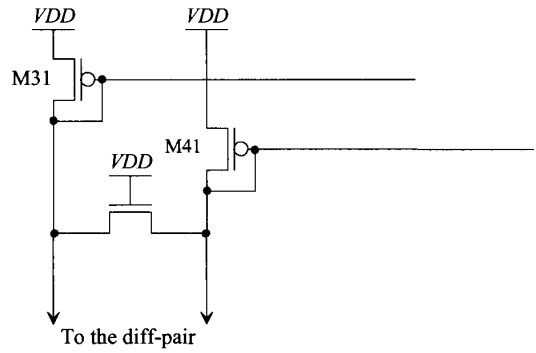


Figure 27.14 Adding a balancing resistor to the input stage of a comparator. See Fig. 27.5

IPR2023-00698
Theta EX2019

Propagation Delay

Ideally, the propagation delay (the time difference between the input, v_p, crossing the reference voltage, v_m, and the output changing logic states) is zero. For the comparator of Fig. 27.9 with simulation results shown in Fig. 27.12, the delay of the comparator is approximately 9 ns (or less with an overdrive). This should be compared with the delay of an op-amp used as a comparator, which may be several hundred nanoseconds (because of the compensation capacitor). Another interesting fact about comparator design is that the delay of a comparator can be reduced by cascading gain stages. In other words, the delay of a single high-gain stage is in general longer than the delay of several low-gain stages. Improvements in comparator sensitivity and speed can be directly related to improvements in the preamplifier. A derivation (similar to that given in Ch. 11 for minimum delay through an inverter string driving a load capacitance, C_L) of the number of stages, N, needed in a preamplifier to reach minimum delay for a given load capacitance and input capacitance results in

$$N = \ln \frac{C_L}{C_{in}} \tag{27.8}$$

In the derivation of this equation, it was assumed that the driving resistance of any stage is $1/g_{mn}$ (the transconductance of the nth differential amplifier stage). In practice, Eq. (27.8) is of little use because one side of the differential amplifier can be off. For this reason (and others such as size and power draw), comparators with more than two or three pre-amp stages are not common.

Minimum Input Slew Rate

The last characteristic we discuss is the minimum input slew rate of a comparator. If the input signals to the comparator vary at a slow rate (e.g., a sine wave generated from the AC line), the output of the comparator may very well oscillate resulting in an output with a metastable state. If a comparator is to be used with slowly varying signals, or in a noisy environment, the decision circuit should have hysteresis. The minimum input slew rate is difficult to simulate with SPICE because of the slow and fast varying signals present at the same instant of time in the circuit. This same situation (metastability) can also occur if the input overdrive is small. The result, for this case, however, is an increase in the delay time of the comparator.

27.1.2 Clocked Comparators

We developed "sense amplifiers" in Sec. 16.2.1. The sense amplifier can be used as a clocked comparator, that is a comparator whose outputs change on the rising (or falling) edge of a clock. Consider the clocked comparator in Fig. 27.15. This circuit is directly taken from Figs. 16.32 and 16.35. It works very well for signal differences that are tens of millivolts. When *clock* is low, the inputs to the NAND SR latch are pulled high. The outputs of the comparator don't change. When *clock* goes high, the two inputs are compared causing the output of the circuit to register which one is higher. As discussed in detail back in Ch. 16, the minimum input voltages for this comparator are above the threshold voltage of the NMOS devices. Further if the input signals get too large, MB1 and MB2 move deep into the triode region. The small differences in their channel resistances adversely affect the quality of the comparator's decisions. Let's modify this topology for better sensitivity, wider input signal swing, and better immunity to kickback.

IPR2023-00698
Theta EX2019

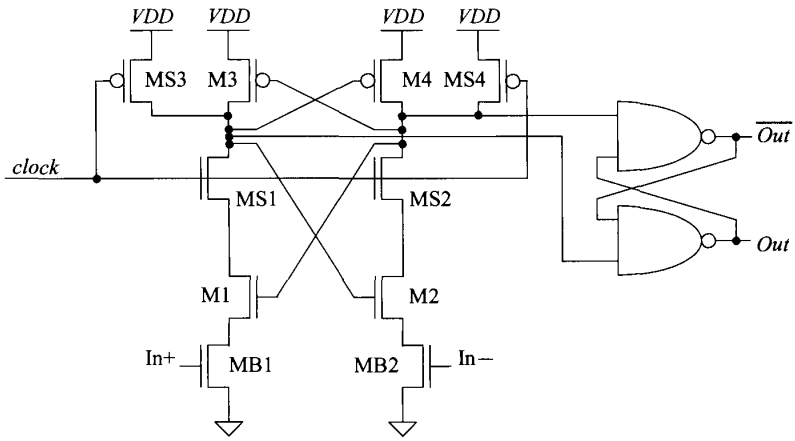


Figure 27.15 A clocked comparator based on Figs. 16.32 and 16.35.

Our modification is seen in Fig. 27.16. Instead of biasing the PMOS and NMOS diff-amps with a bias circuit, here, to be different, we simply use long-length MOSFETs. The bias currents aren't too critical in this application (like they are in an op-amp design). We can still use a bias circuit if controlling the bias current is, for some reason, important. We've removed the triode-operating MOSFETs MB1 and MB2 from the basic cross-coupled latch section. Now we are steering currents from one side of the latch to the other. The differences in the currents causes the latch to switch dependent on the input signals. Note that if either of the diff-amps isn't present in the comparator or if one is off because the comparator's input common-mode voltage is too high or too low, the comparator still works as desired. Only one diff-amp is needed to create the imbalance. Also note, again, that the SR latch is used to make the outputs of the circuit change on the rising edge of the clock signal. If the outputs of the comparator can go high, for a particular application, when the clock signal is low then the NAND gates can be removed.

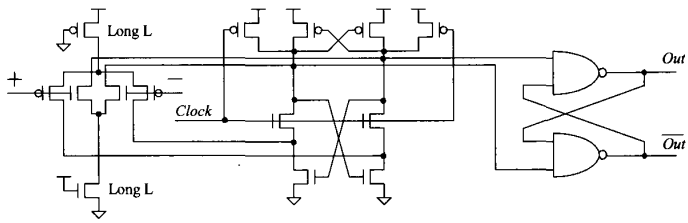


Figure 27.16 Wide-swing clocked comparator. Outputs change on the rising edge of the clock signal.

IPR2023-00698
Theta EX2019

27.1.3 Input Buffers Revisited

When we discussed input buffers back in Ch. 18 we used a self-biased design for the highest speed (see Fig. 18.17 and the associated discussions). As discussed in Ch. 18, it's desirable to have symmetrical propagation delays independent of input slew-rate, amplitudes, or direction (high-to-low or low-to-high). Towards the more ideal input buffer, consider using two NMOS self-biased buffers in parallel, as seen in Fig. 27.17. Here, to maintain good symmetry, we've split the current sources in half and used one side to generate a common-mode feedback signal to balance the outputs. To reduce the power dissipation in the buffer, the triode-operating MOSFETs can have their lengths increased. Further, for rail-to-rail input common-mode range, we can place the PMOS version of this buffer in parallel with the NMOS version seen in Fig. 27.17, as we did in Fig. 18.23. This input buffer can be very useful for low-skew, high-speed design.

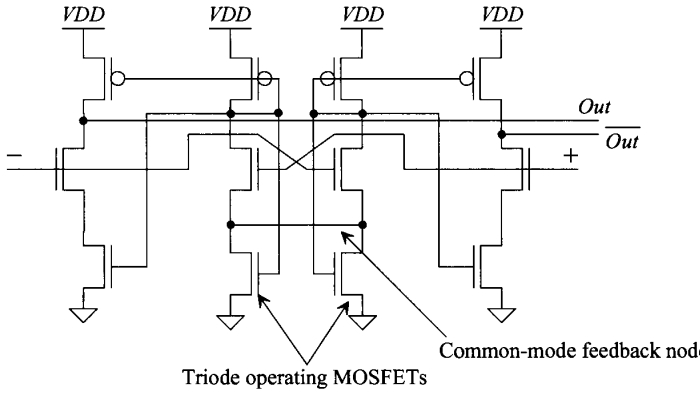


Figure 27.17 A fully-differential input buffer based on the topology discussed in Sec. 18.17.

27.2 Adaptive Biasing

Adaptive biasing can reduce power dissipation in an amplifier while at the same time increasing output current drive capability. Figure 27.18 can be used to help illustrate the idea. When v_{I1} and v_{I2} are equal, the current sources I_{SS1} and I_{SS2} are zero (an open). The diff-amp DC tail current is simply I_{SS}, the same as an ordinarily biased diff-amp. If v_{I1} becomes larger than v_{I2}, the current source I_{SS1} increases above zero, effectively increasing the diff-amp DC bias current. Similarly, if v_{I2} becomes larger than v_{I1}, the current source I_{SS2} increases above zero. The diff-amp output current is normally limited to I_{SS} when one side of the diff-amp shuts off. However, now that the maximum output current is limited to either $I_{SS} + I_{SS1}$ or $I_{SS} + I_{SS2}$. Power dissipation can be reduced using an adaptive bias, and slew-rate problems can be eliminated.

IPR2023-00698
Theta EX2019

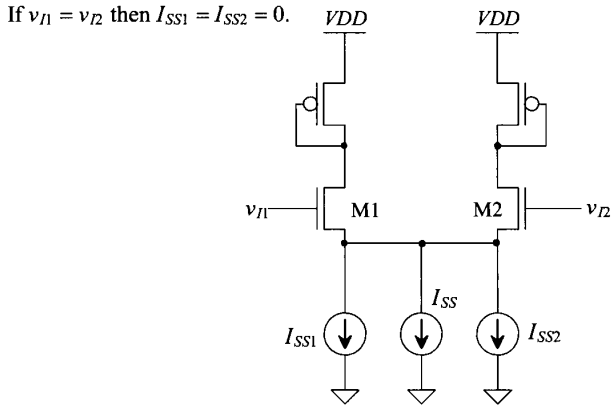


Figure 27.18 Adaptively biased diff-amp.

The current diff-amp of Fig. 27.19 can be used to implement the current source I_{SS1} or I_{SS2}. If the currents I_1 and I_2 are equal, then zero current flows in M3 and M4. Also, if I_2 is greater than I_1, zero current flows in M3 and M4. If I_1 is larger than I_2, the difference between these two currents $(I_1 - I_2)$ flows in M3. Since M4 is K times wider than M3, a current of $K(I_1 - I_2)$ flows in M4 (normally $K < 1$). Two of these diff-amps are needed to implement the adaptive biasing of the diff-amp of Fig. 27.18.

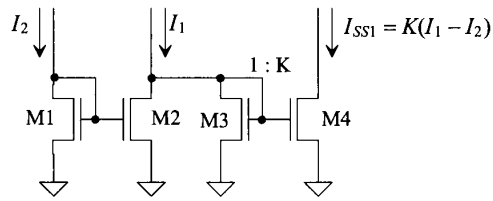


Figure 27.19 Current diff-amp used in adaptive biasing.

Figure 27.20 shows the implementation of adaptive biasing into the diff-amp of Fig. 27.18. P-channel MOSFETs are added adjacent to M3 and M4 to mirror the currents through M1 and M2 (I_1 and I_2). The maximum total current available through M1 occurs when M2 is off. Positive feedback exists through the loop M1, M3, M5–M7. Initially, when M2 shuts off, the current in M1 and M3 is I_{SS}. This is mirrored in M5 and M6, and thus I_{SS1} becomes $K \cdot I_{SS}$. At this particular instance in time, the tail current, which flows through M1, is now $I_{SS} + K \cdot I_{SS}$. However, provided the MOSFETs M1, M3–M7 remain in

IPR2023-00698
Theta EX2019

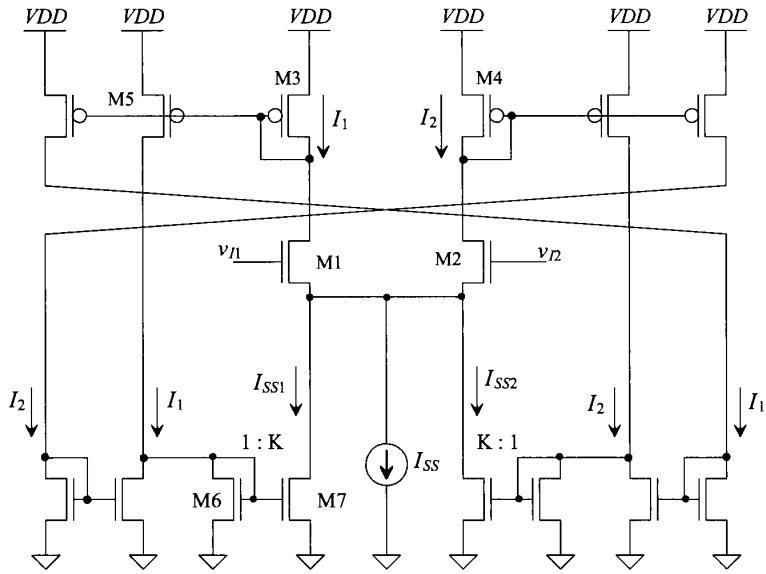


Figure 27.20 Adaptively biased diff-amp.

saturation, this current circles back around the positive feedback loop and increases by K. This continues, resulting in a final or total tail current of

$$I_{tot} = I_{SS} \cdot (1 + K + K^2 + K^3 + ...) \tag{27.9}$$

If $K < 1$, this geometric series can be written as

$$I_{tot} = \frac{I_{ss}}{1 - K} \tag{27.10}$$

Setting $K = 0$ (MOSFET M7 doesn't exist) results in no adaptive biasing and a total tail current of I_{SS}. Setting $K = 1/2$ (M7 half the size of M6) results in a total available tail current of $2 \cdot I_{SS}$. Because the tail current limits the slew rate, when the diff-amp is driving a capacitive load, making K equal to one eliminates slew-rate limitations while at the same time not increasing static power dissipation. In practice, shutting off one side of the diff-pair, M1/M2, is difficult since the adaptive biasing has the effect of lowering the source potentials of the diff-pair, keeping both MOSFETs on. Adaptive biasing can be used in a comparator where M1 or M2 can shut off. However, the static power dissipation will be large; therefore, there is no benefit over the comparators discussed earlier in the chapter. When applying adaptive biasing to an OTA design, the value of K should be unity or less. [Using a K of 1 or 2 can still result in a finite I_{tot} since the MOSFETs have a finite output resistance and the MOSFETs in the diff-pair will not shut off, as was assumed in the derivation of Eq. (27.10)].

IPR2023-00698
Theta EX2019

A final example of an adaptive voltage-follower amplifier is shown in Fig. 27.21. This amplifier can only source current to a load. If v_{in} and v_{out} are equal, the current that flows in M1 and M2 is $I_{SS} + I_{D6}$. If v_{in} is increased, the current in M1 and M3 increases. This causes the currents in M4–M6 to increase, effectively increasing the tail current of the diff-pair. The result is a large current available to drive the load. Note that M7 can be sized larger than the other MOSFETs to increase maximum output current.

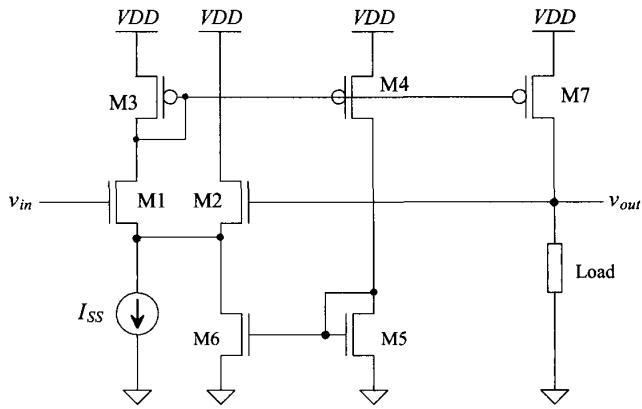


Figure 27.21 Adaptive voltage follower.

27.3 Analog Multipliers

Analog multipliers find extensive use in communication systems. Figure 27.22 shows the voltage characteristics of a four-quadrant multiplier. This multiplier is termed a four-quadrant multiplier because both inputs can be either positive or negative around a common-mode voltage, V_{CM}. The ideal output of the multiplier is related to the inputs by

$$v_{out} = K_m \cdot v_x v_y \tag{27.11}$$

where K_m is the multiplier gain with units of V^{-1}. In reality, imperfections exist in the multiplier gain, resulting in offsets and nonlinearities. The output of the multiplier can be written as

$$v_{out} = K_m(v_x + V_{OSx})(v_y + V_{OSy}) + V_{OSout} + v_x^n + v_y^m \tag{27.12}$$

where V_{OSx}, V_{OSy}, and V_{OSout} are the offset voltages associated with the x-, y-inputs, and the output, respectively. The terms v_x^n and v_y^m represent nonlinearities in the multiplier. Normally, these nonlinearities are specified in terms of the total harmonic-distortion or by specifying the maximum deviation in percentages between a straight line and the actual characteristic curves shown in Fig. 27.22 over some range of input voltages. Although many different techniques exist for implementing analog multipliers in CMOS, we concentrate on a technique useful in high- and low-frequency multiplication.

IPR2023-00698
Theta EX2019

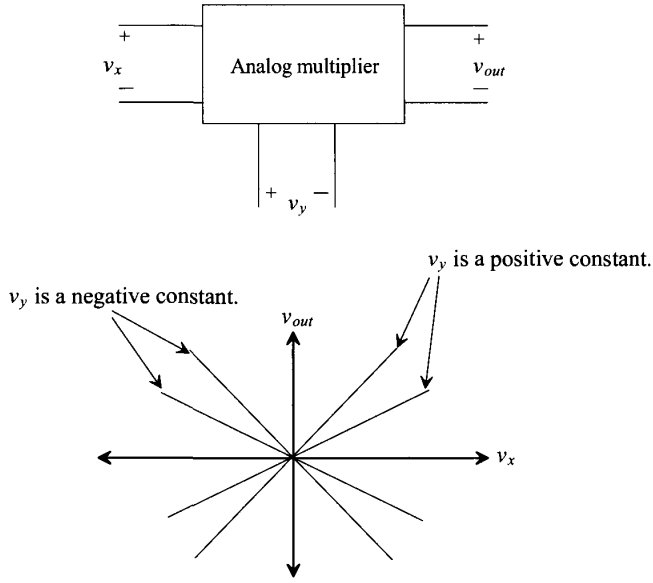


Figure 27.22 Operation of a four-quadrant analog multiplier.

27.3.1 The Multiplying Quad

A CMOS multiplier employing a multiplying quad (M1–M4) is shown in Fig. 27.23. The multiplying quad operates in the triode region, and thus MOSFETs M1–M4 can be thought of as resistors. For the moment we will not consider the biasing of the quad. The negative output voltage of the multiplier is given by

$$v_{om} = -R \cdot (i_{D1} + i_{D2}) \tag{27.13}$$

while the positive output voltage is

$$v_{op} = -R \cdot (i_{D3} + i_{D4}) \tag{27.14}$$

The output voltage of the multiplier is

$$v_{out} = v_{op} - v_{om} = R \cdot (i_{D1} + i_{D2} - i_{D3} - i_{D4}) \tag{27.15}$$

A simplified schematic of the multiplying quad with biasing is shown in Fig. 27.24. The op-amp inputs are at an AC virtual ground and at a DC voltage of V_{CM} (the op-amp output common-mode voltage). In order to minimize the DC input current on the x-axis inputs, the common-mode DC voltage on this input is set to V_{CM}. The DC biasing voltage on the y-input is set to a value large enough to keep the quad in triode. The input signals have been broken into two parts (e.g., $v_x/2$ and $-v_x/2$) to maintain generality. In practice, the minus inputs can be connected directly to the bias voltages at the cost of

IPR2023-00698
Theta EX2019

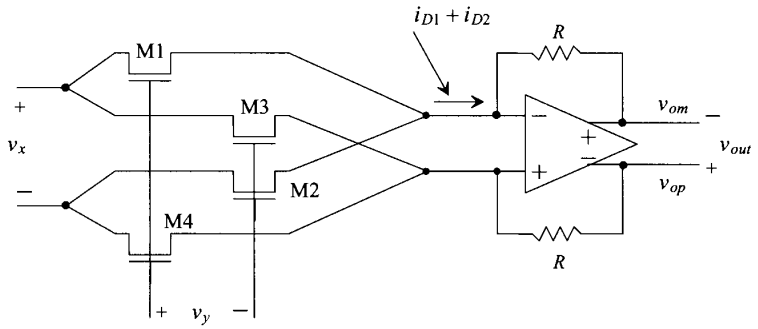


Figure 27.23 CMOS analog multiplier.

large-signal linearity. (The input is not truly differential in this situation.) However, as discussed earlier, a fully differential system has much better coupled noise immunity.

Using Eq. (9.12) and noticing that the DC gate-source voltage of all MOSFETs is the same, the drain currents can be written as

$$i_{D1} = \beta_1 \left[\left(V_{GS} + \frac{v_y}{2} - V_{THN1} \right) \left(\frac{v_x}{2} \right) - \frac{1}{2} \left(\frac{v_x}{2} \right)^2 \right] \qquad (27.16)$$

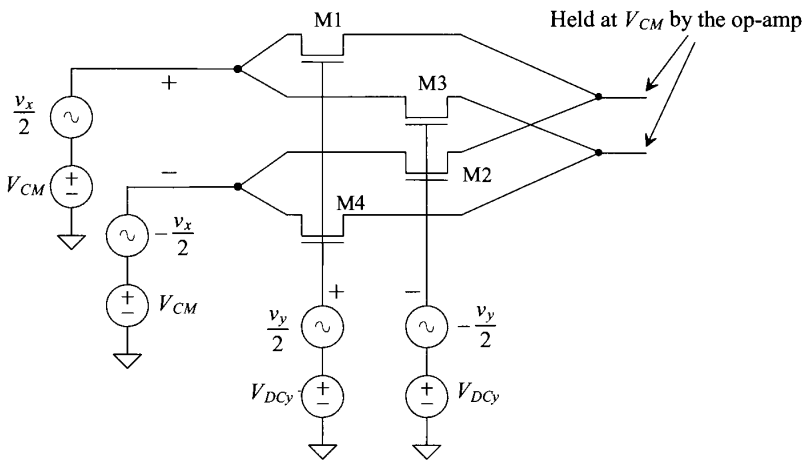


Figure 27.24 Biasing of the multiplying quad.

IPR2023-00698
Theta EX2019

$$i_{D2} = \beta_2\left[\left(V_{GS} - \frac{v_y}{2} - V_{THN2}\right)\left(-\frac{v_x}{2}\right) - \frac{1}{2}\left(-\frac{v_x}{2}\right)^2\right] \quad (27.17)$$

$$i_{D3} = \beta_3\left[\left(V_{GS} - \frac{v_y}{2} - V_{THN3}\right)\left(\frac{v_x}{2}\right) - \frac{1}{2}\left(\frac{v_x}{2}\right)^2\right] \quad (27.18)$$

$$i_{D4} = \beta_4\left[\left(V_{GS} + \frac{v_y}{2} - V_{THN4}\right)\left(-\frac{v_x}{2}\right) - \frac{1}{2}\left(-\frac{v_x}{2}\right)^2\right] \quad (27.19)$$

We can design so that $\beta = \beta_1 = \beta_2 = \beta_3 = \beta_4$. We can use Eq. (27.15) together with Eqs. (27.16)–(27.19) to rewrite the output voltage of the multiplier as

$$v_{out} = R\beta \cdot \left(\frac{v_x}{2}\right)\left[\frac{v_y}{2} - V_{THN1} + \frac{v_y}{2} + V_{THN2} + \frac{v_y}{2} + V_{THN3} + \frac{v_y}{2} - V_{THN4}\right] \quad (27.20)$$

We can see that if $V_{THN1} = (V_{THN2}$ or $V_{THN3})$ and $V_{THN4} = (V_{THN3}$ or $V_{THN2})$, this equation can be rewritten as

$$v_{out} = R\beta \cdot v_x v_y \quad (27.21)$$

The source of a MOSFET (the terminal we label "source" depends on which way current flows in the MOSFET) in the multiplying quad is connected either to the op-amp or to the x inputs. When the sources of the MOSFETs are connected to the op-amp, all of the MOSFETs in the multiplying quad have the same threshold voltage. (Since the source of each MOSFET is tied to the same potential, the body effect changes each MOSFET's threshold voltage by the same amount.) If the positive x-input is sinking a current, then the sources of M1 and M3 are the "+" x-input and thus $V_{THN1} = V_{THN3}$. In any case, the threshold voltages of the MOSFETs cancel and Eq. (27.21) holds. Comparing Eqs. (27.21) and (27.11) results in defining the gain of this multiplier as

$$K_m = R \cdot \beta \quad (27.22)$$

Simulating the Operation of the Multiplier

Simulating the performance and understanding the analog multiplier operation is an important step in the design process. The design of a multiplier consists of designing the op-amp, selecting the sizes of the multiplying quad, and designing the biasing network. Because we covered the design of differential input/output op-amps in the last chapter, it is not covered here. In order to simulate the performance of a multiplier in SPICE without including the limitations of the op-amp, the simple model shown in Fig. 27.25 is used. The sum of the multiplying factors associated with the voltage-controlled voltage sources, E1 and E2, is the open-loop gain of the op-amp. A typical SPICE statement for these VCVS (voltage-controlled voltage source) where the op-amp open loop gain is 20,000 is

```
E1    Voplus 8 4 3   1E4
E2    8 Vominus 4 3  1E4
```

where the nodes correspond to those labeled in Fig. 27.25.

The next problem we encounter in simulating the operation of the multiplier is implementing the differential voltages (e.g., $\pm\ v_x/2$), in addition to the DC biasing voltages. The setup shown in Fig. 27.26 is used to implement the biasing and the differential voltage sources. The op-amp common-mode output voltage, V_{CM}, and the x

IPR2023-00698
Theta EX2019

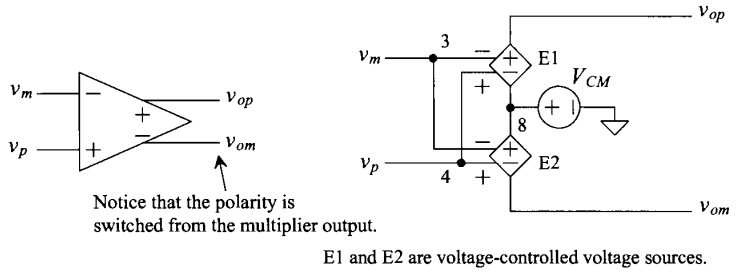


Figure 27.25 SPICE modeling a differential input/output op-amp with common-mode voltage.

input voltage are set to 1.5 V. The lower this voltage, the easier it is to bias the multiplying quad into the triode region. On the other hand, a reduction in the value of V_{CM} limits the op-amp output voltage swing and thus the multiplier output range. The size of the multiplying quad was set to 10/2. The larger the *W/L* ratio of the MOSFETs used in this quad, the easier it is to keep the quad in the triode region. On the other hand, using a large *W/L* increases the required input current. The channel length can be increased to ensure the device operates as a long-channel device and follows Eq. (9.12). Since the quad is part of the feedback around the op-amp, long-channel devices do not affect the

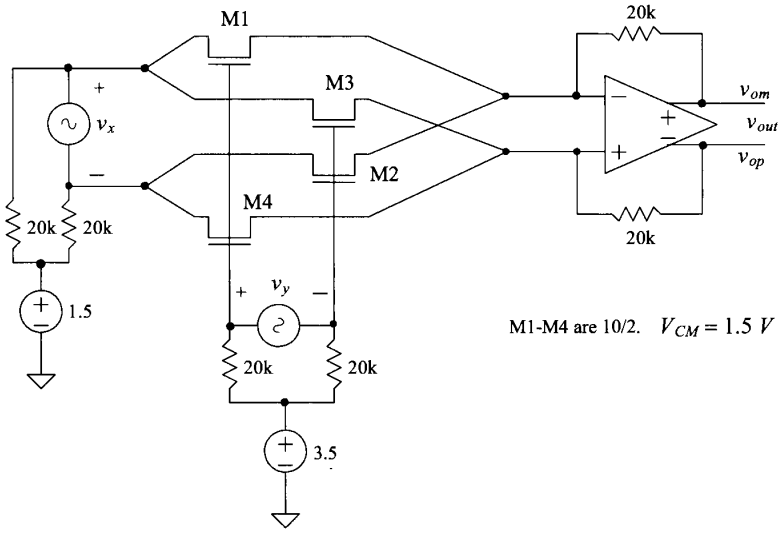


Figure 27.26 SPICE simulation schematic.

IPR2023-00698
Theta EX2019

speed. The DC voltage at the y-inputs was set to 3.5 V as a compromise between keeping the multiplying quad in triode and the y-input voltage range. The gain of the multiplier in Fig. 27.26 is, from Eq. (27.22) and Table 6.2,

$$K_m = 20k \cdot 120 \frac{\mu A}{V^2} \cdot \frac{10}{2} = 12\ V^{-1}$$

A DC sweep showing the operation of the multiplier is shown in Fig. 27.27. The x-input, v_x, was swept from –1 to +1, while at the same time the y-input was stepped from –1 to 1 V in 0.5 V increments. Keeping in mind that the output of the multiplier is $v_{op} - v_{om}$, we can understand the data presented in Fig. 27.27 by considering points A and B. At point A, the y-input is 1 V while the x-input is 0.25 V. The output voltage of the multiplier is the product of the multiplier gain and these two voltages (i.e., 12·1·0.25 = 3 V). The output voltage at point B is 12·0.5·(–0.3) = –1.8 V. Note that this figure was generated using an almost ideal op-amp. The characteristics do not show the limitations of the op-amp. In particular, the limited output swing.

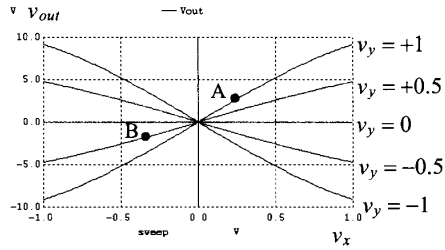


Figure 27.27 DC characteristics of the multiplier of Fig. 26.26.

27.3.2 Multiplier Design Using Squaring Circuits

An analog multiplier can be designed based on the difference between the sum of two voltages squared and the difference of two voltages squared, or

$$V_o = (V_1 + V_2)^2 - (V_1 - V_2)^2 = 4V_1V_2 \qquad (27.23)$$

The basic sum-squaring and difference-squaring circuits are shown in Fig. 27.28. MOSFETs M1 and M4 are source-followers, while MOSFETs M2 and M4 are called squaring MOSFETs. This circuit is designed so that $\beta_1 = \beta_4 = \beta_{14}$, $\beta_2 = \beta_3 = \beta_{23}$, and $\beta_{14} >> \beta_{23}$. This makes almost all of the DC bias currents, I_{S12} and I_{S34}, flow in the MOSFETs M1 and M4, respectively. The squaring current, I_{SQ}, assuming that zero current flows through the resistor when both inputs are zero volts (or whatever the common-mode voltage when a single supply is used), is given by

$$I_{SQ(a)} = \frac{\beta_{23}}{4}(V_1 + V_2)^2 \qquad (27.24)$$

Similarly, the squaring current in the difference-square circuit of Fig. 27.28b is given by

$$I_{SQ(b)} = \frac{\beta_{23}}{4}(V_1 - V_2)^2 \qquad (27.25)$$

IPR2023-00698
Theta EX2019

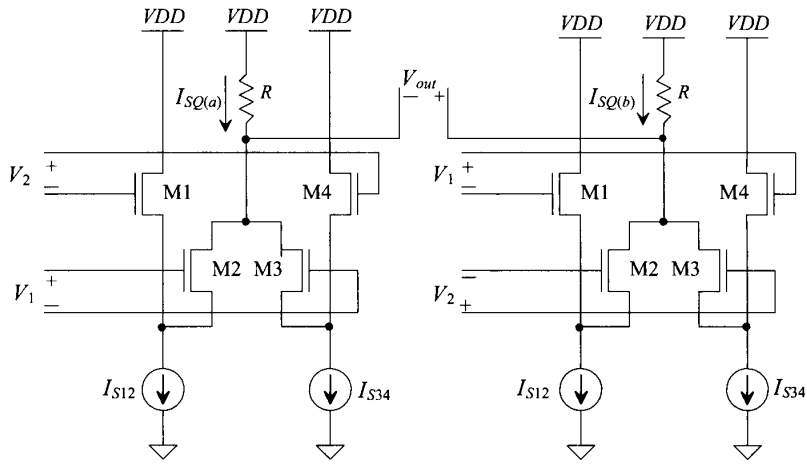


Figure 27.28 (a) Sum-squaring circuit and (b) difference squaring circuit.

The output voltage of the sum-square circuit is given by

$$V_{o-} = VDD - I_{SQ(a)}R \tag{27.26}$$

while the output voltage of the difference-square circuit is given by

$$V_{o+} = VDD - I_{SQ(b)}R \tag{27.27}$$

A multiplier is formed by taking the difference between these voltages. The output voltage of the multiplier of Fig. 27.27 is given by

$$V_{out} = V_{o+} - V_{o-} = R\frac{\beta_{23}}{4}\left[(V_1 + V_2)^2 - (V_1 - V_2)^2\right] \tag{27.28}$$

or, using Eq. (27.23)

$$V_{out} = R\beta_{23} \cdot V_1 V_2 \tag{27.29}$$

The fundamental concern with using this type of multiplier is the fact that modern short-channel CMOS devices don't follow the square-law equations (Eq. [27.24]) we used to derive this result.

ADDITIONAL READING

[1] J. Crols and M. S. J. Steyaert, "A 1.5 GHz Highly Linear CMOS Downconversion Mixer," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 7, pp. 736–742, July 1995.

IPR2023-00698
Theta EX2019

[2] E. A. Vittoz, "Micropower Techniques," Chapter 3 in J. E. Franca and Y. Tsividis (eds.) *Design of Analog-Digital VLSI Circuits for Telecommunications and Signal Processing,* 2nd ed., Prentice Hall, 1994. ISBN 0-13-203639-8.

[3] M. Ismail, S-C. Huang, and S. Sakurai, "Continuous-Time Signal Processing," Chapter 3 in M. Ismail and T. Fiez (eds.), *Analog VLSI: Signal and Information Processing*, McGraw Hill, 1994. ISBN 0-07-032386-0.

[4] H-J. Song and C-K. Kim, "A MOS Four-Quadrant Analog Multiplier Using Simple Two-Input Squaring Circuits with Source Followers," *IEEE Journal of Solid-State Circuits*, vol. 25, no. 3, pp. 841–848, June 1990.

[5] B-S. Song, "CMOS RF Circuits for Data Communications Applications," *IEEE Journal of Solid-State Circuits*, vol. SC-21, no. 2, pp. 310–317, April 1986.

[6] S. Soclof, *Applications of Analog Integrated Circuits*, Prentice Hall, 1985. ISBN 0-13-039173-5.

[7] M. G. Degrauwe, J. Rijmenants, E. A. Vittoz, and H. J. DeMan, "Adaptive Biasing CMOS Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. SC-17, no. 3, pp. 522–528, June 1982.

[8] D. J. Allstot, "A Precision Variable-Supply CMOS Comparator," *IEEE Journal of Solid-State Circuits,* vol. SC-17, no. 6, pp. 1080–1087, December 1982.

PROBLEMS

27.1 Using the long-channel CMOS process, compare the performance (using simulations) of the comparator in Fig. 27.8 with the comparator in Fig. 27.9. Your comparison should include DC gain, systematic offset, delay, sensitivity, and power consumption.

27.2 Show, using simulations, how the addition of a balancing resistor in Fig. 27.14 can be used to improve the response seen in Fig. 27.13.

27.3 Simulate the operation of the comparator in Fig. 27.15 in the short-channel CMOS process. Determine the comparators sensitivity and the kickback noise.

27.4 Repeat problem 27.3 for the comparator in Fig. 27.16. Show that the input common-mode range of the comparator in Fig. 27.16 extends beyond the power supply rails.

27.5 Simulate the operation of the input buffer in Fig. 27.17 in the short-channel CMOS process. How sensitive is the buffer to input slew-rate? How symmetrical are the output rise and fall times? Suggest, and verify with simulations, a method to reduce the power consumed by the input buffer.

27.6 Design a low power clocked comparator for use with a Flash ADC (discussed in Ch. 29). Use the short-channel CMOS process and a clocking frequency of 250 MHz estimate the power dissipated by 256 of these comparators.

IPR2023-00698
Theta EX2019

Chapter

28

Data Converter Fundamentals

Data converters (a circuit that changes analog signals to digital representations or vice-versa) play an important role in an ever-increasing digital world. As more products perform calculations in the digital or discrete time domain, more sophisticated data converters must translate the digital data to and from our inherently analog world. This chapter introduces concepts of data conversion and sampling which surround this useful circuit.

28.1 Analog Versus Discrete Time Signals

Analog-to-digital converters, also known as A/Ds or ADCs, convert analog signals to discrete time or digital signals. Digital-to-analog converters (D/As or DACs) perform the reverse operation. Figure 28.1 illustrates these two operations.To understand the functionality of these data converters, it would be wise first to compare the characteristics of analog versus digital signals.

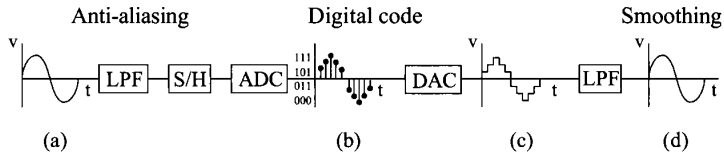


Figure 28.1 Signal characteristics caused by A/D and D/A conversion.

In Fig. 28.1 the original analog signal (a) is filtered by an anti-aliasing filter to remove any high-frequency components that may cause an effect known as aliasing (see Sec. 28.5). The signal is sampled and held and then converted into a digital signal (b). Next the DAC converts the digital signal back into an analog signal (c). Note that the output of the DAC is not as "smooth" as the original signal. A low-pass filter returns the analog signal back to its original form (plus phase shift introduced from the conversions)

IPR2023-00698
Theta EX2019

after eliminating the higher order signal components caused by the conversion. This example illustrates the main differences between analog and digital signals. Whereas the analog signal in Fig. 28.1a is *continuous* and *infinite* valued, the digital signal in (b) is *discrete* with respect to time and *quantized.* The term *continuous-time signal* refers to a signal whose response with respect to time is uninterrupted. Simply stated, the signal has a continuous value for the entire segment of time for which the signal exists. By referring to the analog signal as infinite valued, we mean that the signal can possess any value between the parameters of the system. For example, in Fig. 28.1a, if the peak amplitude of the sine wave was +1V, then the analog signal can be any value between –1 and 1 V (such as 0.4758393848 V). Of course, measuring all the values between –1 and 1 V would require a piece of laboratory equipment with infinite precision.

The digital signal, on the other hand, is discrete with respect to time. This means that the signal is defined for only certain or discrete periods of time. A signal that is quantized can only have certain values (as opposed to an infinitely valued analog signal) for each discrete period. The signal illustrated in Fig. 28.1b illustrates these qualities.

28.2 Converting Analog Signals to Digital Signals

We have already established the differences between analog and digital signals. How is it possible to convert from an analog signal to a digital signal? An example will illustrate the process.

Where you live the temperature in the winter stays between 0° F and 50° F (Fig. 28.2a). Suppose you had a thermometer with only two readings, hot and cold, and you wanted to record the weather patterns and plot the results. The two quantization levels can be correlated with the actual temperature as follows:

If $0° F \leq T < 25° F$	Temperature is recorded as cold
If $25° F \leq T < 50° F$	Temperature is recorded as hot

You take a measurement every day at noon and plot the results after one week. From Fig. 28.2b, it is apparent that your discretized version of the weather is not an accurate representation of the actual weather.

Now suppose that you find another thermometer with four possible temperatures (hot, warm, cool, and cold) and you increase the number of readings to two per day. The result of this reading is seen in Fig. 28.3a. The quantization levels represent four equal bands of temperature as seen below:

If $0° F \leq T < 12.5° F$	Temperature is recorded as cold
If $12.5°F \leq T < 25° F$	Temperature is recorded as cool
If $25° F \leq T < 37.5° F$	Temperature is recorded as warm
If $37.5° F \leq T < 50° F$	Temperature is recorded as hot

Here, the digital version of the weather still looks nothing like the actual weather pattern, but the critical issues in digitizing an analog signal should be apparent. The actual weather pattern is the analog signal. It is continuous with respect to time, and its value can be between 0° F and 50° F (even 33.9638483920398439° F!). The accuracy of the

IPR2023-00698
Theta EX2019

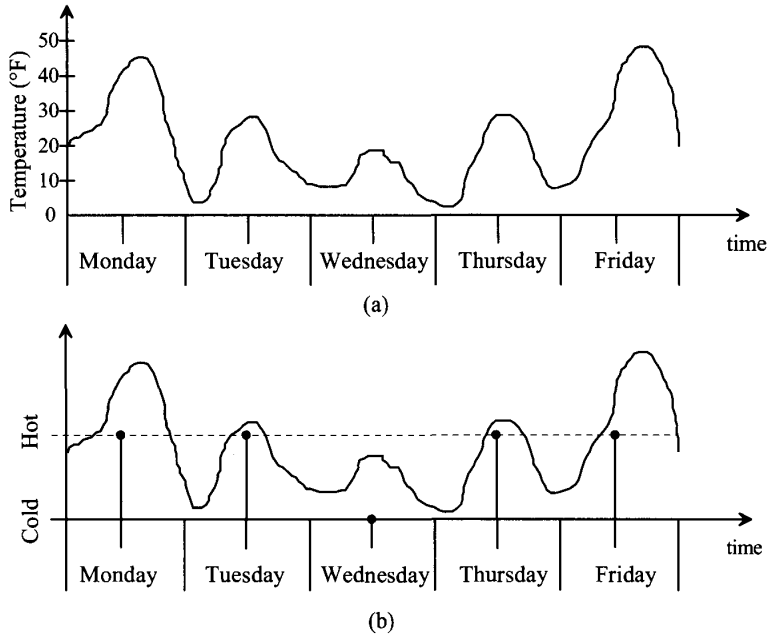


Figure 28.2 (a) An analog signal representing the temperature where you live and (b) a digital representation of the analog signal taking one sample per day with two quantization levels.

digitized signal is dependent on two things: the number of samples taken and the resolution, or number of quantization levels, of the converter. In our example, we need to increase both the number of samples and the resolution of thermometer.

Suppose that finally we obtain a thermometer with 25 temperature readings and that we take a reading eight times per day. Each of the 25 quantization levels now represents a 2° F band of temperature. From Fig. 28.3b, we can see that the digital version of the weather is approaching that of the actual analog signal. If we kept increasing both sampling time and resolution, the difference between the analog and the digital signals would become negligible. This brings up another critical issue: how many samples should one take in order to accurately represent the analog signal?

Suppose a sudden rainstorm swept through your town and caused a sharp decrease in temperature before returning to normal. If that storm had occurred between our sampling times, our experiment would not have shown the effects of the storm. Our sampling time was too slow to catch the change in the weather. If we had increased the number of samples, we would have recognized that something happened which caused the temperature to drop dramatically during that period.

IPR2023-00698
Theta EX2019

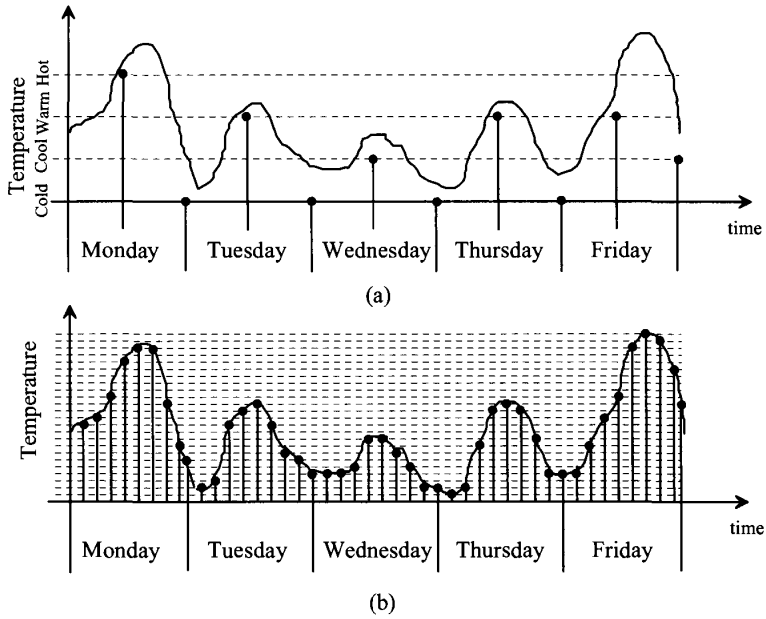


Figure 28.3 Digital representation of the temperature taking (a) two samples per day with four quantization levels and (b) nine samples per day with 25 quantization levels.

As it turns out, the *Nyquist Criterion* defines how fast the sampling rate needs to be to represent an analog signal accurately. This criterion requires that the sampling rate is at least two times the highest frequency contained in the analog signal. In our example, we need to know how quickly the weather can change and then take samples twice as fast as that value. The Nyquist Criterion can be described as

$$f_{sampling} = 2 f_{MAX} \tag{28.1}$$

where $f_{sampling}$ is the sampling frequency required to accurately represent the analog signal and f_{MAX} is the highest frequency of the sampled signal.

How much resolution should we use to represent the analog signal accurately? There is no absolute criterion for this specification. Each application will have its own requirements. In our weather example, if we were only interested in following general trends, then the 25 quantization levels would more than suffice. However, if we were interested in keeping an accurate record of the temperature to within ±0.5° F, we would need to double the resolution to 50 quantization levels so that each quantization level would correspond to each degree ±0.5° F (Fig. 28.4).

IPR2023-00698
Theta EX2019

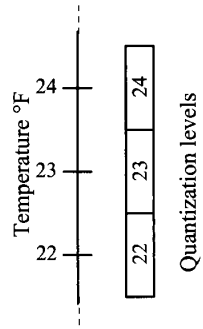


Figure 28.4 Quantization levels overlap actual temperature by ±½° F.

28.3 Sample-and-Hold (S/H) Characteristics

Sample-and-hold (S/H) circuits are critical in converting analog signals to digital signals. The behavior of the S/H is analogous to that of a camera. Its main function is to "take a picture" of the analog signal and hold its value until the ADC can process the information. It is important to characterize the S/H circuit when performing data conversion. Ignoring this component can result in serious error, for both speed and accuracy can be limited by the S/H. Ideally, the S/H circuit should have an output similar to that shown in Fig. 28.5a. Here, the analog signal is instantly captured and held until the next sampling period. However, a finite period of time is required for the sampling to occur. During the sampling period, the analog signal may continue to vary; thus, another type of circuit is called a track-and-hold, or T/H. Here, the analog signal is "tracked" during the time required to sample the signal, as seen in Fig. 28.5b. It can be seen that S/H circuits operate in both static (hold mode) and dynamic (sample mode) circumstances. Thus, characterization of the S/H will be discussed in the context of these two categories. Figure 28.6 presents a summary of the major errors associated with a S/H. A discussion of each error follows.

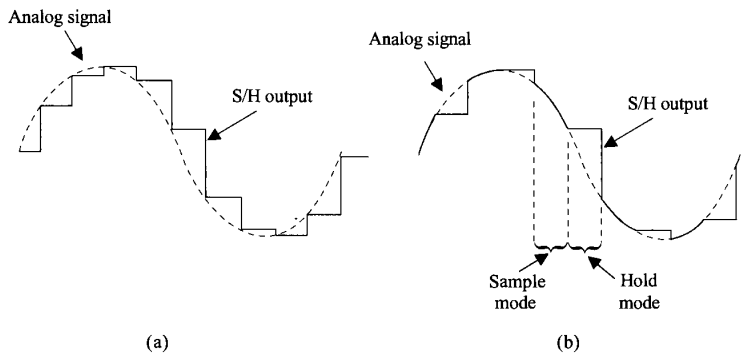


Figure 28.5 The output of (a) an ideal S/H circuit and (b) a track-and-hold (T/H).

IPR2023-00698
Theta EX2019

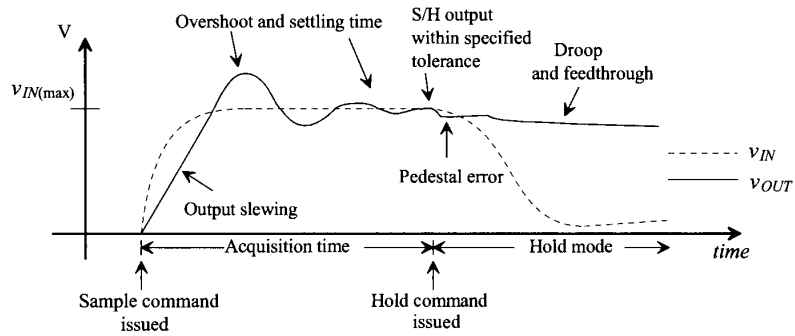


Figure 28.6 Typical errors associated wth an S/H.

Sample Mode

Once the sampling command has been issued, the time required for the S/H to track the analog signal to within a specified tolerance is known as the *acquisition time.* In the worst-case scenario, the analog signal would vary from zero volts to its maximum value, $v_{IN(max)}$. And the worst-case acquisition time would correspond to the time required for the output to transition from zero to $v_{IN(max)}$. Since most S/H circuits use amplifiers as buffers (as seen in Fig. 28.7), it should be obvious that the acquisition will be a function of the amplifier's own specifications. For example, notice that if the input changes very quickly, then the output of the T/H could be limited by the amplifier's slew rate. The amplifier's stability is also extremely critical. If the amplifier is not compensated correctly, and the phase margin is too small, then a large *overshoot* will occur. A large overshoot requires a longer *settling time* for the S/H to settle within the specified tolerance. The error tolerance at the output of the S/H also depends on the amplifier's *offset, gain error* (ideally, the S/H should have a gain of 1) and *linearity* (the gain of the S/H should not vary over the input voltage range).

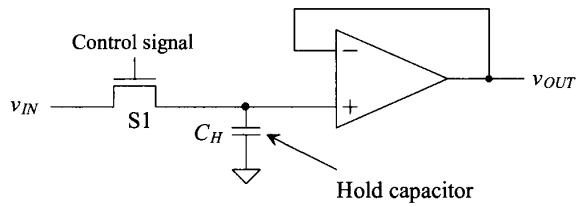


Figure 28.7 Track-and-hold circuit using an output buffer.

IPR2023-00698
Theta EX2019

Hold Mode

Once the hold command is issued, the S/H faces other errors. Pedestal error occurs as a result of charge injection and clock feedthrough. Part of the charge built up in the channel of the switch is distributed onto the capacitor, thus slightly changing its voltage. Also, the clock couples onto the capacitor via overlap capacitance between the gate and the source or drain. Another error that occurs during the hold mode is called *droop*. This error is related to the leakage of current from the capacitor due to parasitic impedances and to the leakage through the reverse-biased diode formed by the drain of the switch. This diode leakage can be minimized by making the drain area as small as can be tolerated. Although the input impedance of the buffer amplifier is very large, the switch has a finite OFF impedance through which leakage can occur. Current can also leak through the substrate. The key to minimizing droop is increasing the value of the sampling capacitor. The trade-off, however, is increased time that's required to charge the capacitor to the value of the input signal.

Aperture Error

A transient effect that introduces error occurs between the sample and the hold modes. A finite amount of time, referred to as aperture time, is required to disconnect the capacitor from the analog input source. The aperture time actually varies slightly as a result of noise on the hold-control signal and the value of the input signal, since the switch will not turn off until the gate voltage becomes less than the value of the input voltage less one threshold voltage drop. This effect is called *aperture uncertainty* or *aperture jitter*. As a result, if a periodic signal were being sampled repeatedly at the same points, slight variations in the hold value would result, thus creating *sampling error*. Figure 28.8 illustrates this effect. Note that the amount of aperture error is directly related to the frequency of the signal and that the worst-case aperture error occurs at the zero crossing, where *dV/dt* is the greatest. This assumes that the S/H circuit is capable of sampling both positive and negative voltages (bipolar). The amount of error that can be tolerated is directly related to the resolution of the conversion. Aperture error will be discussed again in Sec. 28.5 as it relates to the error in an ADC.

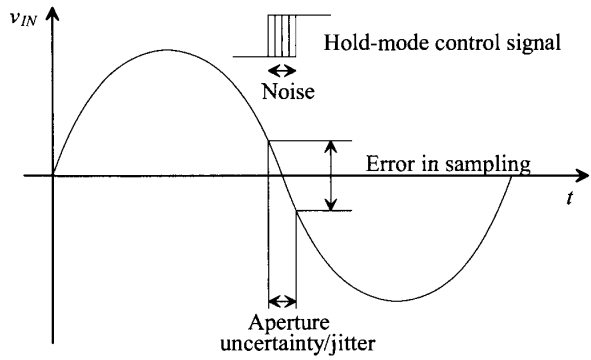


Figure 28.8 Aperture error.

IPR2023-00698
Theta EX2019

Example 28.1
Find the maximum sampling error for a S/H circuit that is sampling a sinusoidal input signal that could be described as

$$v_{IN} = A \sin 2\pi ft$$

where A is 2 V and f = 100 kHz. Assume that the aperture uncertainty is equal to 0.5 ns.

The sampling error due to the aperture uncertainty can be thought of as a slew rate such that

$$\frac{dV}{dt} = \frac{d}{dt} A \sin 2\pi ft = 2\pi fA \cos 2\pi ft$$

with maximum slewing occurring when the cosine term is equal to 1. Therefore,

$$\frac{dV}{dt}(max) = 2\pi fA = (2\pi \cdot 100 \text{ kHz})(2 \text{ V})$$

and the maximum sampling error is

$$Maximum\ Sampling\ Error = dV(max) \quad \text{or}$$

$$(0.5 \times 10^{-9} \text{ s})(2\pi \cdot 100 \text{ kHz})(2 \text{ V}) = 0.628 \text{ mV} \ \blacksquare$$

28.4 Digital-to-Analog Converter (DAC) Specifications

Probably the most popular digital-to-analog converter application is converting stored digital audio and/or video signals. For example, stored digital information in MP3 format can be converted into music via a high-precision DAC. Many characteristics define a DAC's performance. Each characteristic will be discussed before we look at the basic architectures in Ch. 29. This "top-down" approach allows a smoother transition from the data converter characteristics to the actual architectures, since most data converters have similar performance limitations. A discussion of some of the basic definitions associated with DACs follows. It should be noted that DACs and ADCs can use either voltage or current as their analog signal. For purposes of describing specifications, it will be assumed that the analog signal is a voltage.

A block diagram of a DAC can be seen in Fig. 28.9. Here an N-bit digital word is mapped into a single analog voltage. Typically, the output of the DAC is a voltage that is some fraction of a reference voltage (or current), such that

$$v_{OUT} = FV_{REF} \tag{28.2}$$

where v_{OUT} is the analog voltage output, V_{REF} is the reference voltage, and F is the fraction defined by the input word, D, that is N bits wide. The number of input combinations represented by the input word D is related to the number of bits in the word by

$$\text{Number of input combinations} = 2^N \tag{28.3}$$

A 4-bit DAC has a total of 2^4 or 16 total input values. A converter with 4-bit resolution must be able to map a change in the analog output, which is equal to 1 part in 16. The maximum analog output voltage for any DAC is limited by the value of V_{REF}. If the input is an N-bit word, then the value of the fraction, F, can be determined by,

IPR2023-00698
Theta EX2019

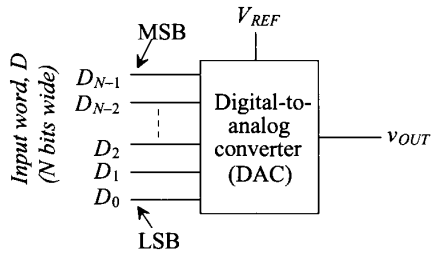


Figure 28.9 Block diagram of the digital-to-analog converter.

$$F = \frac{D}{2^N} \tag{28.4}$$

Therefore, if a 3-bit DAC is being used, the input, D, is $100 = 4_{10}$, and V_{REF} is 5 V, then the value of F is

$$F = \frac{100}{2^3} = \frac{4}{8} \tag{28.5}$$

and the analog voltage that appears at the output becomes,

$$v_{OUT} = \frac{4}{8}(5) = 2.5 \text{ V} \tag{28.6}$$

By plotting the input word, D, versus v_{OUT} as D is incremented from 000 to 111, the *transfer curve* seen in Fig. 28.10 would be generated. The y-axis has been normalized to V_{REF}; therefore, the graduated marks also represent F by Eq. (28.2). Some important characteristics need to be discussed here. First, notice that the transfer curve is not continuous. Since the input is a digital signal, which is inherently discrete, the input signal can only have eight values that must correspondingly produce eight output voltages. If a straight line connected each of the output values, the slope of the line would ideally be one increment/input code value. Also note that the maximum value of the output is 7/8. Since the case where $D = 000$ has to result in an analog voltage of 0 V, and a 3-bit DAC has eight possible analog output voltages, then the analog output will increase from 0 V to only 7/8 V_{REF}.

Again, using Eq. (28.2), this means that the maximum analog output that can be generated by the 3-bit DAC is

$$v_{OUT(\max)} = \frac{7}{8} \cdot V_{REF} \tag{28.7}$$

This maximum analog output voltage that can be generated is known as *full-scale voltage*, V_{FS}, and can be generalized to any N-bit DAC as

$$V_{FS} = \frac{2^N - 1}{2^N} \cdot V_{REF} \tag{28.8}$$

IPR2023-00698
Theta EX2019

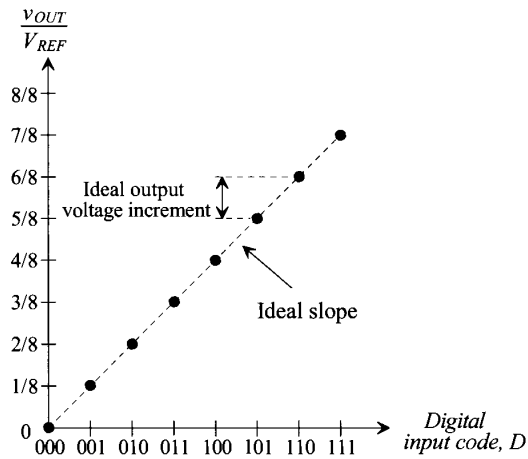


Figure 28.10 Ideal transfer curve for a 3-bit DAC.

The *least significant bit* (*LSB*) refers to the rightmost bit in the digital input word. The LSB defines the smallest possible change in the analog output voltage. The LSB will always be denoted as D_0. One LSB can be defined as

$$1\ LSB = \frac{V_{REF}}{2^N} \tag{28.9}$$

In the previous case of the 3-bit DAC, 1 LSB = 5/8 V, or 0.625 V. Generating an output in multiples of 0.625 V may not seem difficult, but as the number of bits increases, the voltage value of one LSB decreases for a fixed value of V_{REF}.

The *most significant bit* (*MSB*) refers to the leftmost bit of the digital word, *D*. In the previous example, $D = 100$ or $D_2D_1D_0$, with D_2 being the MSB. Generalizing to the *N*-bit DAC, the MSB would be denoted as D_{N-1}. (Since the LSB is denoted as bit 0, the MSB is denoted as *N*-1.) Note that when discussing DACs, the MSB causes the output to change by 1/2 V_{REF}.

When discussing data converters, the term *resolution* describes the smallest change in the analog output with respect to the value of the reference voltage, V_{REF}. This is slightly different from the definition of LSB in that resolution is typically given in terms of bits and represents the *number of unique output voltage levels*, i.e., 2^N.

Example 28.2

Find the resolution for a DAC if the output voltage is desired to change in 1 mV increments while using a reference voltage of 5 V.

The DAC must resolve

$$\frac{1\text{ mV}}{5\text{ V}} = 0.0002 \text{ or } 0.02\% \text{ adjustability}$$

IPR2023-00698
Theta EX2019

Therefore, the *accuracy* required for 1 LSB change over a range of V_{REF} is

$$\frac{1\,LSB}{V_{REF}} = \frac{1}{2^N} = 0.0002 \tag{28.10}$$

and solving N for the resolution yields

$$N = \text{Log}_2(\frac{5\text{ V}}{1\text{ mV}}) = 12.29 \text{ bits}$$

which means that a 13-bit DAC will be needed to produce the accuracy capable of generating 1 mV changes in the output using a 5 V reference. ■

Example 28.3

Find the number of input combinations, values for 1 LSB, the percentage accuracy, and the full-scale voltage generated for a 3-bit, 8-bit, and 16-bit DAC, assuming that V_{REF}= 5 V.

Using Eqs. (28.3), (28.8), (28.9), and (28.10), we can generate the following information:

Resolution	Input combinations	1 LSB	% accuracy	V_{FS}
3	8	0.625 V	12.5	4.375 V
8	256	19.5 mV	0.391	4.985 V
16	65,536	76.29 μV	0.00153	4.9999 V

The value of 1 LSB for an 8-bit converter is 19.5 mV, while 1 LSB for a 16-bit converter is 76.3 μV (a factor of 256)! Increasing the resolution by 1 bit increases the accuracy by a factor of 2. The precision required to map the analog signal at high resolutions is very difficult to achieve. We will examine some of these issues as we examine the limitations of the data converter in Ch. 29.

Note that a data converter may have a resolution of 8 bits, where an LSB is 19.5 mV as above, while having a much higher accuracy. For example, we could require the 8-bit data converter above to have an accuracy of 0.1%. The higher accuracy results in a more ideal (linear) DAC. A typical specification for DAC accuracy is ±½ LSB for reasons discussed below. ■

Differential Nonlinearity

As seen in the ideal DAC in Fig. 28.10, each adjacent output increment should be exactly one-eighth. Since the y-axis is normalized, the values for the increment heights will be unitless. However, the increment heights can be easily converted to volts by multiplying the height by V_{REF}. This corresponds to the ideal increment corresponding to 0.625 V = 1 LSB (assuming V_{REF} = 5 V).

Nonideal components cause the analog increments to differ from their ideal values. The difference between the ideal and nonideal values is known as *differential nonlinearity*, or *DNL* and is defined as

DNL_n = Actual increment height of transition n – Ideal increment height (28.11)

IPR2023-00698
Theta EX2019

where n is the number corresponding to the digital input transition. The DNL specification measures how well a DAC can generate uniform analog LSB multiples at its output.

Example 28.4

Determine the DNL for the 3-bit nonideal DAC whose transfer curve is shown in Fig. 28.11. Assume that $V_{REF} = 5$ V.

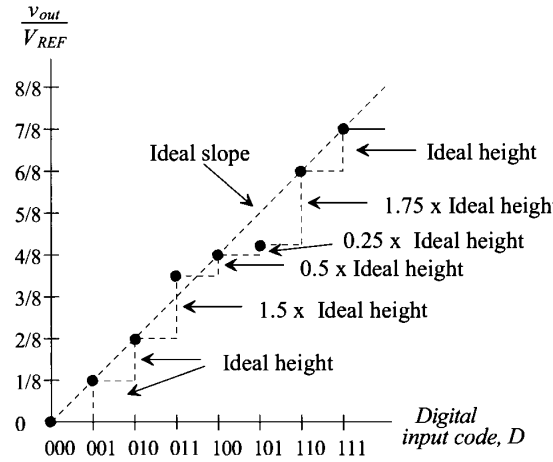


Figure 28.11 Example of differential nonlinearity for a 3-bit DAC.

The actual increment heights are labeled with respect to the ideal increment height, which is 1 LSB, or 1/8 of $\frac{v_{OUT}}{V_{REF}}$. Notice that there is no increment corresponding to 000, since it is desirable to have zero output voltage with a digital input code of 000. The increment height corresponding to 001, however, is equal to the corresponding height of the ideal case seen in Fig. 28.10; therefore, $DNL_1 = 0$. Similarly, DNL_2 is also zero since the increment associated with the transition at 010 is equal to the ideal height. Notice that the 011 increment, however, is not equal to the ideal curve but is 3/16, or 1.5 times the ideal height.

$$DNL_3 = 1.5 \text{ LSB} - 1 \text{ LSB} = 0.5 \text{ LSB}$$

Since we have already determined in Eq. (28.9) that for a 3-bit DAC, 1 LSB = 0.625 V, we can convert the DNL_3 to volts as well. Therefore, $DNL_3 = 0.5$ LSB = 0.3125 V. However, it is popular to refer to DNL in terms of LSBs. The remainder of the digital output codes can be characterized as follows:

$$DNL_4 = 0.5 \text{ LSB} - 1 \text{ LSB} = -0.5 \text{ LSB}$$

$$DNL_5 = 0.25 \text{ LSB} - 1 \text{ LSB} = -0.75 \text{ LSB}$$

$$DNL_6 = 1.75 \text{ LSB} - 1 \text{ LSB} = 0.75 \text{ LSB}$$

$$DNL_7 = 1 \text{ LSB} - 1 \text{ LSB} = 0$$

IPR2023-00698
Theta EX2019

If we were to plot the value of DNL (in LSBs) versus the input digital code, Fig. 28.12 would result. The DNL for the entire converter used in this illustration is ±0.75 LSB since the overall error of the DAC is defined by its worst-case DNL. ■

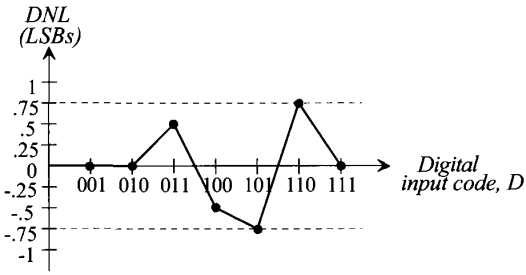


Figure 28.12 DNL curve for the nonideal 3-bit DAC.

Generally, a DAC will have less than ±½ LSB of DNL if it is to be *N*-bit accurate. A 5-bit DAC with 0.75 LSBs of DNL actually has the resolution of a 4-bit DAC. If the DNL for a DAC is less than –1 LSBs, then the DAC is said to be *nonmonotonic*, which means that the analog output voltage does not always increase as the digital input code is incremented. A DAC should always exhibit *monotonicity* if it is to function without error.

Integral Nonlinearity

Another important static characteristic of DACs is called *integral nonlinearity* (*INL*). Defined as the difference between the data converter output values and a reference straight line drawn through the first and last output values, INL defines the linearity of the overall transfer curve and can be described as

$$\text{INL}_n = \text{Output value for input code } n - \text{Output value of the reference line at that point} \tag{28.12}$$

An illustration of this measurement is presented in Fig. 28.13. It is assumed that all other errors due to offset and gain (these will be discussed shortly) are zero. An example follows shortly.

It is common practice to assume that a converter with *N*-bit resolution will have less than ±½ LSB of DNL and INL. The term, *½ LSB*, is a common term that typically denotes the maximum error of a data converter (both DACs and ADCs). For example, a 13-bit DAC having greater than ±½ LSB of DNL or INL actually has the resolution of a 12-bit DAC. The value of ½ LSB in volts is simply

$$0.5\,LSB = \frac{V_{REF}}{2^{N+1}} \tag{28.13}$$

IPR2023-00698
Theta EX2019

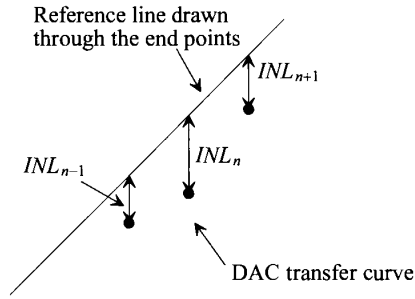


Figure 28.13 Measuring the INL for a DAC transfer curve.

Example 28.5

Determine the INL for the nonideal 3-bit DAC shown in Fig. 28.14. Assume that $V_{REF} = 5$ V.

First, a reference line is drawn through the first and last output values. The INL is zero for every code in which the output value lies on the reference line; therefore, $INL_2 = INL_4 = INL_6 = INL_7 = 0$. Only outputs corresponding to 001, 011, and 101 do not lie on the reference. Both the 001 and the 011 transitions occur ½ LSB higher than the straight-line values; therefore, $INL_1 = INL_3 = 0.5$ LSB. By the same reasoning, $INL_5 = -0.75$ LSB. Therefore, the INL for the DAC is considered to be its worst-case INL of +0.5 LSB and –0.75 LSB. The INL plot for the nonideal 3-bit DAC can be seen in Fig. 28.15. ■

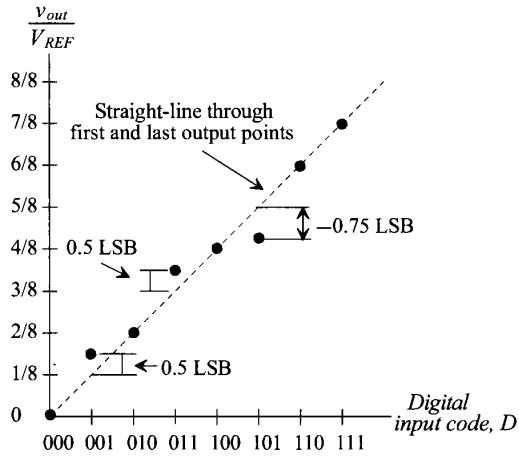


Figure 28.14 Example of integral nonlinearity for a DAC.

IPR2023-00698
Theta EX2019

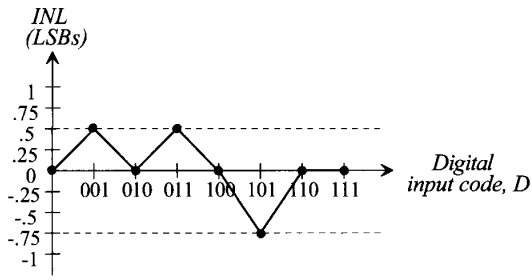


Figure 28.15 INL curve for the nonideal 3-bit DAC.

It should be noted that other methods are used to determine INL. One method compares the output values to the ideal reference line, regardless of the positions of the first and last output values. If the DAC has an offset voltage or gain error, this will be included in the INL determination. Usually, the offset and gain errors are determined as separate specifications.

Another method, described as the "best-fit" method, attempts to minimize the INL by constructing the reference line so that it passes as closely as possible to a majority of the output values. Although this method does minimize the INL error, it is a rather subjective method that is not as widely used as drawing the reference line through the first and last output values.

Offset

The analog output should be 0 V for $D = 0$. However, an offset exists if the analog output voltage is not equal to zero. This can be seen as a shift in the transfer curve as illustrated in Fig. 28.16. This specification is similar to the offset voltage for an operational amplifier except that it is not referred to the input.

Gain Error

A gain error exists if the slope of the best-fit line through the transfer curve is different from the slope of the best-fit line for the ideal case. For the DAC illustrated in Fig. 28.17, the gain error becomes

$$\text{Gain error} = \text{Ideal slope} - \text{Actual slope} \tag{28.14}$$

Latency

This specification defines the total time from the moment that the input digital word changes to the time the analog output value has settled to within a specified tolerance. Latency should not be confused with settling time, since latency includes the delay required to map the digital word to an analog value plus the settling time. It should be noted that settling time considerations are just as important for a DAC as they are for a S/H or an operational amplifier.

Signal-to-Noise Ratio (SNR)

Signal-to-noise (*SNR*) is defined as the ratio of the signal power to the noise at the analog output. In amplifier applications, this specification is typically measured using a sinewave

IPR2023-00698
Theta EX2019

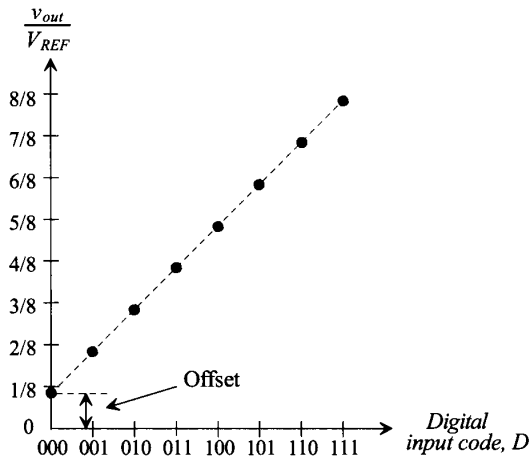


Figure 28.16 Illustration of offset error for a 3-bit DAC.

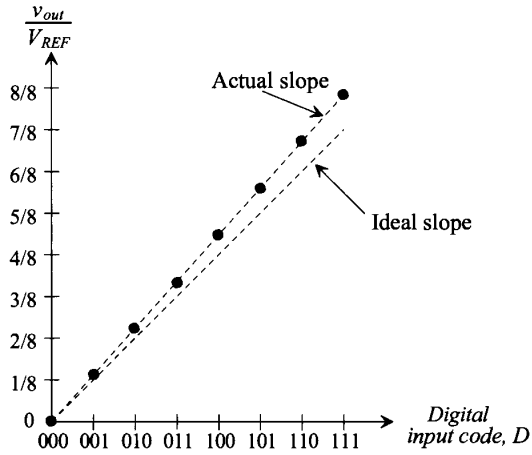


Figure 28.17 Illustration of gain error for a 3-bit DAC.

input. For the DAC, a "digital" sinewave is generated through instrumentation or through an A/D. The *SNR* can reveal the true resolution of a data converter as the effective number of bits can be quantified mathematically. A detailed derivation of the *SNR* is presented in Sec. 28.6, on the discussion of ADC specifications.

IPR2023-00698
Theta EX2019

Dynamic Range

Dynamic range is defined as the ratio of the largest output signal over the smallest output signal. For both DACs and ADCs, the dynamic range is related to the resolution of the converter. For example, an N-bit DAC can produce a maximum output of $2^N - 1$ multiples of LSBs and a minimum value of 1 LSB. Therefore, the dynamic range in decibels is simply

$$DR = 20\text{Log}\left(\frac{2^N - 1}{1}\right) \approx 6.02 \cdot N \text{ dB} \tag{28.15}$$

A 16-bit data converter has a dynamic range of 96.33 dB.

28.5 Analog-to-Digital Converter (ADC) Specifications

Many of the specifications that describe the ADC are similar to those that describe the DAC. However, there are subtle differences. Since the DAC is converting a discrete signal into an analog representation that is also limited by the resolution of the converter, a fixed number of inputs and outputs are generated. However, with the ADC, the input is an analog signal with an infinite number of values, which then has to be quantized into an N-bit digital word (Fig. 28.18). This process is much more difficult than the digital-to-analog process. In fact, many ADC architectures use a DAC as a critical component.

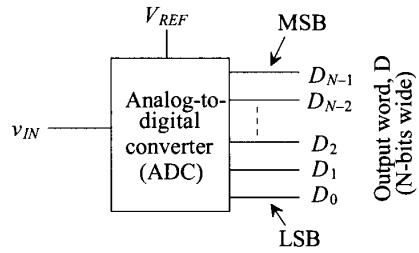


Figure 28.18 Block diagram of the analog-to-digital converter.

For example, in the previous discussion of DACs, it was determined that for a 16-bit DAC, the converter would need to generate output voltages in multiples of 76 μV. However, for the ADC, the converter needs to resolve differences in the analog signal of 76 μV. This means that the ADC must be able to detect changes in the input signal on the order of 1 part in 65,536! In contrast, the DAC had a finite number of input combinations (2^N). The ADC, however, has to "quantize" the infinite-valued analog signal into many segments so that

$$\text{Number of quantization levels} = 2^N \tag{28.16}$$

This distinction is subtle but must be recognized to understand the differences between the two types of conversion.

IPR2023-00698
Theta EX2019

Examine Fig. 28.19a. The digital output, D, of an ideal, 3-bit ADC is plotted versus the analog input, v_{IN}. Note the difference in the transfer curve for the ADC versus the DAC (Fig. 28.10). The y-axis is now the digital output, and the x-axis has been normalized to V_{REF}. Since the input signal is a continuous signal and the output is discrete, the transfer curve of the ADC resembles that of a staircase. Another fact to observe is that the 2^N quantization levels correspond to the digital output codes 0 to 7. Thus, the maximum output of the ADC will be 111 ($2^N - 1$), corresponding to the value for which $\frac{v_{IN}}{V_{REF}} \geq \frac{7}{8}$. Figure 28.19b corresponds to the error caused by the quantization.

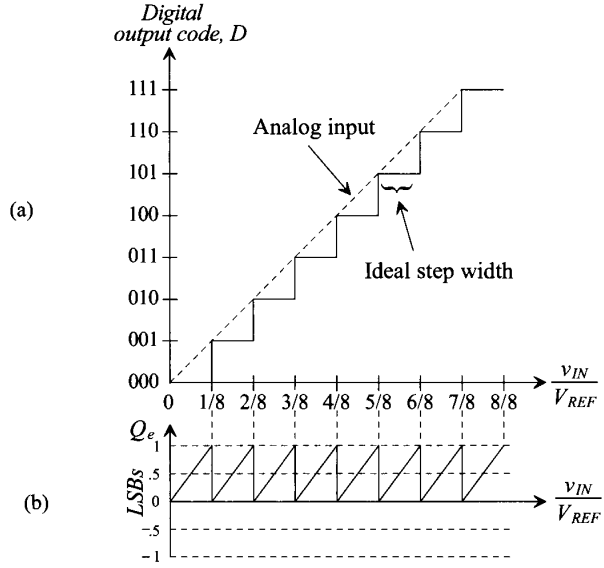


Figure 28.19 (a) Transfer curve for an ideal ADC and (b) its corresponding quantization error.

The value of 1 LSB for this ADC can be calculated using Eq. (28.9) and is the ideal step width (1/8) in Fig. 28.19 (versus the height for the DAC) multiplied by V_{REF}. Therefore, assuming that $V_{REF} = 5$ V,

$$1 \text{ LSB} = 0.625 \text{ V} \tag{28.17}$$

Quantization Error

Since the analog input is an infinite valued quantity and the output is a discrete value, an error will be produced as a result of the quantization. This error, known as *quantization error*, Q_e, is defined as the difference between the actual analog input and the value of the output (staircase) given in voltage. It is calculated as

$$Q_e = v_{IN} - V_{staircase} \tag{28.18}$$

IPR2023-00698
Theta EX2019

where the value of the staircase output, $V_{staircase}$, can be calculated by

$$V_{staircase} = D \cdot \frac{V_{REF}}{2^N} = D \cdot V_{LSB} \tag{28.19}$$

where D is the value of the digital output code and V_{LSB} is the value of 1 LSB in volts, in this case 0.625 V. We can also easily convert the value of Q_e in units of LSBs. In Fig. 28.19a, Q_e can be generated by subtracting the value of the staircase from the dashed line. The result can be seen in Fig. 28.19b. A sawtooth waveform is formed centered about ½ LSBs. Ideally, the magnitude of Q_e will be no greater than one LSB and no less than 0. It would be advantageous if the quantization error were centered about zero so that the error would be at most ±½ LSBs (as opposed to +1LSB). This is easily achieved as seen in Fig. 28.20a and b. Here, the entire transfer curve is shifted to the left by ½ LSB, thus making the codes centered around the LSB increments on the x-axis. This drawing illustrates that at best, an ideal ADC will have quantization error of ±½ LSB.

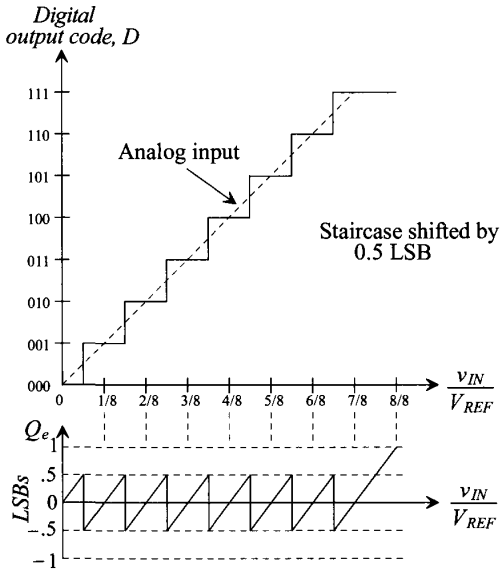


Figure 28.20 (a) Transfer curve for an ideal 3-bit ADC with (b) quantization error centered about zero.

In shifting this curve to the left, notice that the first code transition occurs when $\frac{v_{IN}}{V_{REF}} \geq \frac{1}{16}$. Therefore, the range of $\frac{v_{IN}}{V_{REF}}$ for the digital output corresponding to 000 is half as wide as the ideal step width. The last code transition occurs when $\frac{v_{IN}}{V_{REF}} \geq \frac{13}{16}$ (between 6/8 and 7/8). Note that the step width corresponding to this last code transition is 1.5 times larger than the ideal width and that the quantization error extends up to 1 LSB when $\frac{v_{IN}}{V_{REF}} = 1$. However, the converter would be considered to be out of range once $\frac{v_{IN}}{V_{REF}} \geq \frac{15}{16}$ (halfway between 7/8 and 8/8), so the problem is moot.

IPR2023-00698
Theta EX2019

Differential Nonlinearity

Differential nonlinearity for an ADC is similar to that defined for a DAC. However, for the ADC, DNL is the difference between the actual code *width* of a nonideal converter and the ideal case. Figure 28.21 illustrates the transfer curve for a nonideal 3-bit ADC. The values for the DNL can be solved as follows:

$$\text{DNL} = \text{Actual step width} - \text{Ideal step width} \tag{28.20}$$

Since the step widths can be converted to either volts for LSBs, DNL can be defined using either units. The value of the ideal step is 1/8. Converting to volts, this becomes

$$V_{idealstepwidth} = \tfrac{1}{8} \cdot V_{REF} = 0.625\ V = 1\ \text{LSB} \tag{28.21}$$

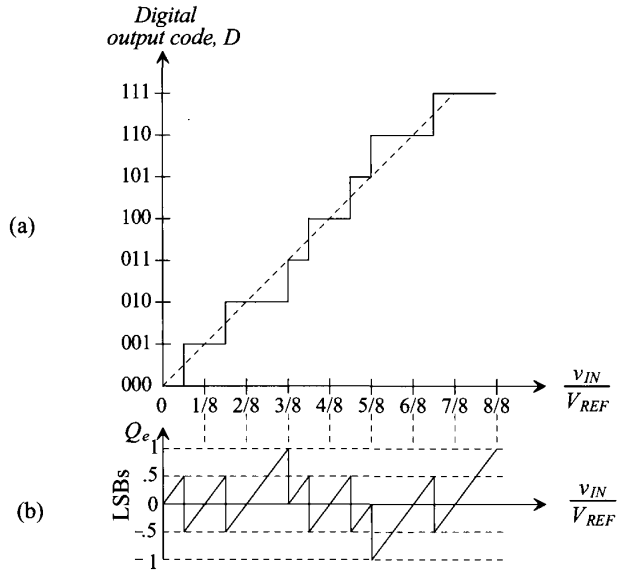


Figure 28.21 (a) Transfer curve for a nonideal 3-bit ADC used in Ex. 28.4 with (b) quantization error illustrating differential nonlinearity.

Example 28.6

Using Fig. 28.21a, calculate the differential nonlinearity of the 3-bit ADC. Assume that $V_{REF} = 5$ V. Draw the quantization error, Q_e, in units of LSBs.

The DNL of the converter can be calculated by examining the step width of each digital output code. Since the ideal step width of the 000 transition is ½ LSB, then $\text{DNL}_0 = 0$. Also note that the step widths associated with 001 and 100 are equal to 1 LSB; therefore, both DNL_1 and DNL_4 are zero. However, the remaining values code widths are not equal to the ideal value but can be calculated as

IPR2023-00698
Theta EX2019

$DNL_2 = 1.5\ LSB - 1\ LSB = 0.5\ LSB$

$DNL_3 = 0.5\ LSB - 1\ LSB = -0.5\ LSB$

$DNL_5 = -0.5\ LSB$

$DNL_6 = 0.5\ LSB$

$DNL_7 = 0\ LSB$ (since the ideal step width is 1.5 LSB wide at this code transition)

The overall DNL for the converter used in this illustration is ±0.5 LSB. Note that the quantization error illustrated in Fig. 28.21b is directly related to the DNL. As DNL increases in either direction, the quantization error worsens. Each "tooth" in the quantization error waveform should ideally be the same size. ■

Missing Codes

It is of interest to note the consequences of having a DNL that is equal to –1 LSB. Figure 28.22 illustrates an ADC for which this is true. The total width of the step corresponding to 101 is completely missing; thus, the value of DNL_5 is –1 LSB. Any ADC possessing a DNL that is equal to –1 LSB is guaranteed to have a missing code. Notice that the step width corresponding to 010 is 2 LSBs and that the value for DNL_2 is +1 LSB. However, there is not a missing code corresponding to 011, since the step width of code 011 depends on the 100 transition. Therefore, an ADC having a DNL greater than +1 LSB is not guaranteed to have a missing code, though in all probability a missing code will occur.

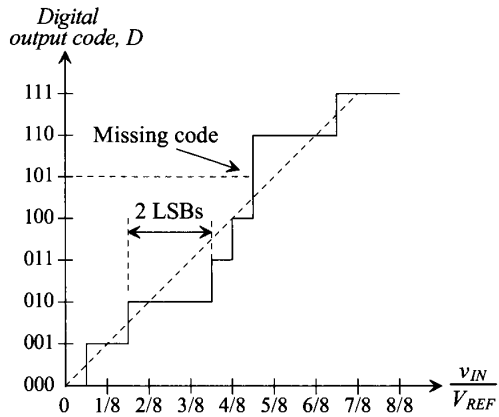


Figure 28.22 Transfer curve for a nonideal 3-bit ADC with a missing code.

Integral Nonlinearity

Integral nonlinearity (INL) is defined similarly to that for a DAC. Again, a "best-fit" straight line is drawn through the end points of the first and last code transition, with INL being defined as the difference between the data converter code transition points and the straight line with all other errors set to zero.

IPR2023-00698
Theta EX2019

Example 28.7

Determine the INL for the ADC whose transfer curve is illustrated in Fig. 28.23a. Assume that V_{REF} = 5 V. Draw the quantization error, Q_e, in units of LSBs.

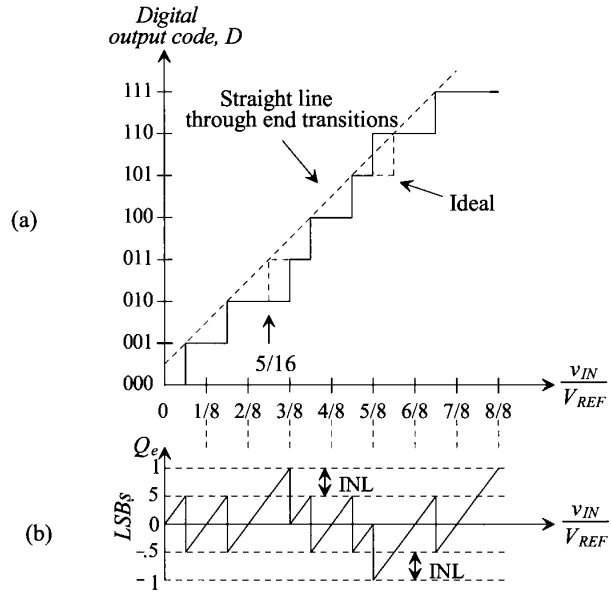


Figure 28.23 (a) Transfer curve of a nonideal 3-bit ADC and (b) its quantization error illustrating INL.

By inspection, it can be seen that all of the transition points occur on the best-fit line except for the transitions associated with code 011 and 110. Therefore,

$$INL_0 = INL_1 = INL_2 = INL_4 = INL_5 = INL_7 = 0$$

The INL corresponding to the remaining codes can be calculated as

$$INL_3 = 3/8 - 5/16 = 1/16 \text{ or } 0.5 \text{ LSB}$$

Similarly, INL_6 can be calculated in the same manner and is found to be –0.5 LSB. Thus, the overall INL for the converter is the maximum value of INL corresponding to ±0.5 LSB.

The INL can also be determined by inspecting the quantization error in Fig. 28.23b. Here, the INL will be the magnitude of the quantization error which lies outside the ±½ LSB band of Q_e. It can be seen that Q_e = 1 LSB, corresponding to the point at which INL = 0.5 LSB for digital output code 011, and that Q_e = –1 LSB at the output code corresponding to INL = –0.5 LSB for digital output code 110. ■

IPR2023-00698
Theta EX2019

Offset and Gain Error

Offset and gain error are identical to the DAC case. *Offset error* occurs when there is a difference between the value of the first code transition and the ideal value of ½ LSBs. As seen in Fig. 28.24a, the offset error is a constant value. Note that the quantization error becomes ideal after the initial offset voltage is overcome. *Gain error* or *scale factor error*, seen in Fig. 28.24b, is the difference in the slope of a straight line drawn through the transfer characteristic and the slope of 1 of an ideal ADC. Causes of offset and gain error are discussed in Ch. 29, but it is important here to understand their overall effects on ADC transfer curves.

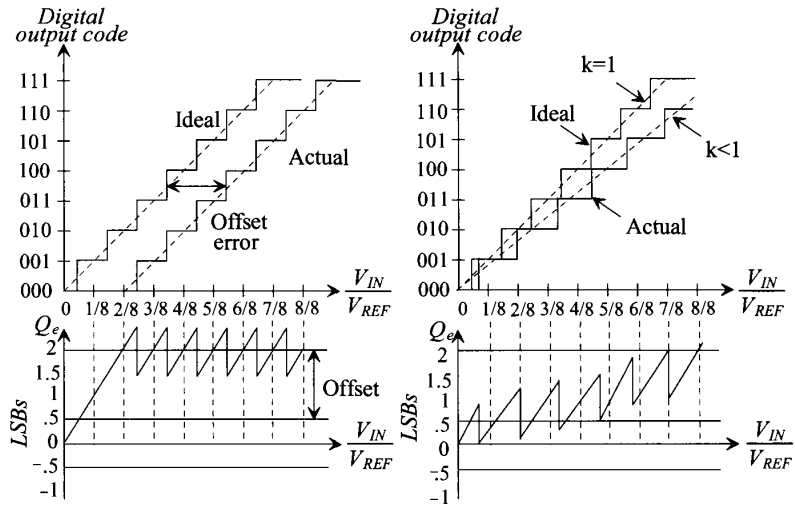


Figure 28.24 Transfer curve illustrating (a) offset error and (b) gain error.

So far, we have examined only the DC characteristics of an ADC. However, examining the dynamic aspects of the converter will lead to a whole new set of errors. Sampling is inherently a dynamic process since the accuracy of the sample is dependent on the speed of the analog signal. Many effects that occur during sampling limit the overall performance of the converter.

Aliasing

As mentioned earlier in the chapter, the Nyquist Criterion requires that a signal be sampled at least two times the highest frequency contained in the signal. What would happen if this criterion were ignored and the sampling rate was actually less than that amount? A phenomenon known as aliasing would occur.

Examine Fig. 28.25. Here, an analog signal is being sampled at a rate slower than the Nyquist Criterion requires. As a result, it appears that a totally different signal (see

IPR2023-00698
Theta EX2019

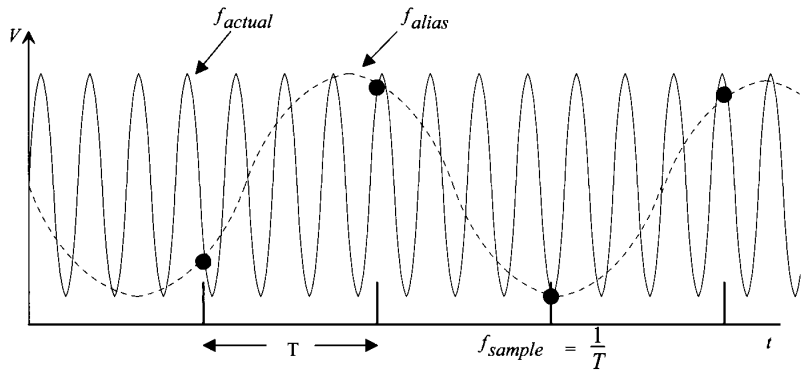


Figure 28.25 Aliasing caused by undersampling.

example dashed line) is being sampled. The different frequency signal is an "alias" of the original signal, and its frequency can be calculated using

$$f_{alias} = f_{actual} + k f_{sample} \quad (k = \ldots -2, -1, 0, 1, 2, 3 \ldots) \tag{28.22}$$

where f_{actual} is the frequency of the analog signal, f_{sample} is the sampling frequency, and f_{alias} is the frequency of the alias signal.

Aliasing can be eliminated by both sampling at higher frequencies and by filtering the analog signal before sampling and removing any frequencies that are greater than one-half the sampling frequency. It is good practice to filter the analog signal before sampling to eliminate any unknown higher order frequency components or noise that could result in aliasing.

A frequency domain analysis may further illustrate the concepts of aliasing. Figure 28.26 shows the analog signal, the *sampling function* (represented by a unit impulse train) and the resulting sampled signal in both the time and frequency domains. The analog signal in Fig. 28.26a is represented as a simple band-limited signal with center frequency, f_o. This simply means that the signal is contained within the frequency range shown. In Fig. 28.26b, the sampling function is shown in both the time and frequency domain. The sampling function simply represents the action of sampling at discrete points in time. The frequency domain version of the sampling function is similar to its time domain counterpart, except that the x-axis is now represented as $f = 1/T$. Since each of the impulses has a value of 1, the resulting sampled signal shown in Fig. 28.26c is the impulse function multiplied by the amplitude of the analog signal at each discrete point in time. Remembering that multiplication in the time domain is equivalent to convolution in the frequency domain, we note that the frequency domain representation of the sampled signal reveals that the overall signal consists of multiple versions of the band-limited signal at multiples of the sampling frequency.

Note in Fig. 28.26b that as the sampling time increases, the sampling frequency decreases and the impulses in the frequency domain become more closely spaced. This results in 28.26d, which illustrates the aliasing as the multiple versions of the

IPR2023-00698
Theta EX2019

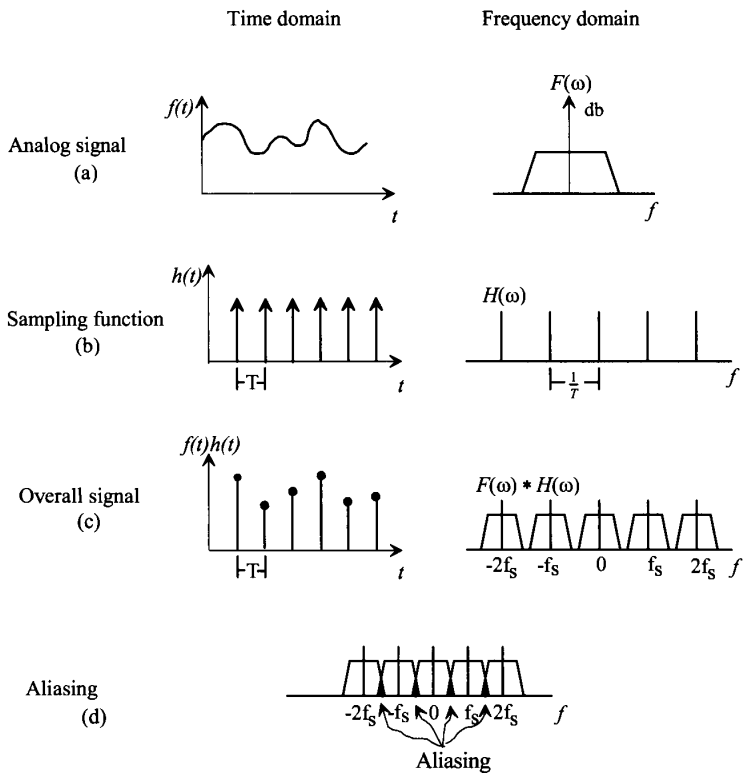


Figure 28.26 Illustration of aliasing in the time and frequency domain. (a) The analog signal; (b) the sampling function; (c) the overall signal; and (d) aliasing in the frequency domain.

band-limited signal begin to overlap. One could also filter the signal "post-sampling" and eliminate the frequencies for which overlap occurs. The point at which the spectra overlap is called the *folding frequency*.

As mentioned earlier, the solutions to aliasing are higher sampling frequency and filtering. Focusing on just one of the solutions may worsen the situation for several reasons. Some noise signals are wide band, which means that they have a large bandwidth. Attempting to increase only the sampling frequency to eliminate the aliasing effects of the noise would be a practical impossibility, not to mention a costly one. However, simply filtering the input signal and the sampled signal adds delays to the overall conversion and increases the expense of the circuit. It is best to use a combination of the two to minimize the problem most efficiently.

IPR2023-00698
Theta EX2019

Signal-to-Noise Ratio

Signal-to-noise (*SNR*) ratios of ADCs represent the value of the largest RMS input signal into the converter over the RMS value of the noise. Typically given in dB, the expression for *SNR* is

$$SNR = 20\text{Log}(\frac{v_{in(max)}}{v_{noise}}) \tag{28.23}$$

If it is assumed that the input signal is a sinewave with a peak-to-peak value equal to the full-scale reference voltage of the converter, then the RMS value for $v_{in(max)}$ becomes

$$v_{in(max)} = \frac{V_{REF}}{2\sqrt{2}} = \frac{2^N(V_{LSB})}{2\sqrt{2}} \tag{28.24}$$

where V_{LSB} is the voltage value of 1 LSB. The value of the noise (if the data converter is considered to be ideal) will be equivalent to the RMS value of the error signal, Q_e (in volts), shown in Fig. 28.20b. The RMS value of Q_e can be calculated to be

$$Q_{e,RMS} = \left[\frac{1}{V_{LSB}} \int_{-0.5V_{LSB}}^{0.5V_{LSB}} (V_{LSB})^2 dV_{LSB}\right]^{0.5} = \frac{V_{LSB}}{\sqrt{12}} \tag{28.25}$$

Therefore, the *SNR* for the ideal ADC will be the ratio of these two RMS values,

$$SNR = 20 \cdot \text{Log} \frac{\frac{2^N(V_{LSB})}{2\sqrt{2}}}{Q_{e,RMS}} \tag{28.26}$$

which can be written in terms of N as simply

$$SNR = 20N\text{Log}(2) + 20\text{Log}\sqrt{12} - 20\text{Log}(2\sqrt{2}) = 6.02N + 1.76 \tag{28.27}$$

Equation (28.27) is an important one relating *SNR* to the resolution of the ADC. For 16-bit data conversion, one must design a circuit that will have an *SNR* of (6.02)(16) + 1.76 = 98.08 dB! Equation (28.27) can also be used in calculating the *signal-to-noise plus distortion ratio*, also known as *SNDR*. Since the output data is digital, we cannot use a spectrum analyzer to calculate this ratio but must instead use a *Discrete Fourier Transform* (*DFT*) and examine the data in the digital domain.

Another useful application of Eq. (28.27) is the determination of effective number of bits given a system with a known *SNR* or *SNDR*. For example, if a 16-bit ADC yielded an SNDR of 88 dB, then the effective resolution of the converter would be

$$N = \frac{88 - 1.76}{6.02} = 14.32 \text{ bits} \tag{28.28}$$

and the ADC would be producing the resolution equivalent to that of a 14-bit converter.

Aperture Error

The aperture error described in Sec. 28.3 (S/H) should be related to the errors associated with the ADC. In the previous discussion, the aperture error resulted in sampling error (Fig. 28.8). However, now that ADC characteristics have been discussed, we can relate the sampling error to the ADC. Since we know that the maximum errors associated with an ADC are related to ½ LSB, we can assume that the maximum sampling error associated with the aperture uncertainty can be no larger than ½ LSB.

IPR2023-00698
Theta EX2019

Example 28.8
Find the maximum resolution of an ADC which can use the S/H described in Ex. 28.1 while maintaining a sampling error less than ½ LSB.

Since it was determined that the maximum sampling error produced by the given aperture uncertainty was 0.628 mV, we can relate this value to the highest resolution of an ADC by assuming that 0.628 mV will be less than or equal ½ LSB. Therefore,

$$0.628 \text{ mV} \leq .5\ LSB = \frac{V_{REF}}{2^{N+1}} = \frac{5}{2^{N+1}}$$

or

$$2^{N+1} \leq 7961.8$$

which, solving for N (limited to an integer), yields a maximum resolution of 11 bits. ■

28.6 Mixed-Signal Layout Issues

Naturally, analog ICs are more sensitive to noise than digital ICs. For any analog design to be successful, careful attention must be paid to layout issues, particularly in a digital environment. Sensitive analog nodes must be protected and shielded from any potential noise sources. Grounding and power supply routing must also be considered when using digital and analog circuitry on the same substrate. Since a majority of ADCs use switches controlled by digital signals, separate routing channels must be provided for each type of signal.

Techniques used to increase the success of mixed-signal designs vary in complexity and priority. Strategies regarding the systemwide minimization of noise should always be considered foremost. A mixed-signal layout strategy can be modeled as seen in Fig. 28.27. The lowest issues are foundational and must be considered before each succeeding step. The successful mixed-signal design will always minimize the effect of the digital switching on the analog circuits.

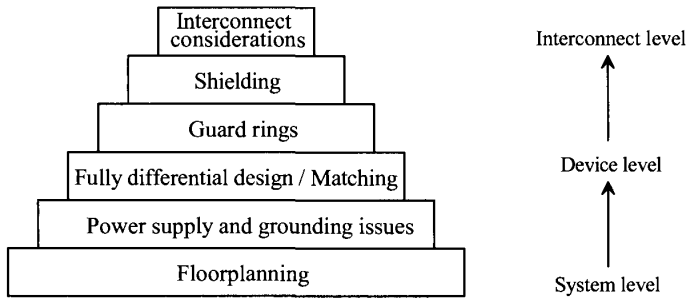


Figure 28.27 Mixed-signal layout strategy.

IPR2023-00698
Theta EX2019

Floorplanning

The placement of sensitive analog components can greatly affect a circuit's performance. Many issues must be considered. In designing a mixed-signal system, strategies regarding the "floorplan" of the circuitry should be thoroughly analyzed well before the layout is to begin.

The analog circuitry should be categorized by the sensitivity of the analog signal to noise. For example, low-level signals or high-impedance nodes typically associated with input signals are considered to be sensitive nodes. These signals should be closely guarded and shielded, especially from digital output buffers. High-swing analog circuits such as comparators and output buffer amplifiers should be placed between the sensitive analog and the digital circuitry.

The digital circuitry should also be categorized by speed and function. Obviously, since digital output buffers are usually designed to drive capacitive loads at very high rates, they should be kept farthest from the sensitive analog signals. Next, the high and lower speed digital should be placed between the insensitive analog and the output buffers. An example of this type of strategy can be seen in Fig. 28.28 [1]. Notice that *the sensitive analog is as far away as possible from the digital output buffers* and that the least sensitive analog circuitry is next to the least offensive digital circuitry.

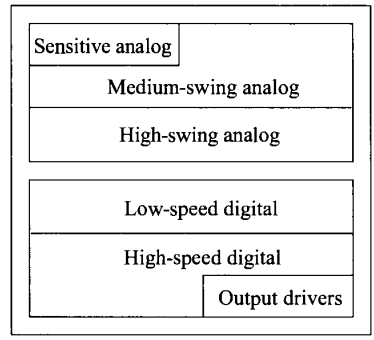


Figure 28.28 Example of a mixed-signal floorplan [1].

Power Supply and Grounding Issues

When analog and digital circuits exist together on the same die, danger exists of injecting noise from the digital system to the sensitive analog circuitry through the power supply and ground connections. Much of the intercoupling can be minimized by carefully considering how power and ground are supplied to both the analog and digital circuits.

In Fig. 28.28a analog and digital circuitry share the same routing to a single pad for power and ground. The resistors, R_{I1} and R_{I2}, represent the small, nonnegligible resistance of the interconnect to the pad. The inductors, L_{B1} and L_{B2}, represent the inductance of the bonding wire which connects the pads to the pin on the lead frame.

IPR2023-00698
Theta EX2019

Since digital circuitry is typified by high amounts of transient currents due to switching, a small amount of resistance associated with the interconnect can result in significant voltage spikes. Low-level analog signals are very sensitive to such interference, thus resulting in a contaminated analog system. Another significant voltage spike can occur due to the inductance of the bonding wire. Since the voltage across the inductor is proportional to the change in current through it, voltage spikes equating to hundreds of millivolts can result! Both of these voltage effects are true for both the power and ground connection.

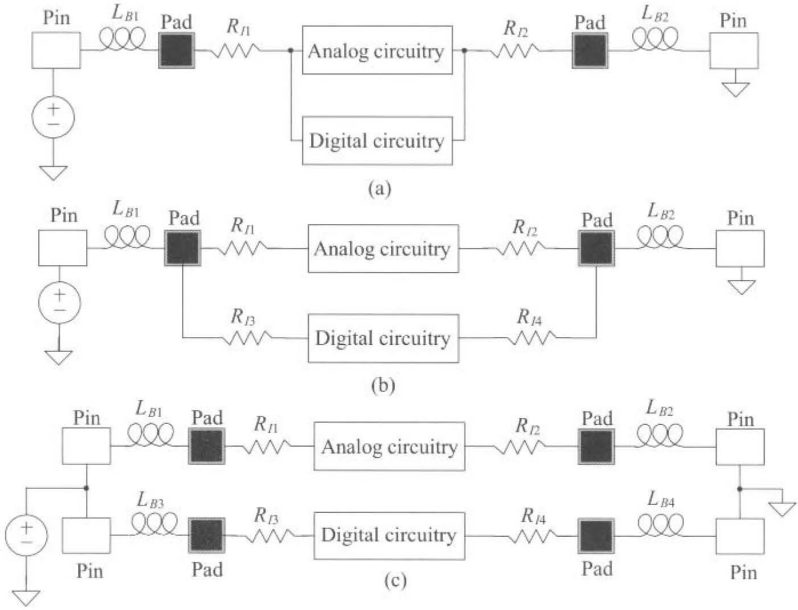


Figure 28.29 Power and ground connection examples, (a) poor noise immunity and (b) better noise immunity, and (c) using separate power and ground pins to achieve even better immunity.

One way to reduce the interference, seen in Fig. 28.29b, is to prohibit the analog and digital circuit from sharing the same interconnect. The routing for the supply and ground for both the analog and digital sections are provided separately. Although this eliminates the parasitic resistance due to the common interconnect, there is still a common inductance due to the bonding wire which causes interference.

Another method that minimizes interference even more than the previous case is seen in Fig. 28.29c. *By using separate pads and pins, the analog and the digital circuits are completely decoupled.* The current through the analog interconnect is much less abrupt than the digital; thus, the analog circuitry now has a "quiet" power and ground. However, this technique depends on whether extra pins and pads are available for this

IPR2023-00698
Theta EX2019

use. The separate power supply and ground pins are then connected externally. *It is not wise to use two separate power supplies because if both types of circuits are not powered up simultaneously, latch-up could easily result.*

In cases presented in Fig. 28.29b and c, *the resistance associated with the analog connection to ground or supply can be reduced by making the power supply and ground bus as wide as feasible.* This reduces the overall resistance of the metal run, thus decreasing the voltage spikes that occur across the resistor. The inductor itself is impossible to eliminate, though it can be minimized with careful planning. Since the length of the bonding wire depends on the distance from the pad to the lead frame, *one could reduce the effect of the wire inductance by reserving pins closest to the die for sensitive connections such as analog supply and ground.* This, again, illustrates the importance of floorplanning.

Fully Differential Design

Fully differential operational amplifiers were discussed earlier, Fig. 28.30. The noise sources represent the noise from digital circuitry coupled through the parasitic stray capacitors. If equal amounts of noise are injected into the differential amplifiers, then the common-mode rejection inherent in the amplifiers will eliminate most or all of the noise. This, of course, depends on the symmetry of the amplifiers, meaning that matching the transistors in the amplifier becomes crucial. Therefore, *in a mixed-signal environment, layout techniques should be used to improve matching such as common-centroid and interdigitated techniques.*

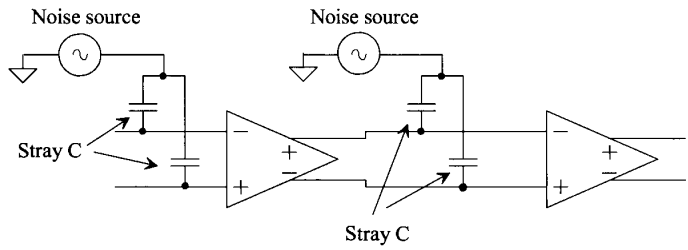


Figure 28.30 Differential output op-amps showing parasitic coupling to noise sources.

Guard Rings

Guard rings should be used wisely throughout a mixed-signal environment. Circuits that process sensitive signals should be placed in a separate well (if possible) with guard rings attached to the analog *VDD* supply. In the case of an n-well (only) process, the n-type devices outside the well should have guard rings attached to analog ground placed around them. Digital circuits should be placed in their own well with guard rings attached to digital *VDD*. Guard rings placed around the n-channel digital devices also help minimize the amount of noise transmitted from the digital devices.

IPR2023-00698
Theta EX2019

Shielding

A number of techniques exist which can shield sensitive, low-level analog signals from noise resulting from digital switching. A shield can take the form of a layer tied to analog ground placed between two other layers, or it can be a barrier between two signals running in parallel.

If at all possible, one should avoid crossing sensitive analog signals, such as low-level analog input signals, with any digital signals. The parasitic capacitance coupling the two signal lines can be as much as a couple of fF, depending on the process. If it cannot be avoided, then attempt to carry the digital signal using the top layer of metal (such as metal2). If the analog signal is an input signal, then it will most likely be carried by the poly layer (or a lower level of metal). A strip of metal1 can be placed between the two layers and connected to analog ground (see Fig. 28.31).

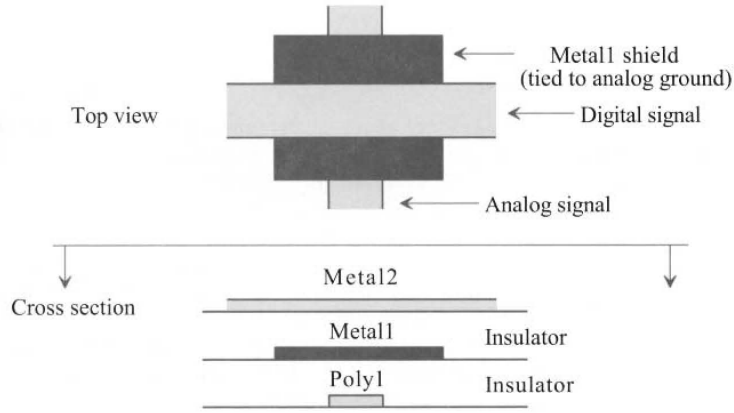


Figure 28.31 Shielding a sensitive analog signal from a digital signal crossover using a metal 1 shield layer.

Another situation that should be avoided is running an interconnect containing sensitive analog signals parallel and adjacent to any interconnect carrying digital signals. Coupling occurs due to the parasitic capacitance between the lines. If this situation cannot be avoided, then an additional line connected to analog ground should be placed between the two signals, as seen in Fig. 28.32. This method can also be used to partition the analog and digital sections of the chip.

In addition, the n-well can be used as a bottom-plate shield to protect analog signals from substrate noise. Poly resistors (or capacitors) used for sensitive analog signals can be shielded by placing an n-well beneath the components and connecting the well to analog *VDD*.

IPR2023-00698
Theta EX2019

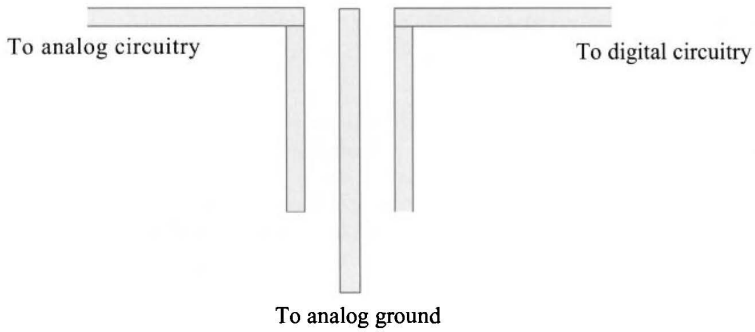


Figure 28.32 Using a dummy metal strip to provide shielding to two parallel signals.

Other Interconnect Considerations

Finally, some other layout strategies will incrementally improve the performance of the analog circuitry. However, if the previous strategies are not followed, these suggestions will be useless. *When routing the analog circuitry, minimize the lengths of current carrying paths.* This will simply reduce the amount of voltage drop across the path due to the metal1 or metal2 resistance. *Vias and contacts should also be used very liberally whenever changing layers.* Not only does this minimize resistance in the path, but it also improves fabrication reliablity. *Avoid using poly to route current carrying signal paths.* Not only is the poly higher in resistance value, but also the additional contact resistance required to change layers will not be insignificant. If the poly is made wider to lower the resistance, additional parasitic capacitance will be added to the node. *Use poly to route only high-impedance gate nodes that carry virtually no current.*

ADDITIONAL READING

[1] Y. Tsividis, *Mixed Analog-Digital VLSI Devices and Technology: An Introduction*, McGraw-Hill Publishing Co., 1996.

[2] B. Razavi, *Principles of Data Conversion System Design*, IEEE Press, 1995.

[3] M. J. Demler, *High-Speed Analog-to-Digital Conversion*, Academic Press, 1991.

[4] R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, McGraw-Hill Publishing Co., 1990.

[5] D. H. Sheingold, *Analog-Digital Conversion Handbook*, Prentice-Hall Publishing, 1986.

[6] S. K. Tewksbury, et al., "Terminology Related to the Performance of S/H, A/D and D/A Circuits," *IEEE Transactions on Circuits and Systems*, CAS-25, vol. CAS-25, pp. 419–426, July 1978.

IPR2023-00698
Theta EX2019

PROBLEMS

28.1 Determine the number of quantization levels needed if one wanted to make a digital thermometer that was capable of measuring temperatures to within 0.1 °C accuracy over a range from –50 °C to 150 °C. What resolution of ADC would be required?

28.2 Using the same thermometer as above, what sampling rate, in samples per second, would be required if the temperature displayed a frequency of $15°\cdot\sin(0.01\cdot 2\pi t)$?

28.3 Determine the maximum droop allowed in an S/H used in a 16-bit ADC assuming that all other aspects of both the S/H and ADC are ideal. Assume $V_{ref} = 5$ V.

28.4 An S/H circuit settles to within 1 percent of its final value at 5 μs. What is the maximum resolution and speed with which an ADC can use this data assuming that the ADC is ideal?

28.5 A digitally programmable signal generator uses a 14-bit DAC with a 10-volt reference to generate a DC output voltage. What is the smallest incremental change at the output that can occur? What is the DAC's full-scale value? What is its accuracy?

28.6 Determine the maximum DNL (in LSBs) for a 3-bit DAC, which has the following characteristics. Does the DAC have 3-bit accuracy? If not, what is the resolution of the DAC having this characteristic?

Digital Input	Voltage Output
000	0 V
001	0.625 V
010	1.5625 V
011	2.0 V
100	2.5 V
101	3.125 V
110	3.4375 V
111	4.375 V

28.7 Repeat Problem 28.6 calculating the INL (in LSBs).

28.8 A DAC has a reference voltage of 1,000 V, and its maximum INL measures 2.5 mV. What is the maximum resolution of the converter assuming that all the other characteristics of the converter are ideal?

28.9 Determine the INL and DNL for a DAC that has a transfer curve shown in Fig. 28.33.

28.10 A DAC has a full-scale voltage of 4.97 V using a 5 V reference, and its minimum output voltage is limited by the value of one LSB. Determine the resolution and dynamic range of the converter.

IPR2023-00698
Theta EX2019

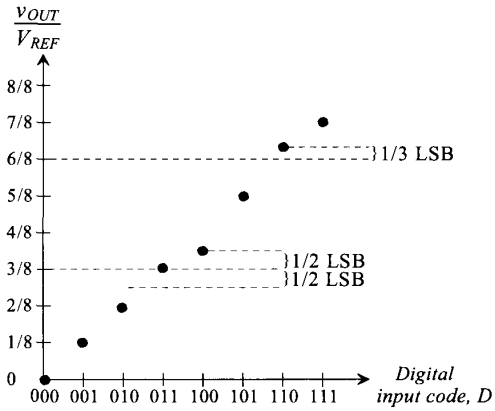


Figure 28.33 Transfer curves for Problem 28.9.

28.11 Prove that the RMS value of the quantization noise shown in Fig. 28.20b is as stated in Eq. (28.25).

28.12 An ADC has a stated *SNR* of 94 dB. Determine the effective number of bits of resolution of the converter.

28.13 Discuss the methods used to prevent aliasing and the advantages and disadvantages of each.

IPR2023-00698
Theta EX2019

Chapter
29

Data Converter Architectures

Applications such as wireless communications and digital audio and video have created the need for cost-effective data converters that will achieve higher speed and resolution. The needs required by digital signal processors continually challenge analog designers to improve and develop new ADC and DAC architectures. There are many different types of architectures, each with unique characteristics and different limitations. This chapter presents a basic overview of the more popular data converter architectures and discusses the advantages and disadvantages of each along with their limitations.

Now that we have defined the operating characteristics of ADCs in Ch. 28, a more detailed examination of the basic architectures will be discussed using a top-down approach. Because many of the converters use op-amps, comparators, and resistor and capacitor arrays, the top-down approach will allow a broader discussion of the key component limitations in later sections.

29.1 DAC Architectures

A wide variety of DAC architectures exist, ranging from very simple to complex. Each, of course, has its own merits. Some use voltage division, whereas others employ current steering and even charge scaling to map the digital value into an analog quantity.

29.1.1 Digital Input Code

In many cases, the digital signal is not provided in binary code but is any one of a number of codes: binary, BCD, thermometer code, Gray code, sign-magnitude, two's complement, offset binary, and so on. (See Fig. 29.1 for a comparison of some of the more commonly used digital input codes.) For example, it may be desirable to allow only one bit to change value when changing from one code to the next. If that is the case, a Gray code will suffice. The thermometer code is used quite frequently and can also be seen in Fig. 29.1. Notice that it requires $2^N - 1$ bits to represent an N-bit word. The choice of code depends on the application, and the reader should be aware that many types of codes are available.

IPR2023-00698
Theta EX2019

Decimal	Binary	Thermometer	Gray	Two's Complement
0	000	0000000	000	000
1	001	0000001	001	111
2	010	0000011	011	110
3	011	0000111	010	101
4	100	0001111	110	100
5	101	0011111	111	011
6	110	0111111	101	010
7	111	1111111	100	001

Figure 29.1 Comparison of digital input codes.

29.1.2 Resistor String

The most basic DAC is seen in Fig. 29.2a. Comprised of a simple resistor string of 2^N identical resistors and switches, the analog output is simply the voltage division of the resistors at the selected tap. Note that a N:2^N decoder is required to provide the 2^N signals controlling the switches. This architecture typically results in good accuracy, provided that no output current is required and that the values of the resistors are within the specified error tolerance of the converter. One big advantage of a resistor string is that the output is always guaranteed to be monotonic.

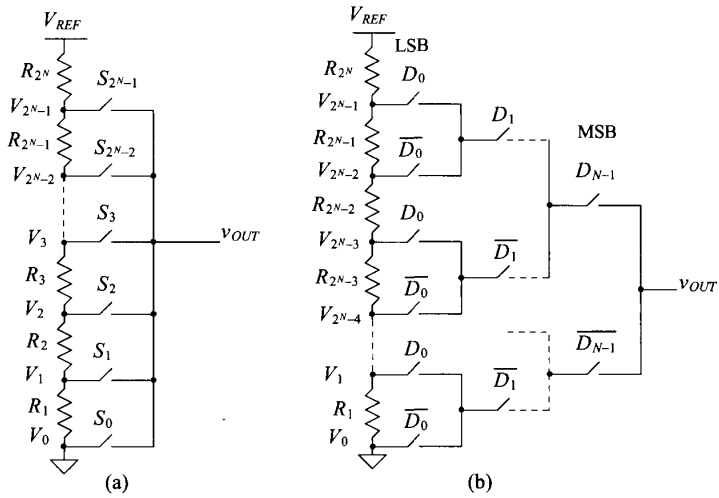


Figure 29.2 (a) A simple resistor-string DAC and (b) the use of a binary switch array to lower the output capacitance.

IPR2023-00698
Theta EX2019

One problem with this converter is that the converter output is always connected to $2^N - 1$ switches that are off and one switch that is on. For larger resolutions, a large parasitic capacitance appears at the output node, resulting in slower conversion speeds. A better alternative for the resistor-string DAC is seen in Fig. 29.2b. Here, a binary switch array ensures that the output is connected to at most N switches that are on and N switches that are off, thus increasing conversion speed. The input to this switch array is a binary word since the decoding is inherent in the binary tree arrangement of the switches.

Another problem with the resistor-string DAC is the balance between area and power dissipation. An integrated version of this converter leads to a large chip area for higher bit resolutions because of the large number of passive components needed. Active resistors such as the n-well resistor can be used for low-resolution applications. However, as the resolution increases, the relative accuracy of the resistors becomes an important factor. Although the value of R could always be made small to minimize the chip area required, power dissipation would then become the critical issue as current flows through the resistor string at all times.

Example 29.1

Design a 3-bit resistor-string ladder using a binary switch array. Assume that V_{REF} = 5 V and that the maximum power dissipation of the converter is to be 5 mW (not including the power required by the digital logic). Determine the value of the analog voltage for each of the possible digital input codes.

The power dissipation will determine the current flowing through the resistor string by

$$I_{MAX} = \frac{5 \times 10^{-3}\ \text{W}}{5\ \text{V}} = 1\ \text{mA}$$

Since a 3-bit converter will have eight resistors, the value of R is

$$R = \frac{1}{8} \cdot \frac{5\ \text{V}}{1\ \text{mA}} = 625\ \Omega$$

The converter can be seen in Fig. 29.3. Examine the switch array if the input code is $D_2D_1D_0$ = 100 or 4_{10}. Since D_2 is high, the top switch will be closed and the lower switch, $\overline{D_2}$, will be open. In the row corresponding to D_1, since $D_1 = 0$, both of the switches marked $\overline{D_1}$ will be closed and the other two will be open. The LSB controls the largest number of switches; therefore, since D_0 is low, all of the $\overline{D_0}$ switches will be closed and all of the D_0 switches will be open. There should be only one path connecting a single tap on the resistor string to the output. This is bolded, with the resistor string tapped in the middle of the string. Therefore, v_{OUT} = ½ V_{REF} = 2.5 V. The remaining outputs can be seen in Fig. 29.4. ■

Mismatch Errors Related to the Resistor-String DAC

The accuracy of the resistor string is obviously related to matching between the resistors, which ultimately determines the INL and DNL for the entire DAC. Suppose that the i-th resistor, R_i, has a mismatch error associated with it so that

$$R_i = R + \Delta R_i \tag{29.1}$$

IPR2023-00698
Theta EX2019

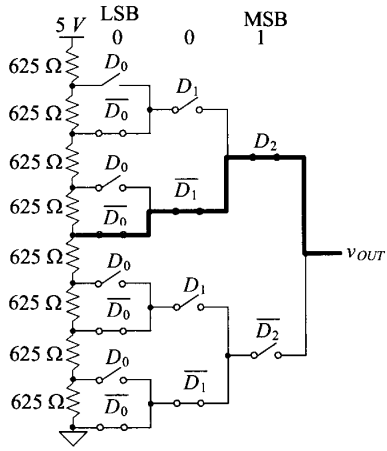


Figure 29.3 A 3-bit resistor-string DAC used in Ex. 29.1

where R is the ideal value of the resistor and ΔR_i is the mismatch error. Also suppose that the mismatches were symmetrical about the string so that the sum of all the mismatch terms were zero, or

$$\sum_{i=1}^{2^N} \Delta R_i = 0 \tag{29.2}$$

The value of the voltage at the tap associated with the i-th resistor should ideally be

$$V_{i,ideal} = \frac{(i)V_{REF}}{2^N}, \text{ for } i = 0, 1, 2, ..., 2^N - 1 \tag{29.3}$$

$D_2 D_1 D_0$	v_{OUT}
000	0
001	0.625
010	1.25
011	1.875
100	2.5
101	3.125
110	3.75
111	4.375

Figure 29.4 Output voltages generated from the 3-bit DAC in Ex. 29.1.

IPR2023-00698
Theta EX2019

However, including the mismatch, the actual value of the i-th voltage is the sum of all of the resistances up to and including resistor i, divided by the sum of all of the resistances in the string. This can be represented by

$$V_i = V_{REF} \cdot \frac{\sum_{k=1}^{i} R_k}{\sum_{k=1}^{2^N} R_k} = V_{REF} \cdot \frac{\sum_{k=1}^{i} R + \Delta R_k}{2^N R} \tag{29.4}$$

The denominator does not include any mismatch error since it was assumed that the mismatches sum to zero as defined in Eq. (29.2). Notice that there is no resistor, R_0, corresponding to V_0 (see Fig. 29.2), and it is assumed that V_0 is ground. Equation (29.4) can be rewritten as

$$V_i = \frac{V_{REF}}{2^N R}\left[(i)R + \sum_{k=1}^{i} \Delta R_k\right] = \frac{(i)V_{REF}}{2^N} + \frac{V_{REF}}{2^N R}\sum_{k=1}^{i} \Delta R_k \tag{29.5}$$

or finally, the value of the voltage at the i-th tap is

$$V_i = V_{i,ideal} + \frac{V_{REF}}{2^N} \cdot \sum_{k=1}^{i} \frac{\Delta R_k}{R} \tag{29.6}$$

Equation (29.6) is not of much importance by itself, but it can be used to help determine the nonlinearity errors.

Integral Nonlinearity of the Resistor-String DAC

Integral nonlinearity (INL) is defined as the difference between the actual and ideal switching points, or

$$INL = V_i - V_{i,ideal} \tag{29.7}$$

and plugging in Eqs. (29.6) and (29.3) into (29.7) yields,

$$INL = \frac{V_{REF}}{2^N} \cdot \sum_{k=1}^{i} \frac{\Delta R_k}{R} \tag{29.8}$$

Equation (29.8) is a general expression for the INL for a given resistor, R_i, and requires that the mismatch of all the resistances used in the summation are known. However, this equation does not illuminate how to determine the *worst-case* or maximum INL for a resistor string.

Intuitively, one would think that the worst-case INL would occur at the top of the resistor string (i=2^N) with all of the ΔR_k's at their maximum values. However, the previous derivation was performed with the assumption that the mismatches summed to zero. With this restriction, the maximum INL occurs at the midpoint of the string where i = 2^{N-1}, corresponding to the case where the MSB was a one and all other bits were zero. Another condition that will ensure a worst-case scenario is to consider the lower half resistors at their maximum positive mismatch value and the upper half resistors at their maximum negative mismatch value, or vice versa.

If the resistors on a string were known to have 2% matching, then ΔR_k would be constrained to the bounds of

$$-0.02R \le \Delta R_k \le 0.02R \tag{29.9}$$

IPR2023-00698
Theta EX2019

and the worst-case INL (again, % matching = 0.02) using Eq. (29.8) would be,

$$|INL|_{max} = \frac{V_{REF}}{2^N} \cdot \sum_{k=1}^{2^{N-1}} \frac{\Delta R_k}{R} = \frac{V_{REF}}{2^N} \cdot \frac{2^{N-1} \cdot \Delta R_k}{R} = \frac{1}{2} LSB \cdot 2^N \cdot (\%\ \text{matching}) = 0.01 V_{REF} \tag{29.10}$$

which for $INL < 0.5\ LSB$ requires $1/2^N >$ (% matching). For 2% matching, the maximum number of bits, N, is then 5! For better than 0.2% matching $N = 9$ bits.

Because the worst-case analysis was performed, the maximum INL occurs at the middle of the string. We can improve this specification on paper by using the "best-fit" approach to measuring INL. In this case, the reference line is simply shifted up slightly (refer to Ch. 28) so that it no longer passes through the end points, but instead minimizes the INL.

Example 29.2

Determine the effective number of bits for a resistor-string DAC, which is assumed to be limited by the INL. The resistors are passive poly resistors with a known relative matching of 1%, and $V_{REF} = 5$ V.

Using Eq. (29.10), the maximum INL will be

$$|INL|_{max} = 0.005 \cdot V_{REF} = 0.025 \text{ V}$$

Since we know that this maximum INL should be equal to ½ LSB in the worst case,

$$\frac{1}{2} LSB = \frac{5}{2^{N+1}} = 0.025 \text{ V}$$

and solving for N yields

$$N = \log_2\left(\frac{5}{0.025}\right) - 1 = 6.64 \text{ bits}$$

This means that the resolution for a DAC containing a resistor string matched to within 1% will be, at most 6 bits. ■

Differential Nonlinearity of the Worst-Case Resistor-String DAC

Resistor-string matching is not as critical when determining the DNL. Remembering that the definition of DNL is simply the actual height of the stair-step in the DAC transfer curve minus the ideal step height, we can write this in terms of the voltages at the taps of adjacent resistors on the string. Using Eq. (29.5), we can express this as,

$$|V_i - V_{i-1}| = \left| \left[\frac{(i)V_{REF}}{2^N} + \frac{V_{REF}}{2^N} \cdot \sum_{k=1}^{i} \frac{\Delta R_k}{R} \right] - \left[\frac{(i-1)V_{REF}}{2^N} + \frac{V_{REF}}{2^N} \cdot \sum_{k=1}^{i-1} \frac{\Delta R_k}{R} \right] \right|$$

which can be simplified to

$$|V_i - V_{i-1}| = \left| \frac{V_{REF}}{2^N} \left(1 + \frac{\Delta R_i}{R}\right) \right| \tag{29.11}$$

The DNL can then be determined by subtracting the ideal step height from Eq. (29.11),

$$DNL_i = \left| \frac{V_{REF}}{2^N} \left(1 + \frac{\Delta R_i}{R}\right) - \frac{V_{REF}}{2^N} \right| = \left| \frac{V_{REF}}{2^N} \cdot \frac{\Delta R_i}{R} \right| \tag{29.12}$$

IPR2023-00698
Theta EX2019

and the maximum DNL will occur at the value of i for which ΔR is at its maximum value. If it is assumed once again that the resistors are matched to within 2 percent, the worst-case DNL will be

$$DNL_{max} = \left| 0.02 \cdot \frac{V_{REF}}{2^N} \right| = 0.02\ LSB \tag{29.13}$$

which is well below the ½ LSB limit. The INL is obviously the limiting factor in determining the resolution of a resistor-string DAC as its maximum value is 2^N times larger than the DNL.

29.1.3 *R-2R* Ladder Networks

Another DAC architecture that incorporates fewer resistors is called the *R-2R* ladder network. This configuration consists of a network of resistors alternating in value of R and $2R$. Figure 29.5 illustrates an N-bit *R-2R* ladder. Starting at the right end of the network, notice that the resistance looking to the right of any node to ground is $2R$. The digital input determines whether each resistor is switched to ground (noninverting input) or to the inverting input of the op-amp. Each node voltage is related to V_{REF}, by a binary-weighted relationship caused by the voltage division of the ladder network. The total current flowing from V_{REF} is constant, since the potential at the bottom of each switched resistor is always zero volts (either ground or virtual ground). Therefore, the node voltages remain constant for any value of the digital input.

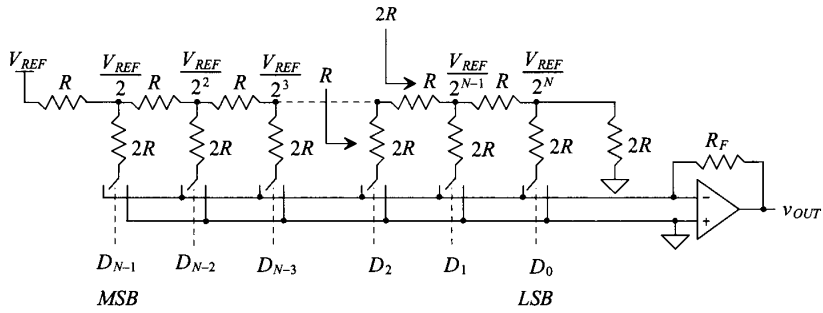


Figure 29.5 An *R-2R* digital-to-analog converter.

The output voltage, v_{OUT}, depends on currents flowing through the feedback resistor, R_F, such that

$$v_{OUT} = -i_{TOT} \cdot R_F \tag{29.14}$$

where i_{TOT} is the sum of the currents selected by the digital input by

$$i_{TOT} = \sum_{k=0}^{N-1} D_k \cdot \frac{V_{REF}}{2^{N-k}} \cdot \frac{1}{2R} \tag{29.15}$$

where D_k is the k-th bit of the input word with a value that is either a 1 or a 0.

IPR2023-00698
Theta EX2019

This architecture, like the resistor-string architecture, requires matching to within the resolution of the converter. Therefore, the switch resistance must be negligible, or a small voltage drop will occur across each switch, resulting in an error. One way to eliminate this problem is to add dummy switches. Assume that the resistance of each switch connected to the $2R$ resistors is ΔR, as seen in Fig. 29.6. Dummy switches with one-half the resistance of the real switches are "hard-wired" so that they are always on and placed in series with each of the horizontal resistors. The total resistance of any horizontal branch, R', is

$$R' = R + \frac{\Delta R}{2} \tag{29.16}$$

The resistance of any vertical branch is $2R + \Delta R$, which is twice the value of the horizontal branch. Therefore, a $R' - 2R'$ relationship is maintained. Of course, a dummy switch equal to the switch size of a $2R$ switch will have to be placed in series with the terminating resistor as well.

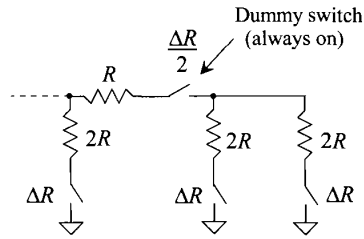


Figure 29.6 Use of dummy switches to offset switch resistance.

Example 29.3

Design a 3-bit DAC using an R-$2R$ architecture with R = 1 kΩ, R_F = 2 kΩ, and V_{REF} = 5 V. Assume that the resistances of the switches are negligible. Determine the value of i_{TOT} for each digital input and the corresponding output voltage, v_{OUT}.

Figure 29.7 shows the 3-bit DAC for a digital input of 001. The voltages at each node in the resistor network are labeled. For each switch, if the digital input bit is a zero, then the resistor is attached to the ground. If the bit is a one, then the resistor is attached to the virtual ground of the inverting input and current flows to the output of the op-amp. Therefore, for $D_2D_1D_0$ = 000, all of the switches are connected to ground, no current flows through the feedback resistor, and the output voltage, v_{OUT}, is zero.

When $D_2D_1D_0$ = 001, the rightmost resistor is switched to the op-amp inverting input and the other two resistors remain attached to ground. Therefore, the total current flowing through the feedback resistor is simply the current through the rightmost resistor, which is defined by Eq. (29.15) as

IPR2023-00698
Theta EX2019

$$\frac{V_{REF}}{8} \cdot \frac{1}{2000} = 0.3126 \text{ mA}$$

and the output voltage, by Eq. (29.14) becomes,

$$v_{OUT} = -(0.3126 \text{ mA})(2000\ \Omega) = -0.625 \text{ V}$$

which is to be expected. The other values for the output voltage can be calculated using Eqs. (29.14) and (29.15) and are seen in Fig. 29.8. ■

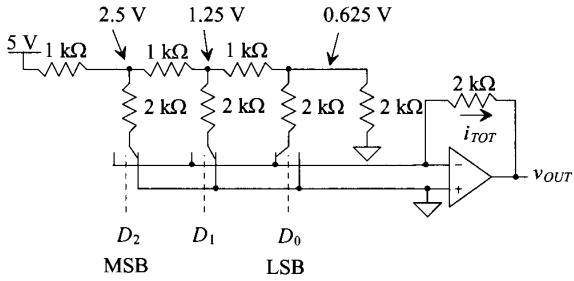


Figure 29.7 A 3-bit *R*-2*R* digital-to-analog converter used in Ex. 29.3.

$D_2D_1D_0$	i_{TOT} (mA)	v_{OUT} (V)
000	0	0
001	0.3125	– 0.625
010	0.625	– 1.25
011	0.625 + 0.3125 = 0.9375	– 1.875
100	1.25	– 2.5
101	1.25 + 0.3125 = 1.5625	– 3.125
110	1.25 + 0.625 = 1.875	– 3.75
111	1.25 + 0.625 + 0.3125 = 2.1875	– 4.375

Figure 29.8 Output voltages generated from the 3-bit DAC in Example 29.3.

29.1.4 Current Steering

In the previous section, a voltage was converted into a current, which then generated a voltage at the output. Another DAC method uses current throughout the conversion. Known as *current steering*, this type of DAC requires precision current sources that are summed in various fashions.

IPR2023-00698
Theta EX2019

Figure 29.9 illustrates a generic current-steering DAC. This configuration requires a set of current sources, each having a unit value of current, I. Since there are no current sources generating i_{OUT} when all the digital inputs are zero, the MSB, D_{2^N-2}, is offset by two index positions instead of one. For example, for a 3-bit converter, seven current sources will be needed, labeled from D_0 to D_6. The binary signal controls whether or not the current sources are connected to either i_{OUT} or some other summing node (in this case ground). The output current, i_{OUT}, has the range of

$$0 \le i_{OUT} \le (2^N - 1) \cdot I \tag{29.17}$$

and can be any integer multiple of I in between. An interesting issue to note is the format of the digital code required to drive the switches. Since there are $2^N - 1$ current sources, the digital input will be in the form of a *thermometer code*. This code will be all ones from the LSB up to the value of the k-th bit, D_k, and all zeros above it. The point at which the input code changes from all ones to all zeros "floats" up or down and resembles the action of a thermometer, hence the name. Typically, a thermometer encoder is used to convert binary input data into a thermometer code.

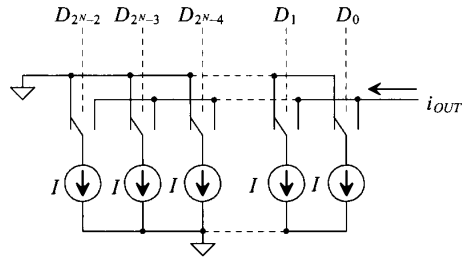


Figure 29.9 A generic current-steering DAC.

Another current-steering architecture is seen in Fig. 29.10. This architecture uses binary-weighted current sources, thus requiring only N current sources of various sizes versus $2^N - 1$ sources in the previous example. Since the current sources are binary weighted, the input code can be a simple binary number with no thermometer encoder needed.

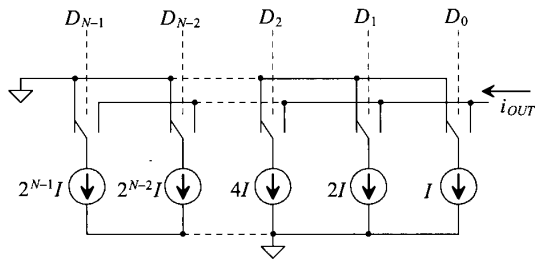


Figure 29.10 A current-steering DAC using binary-weighted current sources.

IPR2023-00698
Theta EX2019

One advantage of the current-steering DACs is the high-current drive inherent in the system. Since no output buffers are necessary to drive resistive loads, these DACs are typically used in high-speed applications. Traditionally, high-speed current-steering DACs have been fabricated using bipolar technology. However, the ability to generate matched current mirrors makes CMOS an enticing alternative. Of course, the precision needed to generate high resolutions depends on how well the current sources can be matched or the degree to which they can be made binary weighted. For example, if a 13-bit DAC was designed using these architectures, there would have to be 8,191 current sources resident on the chip, not an insignificant amount. For the binary-weighted sources, only 13 current sources would be needed. Yet the size of the largest current source would have to be 4,096 or 2^{N-1} times larger than the smallest. Even if the unit current, I, was chosen to be 5 μA, the largest current source would be 20.48 mA!

Another problem associated with this architecture is the error due to the switching. Since the current sources are in parallel, if one of the current sources is switched off and another is switched on, a "glitch" could occur in the output if the timing was such that both of them were on or both were off for an instant. While this may not seem significant, if the converter is switching from 0111111 to 10000000, the output will spike toward ground and then back to the correct value if all the switches turn off for an instant. If the DAC is driving a resistive load and the output current is converted to a voltage, a substantial voltage spike will occur at the output.

Example 29.4

Construct a table showing the thermometer code necessary to generate the output shown in Fig. 29.11a for a 3-bit current-steering DAC using unit current sources.

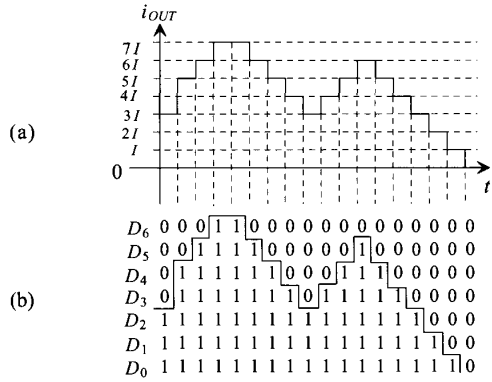


Figure 29.11 (a) Output of a 3-bit current-steering DAC and (b) the thermometer code input.

The thermometer code can be seen in Fig. 29.11b. When the code is all zeros, the output is 0 volts. Therefore, only 7 bits are needed to represent the 2^N or 8 states of a 3-bit DAC. Note how the interface between all ones and all zeros actually resembles the output signal itself. ■

IPR2023-00698
Theta EX2019

Mismatch Errors Related to Current-Steering DACs

Analysis of the mismatch associated with the current sources is similar to the resistor string analysis. It is assumed that each current source in Fig. 29.9 is

$$I_k = I + \Delta I_k \text{ for } k = 1, 2, 3, ..., 2^N - 1 \tag{29.18}$$

where I is the ideal value of the current and ΔI_k is the error due to mismatch. If it is again assumed that the ΔI_K terms sum to zero and that one-half of the current sources contain the maximum positive mismatch, ΔI_{max}, and the other half contains the maximum negative mismatch, $-\Delta I_{max}$, (or vice versa), then the worst-case condition will occur at midscale with the actual output current being

$$I_{out} = \sum_{k=1}^{2^{N-1}} (I + \Delta I_k) = 2^{N-1} \cdot I + 2^{N-1} \cdot |\Delta I|_{max} = I_{out,ideal} + 2^{N-1} \cdot |\Delta I|_{max} \tag{29.19}$$

Since the INL is simply the actual output current minus the ideal, the worst-case INL is

$$|INL|_{max} = 2^{N-1} \cdot |\Delta I|_{max,INL} \tag{29.20}$$

The term, $|\Delta I|_{max,INL}$ represents the maximum current source mismatch error that will keep the INL less than ½ LSB. Each current source represents the value of 1 LSB; therefore, ½ LSB is equal to 0.5 I. Because the maximum INL should correspond to the ½ LSB, equating Eq. (29.20) to ½ I results in the value for $|\Delta I|_{max,INL}$,

$$|\Delta I|_{max,INL} = \frac{0.5I}{2^{N-1}} = \frac{I}{2^N} \tag{29.21}$$

Equation (29.21) illustrates the difficulty of using this architecture at high resolutions. If the value of I is set to be 5 μA, and the N is desired to be 12 bits, then

$$|\Delta I|_{max,INL} = \frac{5 \times 10^{-6}}{2^{12}} = 1.221 \text{ nA!} \tag{29.22}$$

which means that each of the 5 μA current sources must lie between the bounds of,

$$4.99878\ \mu A \le I_k \le 5.001221\ \mu A \tag{29.23}$$

to achieve a worst-case INL, which is within ½ LSB error.

The DNL is easily obtained since the step height in the transfer curve is equivalent to the value of the ideal current source, I. The maximum difference between any two adjacent values of output current will simply be the value of the single source, I_k, which contains the largest mismatch error for which the DNL will be less than ½ LSB, $|\Delta I|_{max,DNL}$:

$$I_{out(x)} - I_{out(x-1)} = I_k + |\Delta I|_{max,DNL} \tag{29.24}$$

Therefore, the DNL is simply

$$|DNL|_{max} = I_k + |\Delta I|_{max,DNL} - I_k = |\Delta I|_{max,DNL} \tag{29.25}$$

Equating the maximum DNL to the value of ½ LSB,

$$|\Delta I|_{max,DNL} = \tfrac{1}{2} LSB = \tfrac{1}{2} I \tag{29.26}$$

which is much easier to attain than the requirement for the INL.

IPR2023-00698
Theta EX2019

For the binary-weighted current sources seen in Fig. 29.10, a slightly different analysis is needed to determine the requirements for INL and DNL. In this case, it will be assumed that the current source corresponding to the MSB (D_{N-1}) has a maximum positive mismatch error value and that the remainder of the bits (D_0 to D_{N-2}) contain a maximum negative mismatch error, so that the sum of all the errors equals zero. The INL is

$$|INL|_{max} = 2^{N-1}(I + |\Delta I|_{max,INL}) - 2^{N-1} \cdot I = 2^{N-1} \cdot |\Delta I|_{max,INL} \quad (29.27)$$

which is equivalent to the value of the current steering array in Fig. 29.9.

The DNL is slightly different because of the binary weighting of the current sources. One cannot add a single current source with each incremental increase in the digital input code. However, the worst-case condition for binary-weighted arrays tends to occur at midscale when the code transitions from 011111....111 to 100000....000. The worst-case DNL at this point is

$$DNL_{max} = \left[2^{N-1} \cdot (I + |\Delta I|_{max,DNL}) - \sum_{k=1}^{N-1} 2^{k-1} \cdot (I - |\Delta I|_{max,DNL}) \right] - I \quad (29.28)$$

which can be written as

$$DNL_{max} = 2^{N-1} \cdot (I + |\Delta I|_{max,DNL}) - (2^{N-1} - 1) \cdot (I - |\Delta I|_{max,DNL}) - I = (2^N - 1) \cdot |\Delta I|_{max,DNL} \quad (29.29)$$

and setting this value equal to ½ LSB and solving for ΔI_{max},

$$|\Delta I|_{max,DNL} = \frac{0.5I}{2^N - 1} = \frac{I}{2^{N+1} - 2} \quad (29.30)$$

Therefore, the DNL requirements for the binary-weighted current source array is more stringent than the INL requirements.

One interesting issue regarding the previous derivation is that the challenging accuracy requirements in Eq. (29.30) are placed only on the MSB current source. For each of the remaining binary-weighted sources, the DNL requirements become more relaxed. This is simply because the size of the MSB source is equivalent to all of the other sources combined, and so its value plays the most important role in the DAC's accuracy.

Example 29.5

Determine the tolerance of the MSB current source on a 10-bit binary-weighted current source array with a unit current source of 1 μA, which will result in a worst-case DNL that is less than ½ LSB.

Since Eq. (29.30) defines the maximum $|\Delta I|$ needed to keep the DNL less than ½ LSB, we must first use this equation,

$$|\Delta I|_{max,DNL} = \frac{1 \times 10^{-6}}{2^{11} - 2} = 0.4888 \text{ nA}$$

For a 10-bit DAC, the MSB current source will have a value that is 2^9 times larger than the unit current source, or 0.512 mA. Therefore, the range of values for which this array will have a DNL that is less than ½ LSB is

$$0.51199995 \text{ mA} \leq I_{MSB} \leq 0.5120004888 \text{ mA} \blacksquare$$

IPR2023-00698
Theta EX2019

29.1.5 Charge-Scaling DACs

A very popular DAC architecture used in CMOS technology is the charge-scaling DAC. Shown in Fig. 29.12a, a parallel array of binary-weighted capacitors, totaling $2^N C$, is connected to an op-amp. The value, C, is a unit capacitance of any value. After initially being discharged, the digital signal switches each capacitor to either V_{REF} or ground, causing the output voltage, v_{OUT}, to be a function of the voltage division between the capacitors.

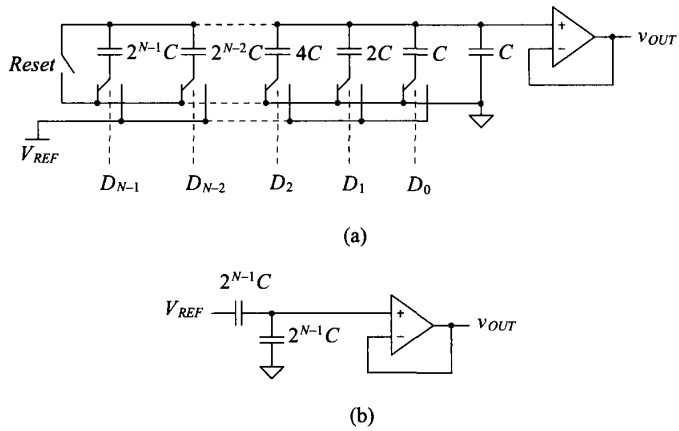


Figure 29.12 (a) A charge-scaling DAC, (b) the equivalent circuit with the MSB = 1, and all other bits set to zero.

The capacitor array totals $2^N C$. Therefore, if the MSB is high and the remaining bits are low, then a voltage divider occurs between the MSB capacitor and the rest of the array. The analog output voltage, v_{OUT}, becomes

$$v_{OUT} = V_{REF} \cdot \frac{2^{N-1}C}{(2^{N-1}+2^{N-2}+2^{N-3}+\ldots+4+2+1+1)C} = V_{REF} \cdot \frac{2^{N-1}C}{2^N C} = \frac{V_{REF}}{2} \tag{29.31}$$

which confirms the fact that the MSB changes the output of a DAC by ½ V_{REF}. Figure 29.12b shows the equivalent circuit under this condition. The ratio between v_{OUT} and V_{REF} due to each capacitor can be generalized to

$$v_{OUT} = \frac{2^k C}{2^N C} \cdot V_{REF} = 2^{k-N} \cdot V_{REF} \tag{29.32}$$

where it is assumed that the k-th bit, D_k, is one and all other bits are zero. Superposition can then be used to find the value of v_{OUT} for any digital input word by

$$v_{OUT} = \sum_{k=0}^{N-1} D_k 2^{k-N} \cdot V_{REF} \tag{29.33}$$

IPR2023-00698
Theta EX2019

One limitation of this architecture as shown in Fig. 29.12a is the existence of a parasitic capacitance at the top plate of the capacitor array due to the op-amp. This will prohibit its use as a high-resolution data converter. A better implementation would include the use of a parasitic insensitive, switched-capacitor integrator (see Ch. 25) as the driving circuit. However, the capacitor array itself is the critical component of this data converter and is used in charge redistribution ADCs (Sec. 29.2.5).

The INL and DNL calculations for the binary-weighted capacitor array are identical to those for the binary-weighted current source array, except that the unit current source, I, and its corresponding error term, ΔI, are replaced by C and ΔC in Eqs. (29.27) – (29.30).

Example 29.6

Design a 3-bit charge-scaling DAC and find the value of the output voltage for $D_2D_1D_0 = 010$ and 101. Assume that $V_{REF} = 5$ V and $C = 0.5$ pF.

The 3-bit DAC can be seen in Fig. 29.13a. The equivalent circuits for the capacitor array can be seen in Fig. 29.13b and c. The value of the output voltage can be calculated by either using Eq. (29.32) or the equivalent circuits and performing the voltage division. For $D = 010$, the equivalent circuit in Fig. 29.13b yields

$$v_{OUT} = V_{REF} \cdot (\frac{1}{4}) = 1.25 \text{ V}$$

Using Eq. (29.33) to calculate v_{OUT} for $D = 101$ yields

$$v_{OUT} = \sum_{k=0}^{N-1} D_k 2^{k-N} \cdot V_{REF} = [1 \cdot (2^{-3}) + 0 \cdot (2^{-2}) + 1 \cdot (2^{-1})] \cdot 5 = \left(\frac{1}{8} + \frac{1}{2}\right) \cdot 5 = 3.125 \text{ V}$$

which is the result expected. ■

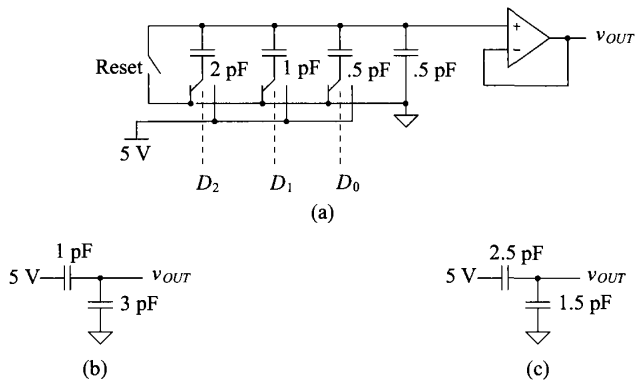


Figure 29.13 (a) A 3-bit charge-scaling DAC used in Ex. 29.6 and the equivalent circuits inputs equal to (b) 010 (c) 101.

IPR2023-00698
Theta EX2019

Layout Considerations for a Binary-Weighted Capacitor Array

One problem with this converter is the need for precisely ratioed capacitors. As the number of bits increase, the ratio of the MSB capacitor to the LSB capacitor becomes more difficult to control. For example, Fig. 29.14a shows a 3-bit binary capacitor array using three capacitors. When the capacitor is fabricated, *undercutting* of the mask causes an error in the ratio of the capacitors, creating potentially large DNL and INL errors as N increases.

One solution to this problem is seen in Fig. 20.14b. Here, each capacitor in the array is constructed out of a unit capacitance. Undercutting then affects all of the capacitors in the same way, and the ratio between capacitors is maintained. Another problem that affects even this layout strategy is a nonuniform oxide growth. Gradients result in errors in the ratios of the capacitors. Figure 29.14c illustrates another layout strategy that overcomes this issue. The capacitors are laid out in a common-centroid scheme so that the first-order oxide errors average out to be the same for each capacitor.

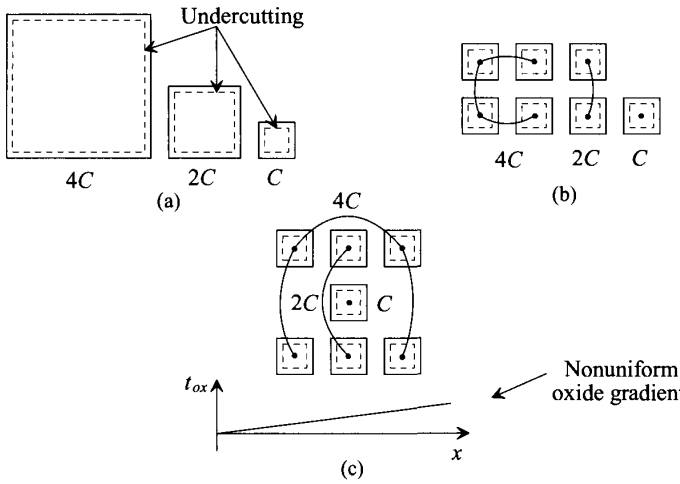


Figure 29.14 Layout of a binary-weighted capacitor array using (a) single capacitors (b) unit capacitors to minimize undercutting effect, and (c) common-centroid to minimize oxide gradients.

The Split Array

The charge-scaling architecture is very popular among CMOS designers because of its simplicity and relatively good accuracy. Although a linear capacitor is required using poly2, high resolutions in the 10- to 12-bit range can be achieved. Passive, double-poly capacitors have good matching accuracy as well. However, as the resolution increases, the size of the MSB capacitor becomes a major concern. For example if the unit capacitor, C,

IPR2023-00698
Theta EX2019

were 0.5 pF, and a 16-bit DAC were to be designed, the MSB capacitor would need to be

$$C_{MSB} = 2^{N-1} \cdot 0.5 \text{ pF} = 16.384 \text{ nF} \tag{29.34}$$

If the capacitance between poly1 and poly2 is nominally 25 fF/μm^2, then the area required for this one capacitor is (roughly) 800 by 800 μm^2.

One method of reducing the size of the capacitors is to use a split array. A 6-bit example of the array is pictured in Fig. 29.15. This architecture is slightly different from the charge-scaling DAC pictured in Fig. 29.13 in that the output is taken off a different node and an additional attenuation capacitor is used to separate the array into a LSB array and a MSB array. Note that the LSB, D_0, now corresponds to the leftmost switch and that the MSB, D_5, corresponds to the rightmost switch. The value of the attenuation capacitor can be found by

$$C_{atten} = \frac{sum\ of\ the\ LSB\ array\ capacitors}{sum\ of\ the\ MSB\ array\ capacitors} \cdot C \tag{29.35}$$

where the sum of the MSB array equals the sum of LSB capacitor array minus C. The value of the attenuation capacitor should be such that the series combination of the attenuation capacitor and the LSB array, assuming all bits are zero, equals C.

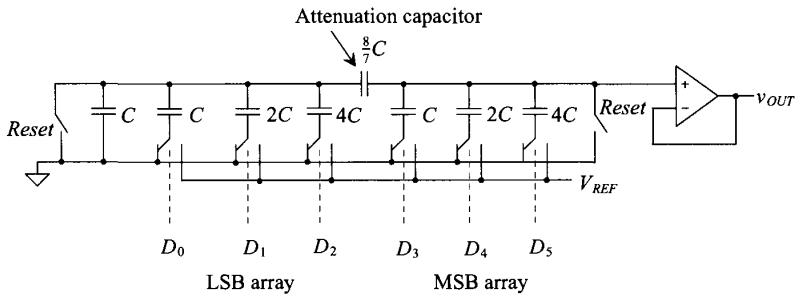


Figure 29.15 A charge-scaling DAC using a split array.

Example 29.7

Using the 6-bit charge-scaling DAC shown in Fig. 29.15, (a) show that the output voltage will be $\frac{1}{2} \cdot V_{REF}$ if (a) $D_5 D_4 D_3 D_2 D_1 D_0$ = 100000 and (b) the output will be $\frac{1}{64} \cdot V_{REF}$ if $D_5 D_4 D_3 D_2 D_1 D_0 = 000001$.

(a) If $D_5 = 1$ and the remaining bits are all zero, then the equivalent circuit for the DAC can be represented by Fig. 29.16a. The expression for the output voltage then becomes

$$v_{OUT} = \frac{4}{\left(\frac{8}{7} \text{ in series with } 8\right) + 3 + 4} \cdot V_{REF} = \frac{1}{2} \cdot V_{REF}$$

IPR2023-00698
Theta EX2019

(b) For the second case, the equivalent circuit can be seen in Fig. 29.16b. The intermediate node voltage, V_A, is simply the voltage division between the C associated with D_0 and the remainder of the circuit, or

$$V_A = V_{REF} \cdot \frac{1}{\left(7 + \frac{\frac{8}{7} \cdot 7}{\frac{8}{7}+7}\right) + 1} = \frac{1}{8 + \frac{56}{57}} \cdot V_{REF} \tag{29.36}$$

The output voltage can be written as

$$v_{OUT} = V_A \cdot \frac{\frac{8}{7}}{\frac{8}{7}+7} = \frac{8}{57} \cdot V_A \tag{29.37}$$

Plugging Eq. (29.36) into Eq. (29.37) yields

$$v_{OUT} = V_{REF} \cdot \frac{8}{(8 \cdot 57) + 56} = \frac{V_{REF}}{64} \tag{29.38}$$

which is the desired result. ■

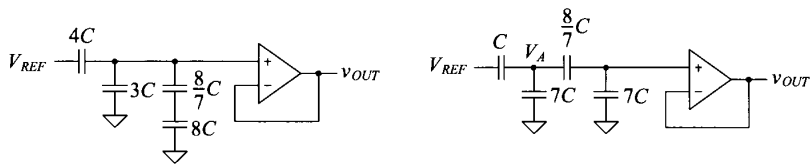


Figure 29.16 Equivalent circuits for Example 29.7.

29.1.6 Cyclic DAC

The cyclic DAC uses only a couple of simple components to perform the conversion. As seen in Fig. 29.17, a summer adds V_{REF} or ground to the feedback signal depending on the input bits. An amplifier with a gain of 0.5 feeds the output voltage back to the summer such that the output at the end of each cycle depends on the value of the output during the cycle before. Notice that the input bits must be read in a serial fashion. Therefore, the conversion is performed one bit at a time, resulting in N cycles required for each conversion. The voltage output at the end of the n-th cycle of the conversion can be written as

$$v_{OUT}(n) = \left(D_{n-1} \cdot V_{REF} + \frac{1}{2} \cdot v_A(n-1)\right) \cdot \frac{1}{2} \tag{29.39}$$

with a condition such that the output of the S/H is initially zero [$v_A(0) = 0$ V].

The accuracy of this converter is dependent on several factors. The gain of the 0.5 amplifier needs to be highly accurate (to within the accuracy of the DAC) and is usually generated with passive capacitors. Similarly, the summer and the sample-and-hold also need to be N-bit accurate. Limitations of the converter due to these fundamental building blocks will be discussed in more detail in Sec. 29.2.3. Since this converter uses a pseudo-"sampled-data" approach, implementing this architecture using switched capacitors is relatively easy.

IPR2023-00698
Theta EX2019

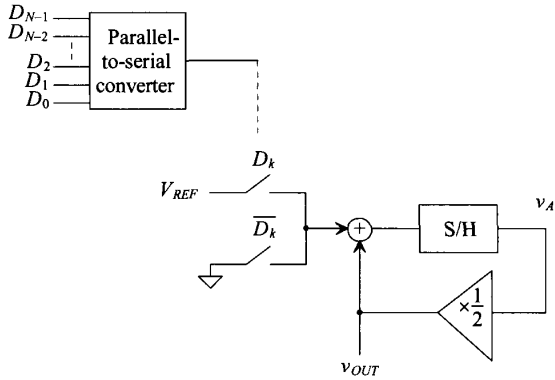


Figure 29.17 A cyclic digital-to-analog converter.

Example 29.8

Show the value of the output voltage at the end of each cycle for a 6-bit cyclic DAC with an input value of $D_5D_4D_3D_2D_1D_0 = 110101$. Assume that $V_{REF} = 5$ V.

We can predict the value of the output based on our previous experience with DACs. The digital input 110101 corresponds to 53_{10}. Therefore, the output voltage due to this input should be

$$v_{OUT} = \frac{53}{64} \cdot V_{REF} = 4.140625 \text{ V}$$

Now examine the cyclic converter in Fig. 29.17. By performing a 6-bit conversion and using Eq. (29.39), the outputs occurring at the end of each cycle can be seen in Fig. 29.18.

The output voltage at the end of the sixth cycle is precisely what was predicted. Note that had this been a 3-bit conversion, the output voltage at the end of cycle 3 would correspond to the value of the 3-bit DACs studied previously with an input of 101. ■

Cycle Number, n	D_{n-1}	$v_A(n\text{-}1)$	$v_{OUT}(n)$
1	1	0	½ (5 + 0) = 2.5 V
2	0	5	½ (0 + 2.5) = 1.25 V
3	1	2.5	½ (5 +1.25) = 3.125 V
4	0	6.25	½ (0 + 3.125) = 1.5625 V
5	1	3.125	½ (5 + 1.5625) = 3.28125 V
6	1	6.5625	½ (5 + 3.28125) = 4.140625 V

Figure 29.18 Output from the 6-bit cyclic DAC used in Ex. 29.8.

IPR2023-00698
Theta EX2019

29.1.7 Pipeline DAC

The cyclic converter presented in the last section takes N clock cycles per N-bit conversion. Instead of recycling the output back to the input each time, we could extend the cyclic converter to N stages, where each stage performs one bit of the conversion. This extension of the cyclic converter is called a *pipeline* DAC and is seen in Fig. 29.19. Here, the signal is passed down the "pipeline," and as each stage works on one conversion, the previous stage can begin processing another. Therefore, an initial N clock cycle delay is experienced as the signal makes its way down the pipeline the very first time. After the N clock cycle delay, a conversion takes place at every clock cycle.

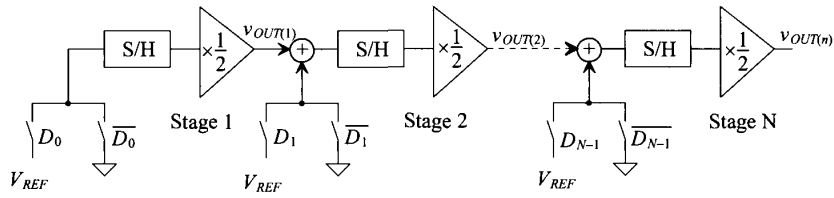


Figure 29.19 A pipeline digital-to-analog converter.

Besides the N clock cycle delay, this architecture can be very fast. However, the amplifier gains must be very accurate to produce high resolutions. Also, this architecture uses N times more circuitry than that of the cyclic, so there is a trade-off between speed and chip area. The output voltage of the n-th stage in the converter can be written as

$$v_{OUT(n)} = [D_{n-1} \cdot V_{REF} + v_{OUT(n-1)}] \cdot \frac{1}{2} \tag{29.40}$$

The operation of each stage in the pipeline can be summarized as follows: if the input bit is a 1, add V_{REF} to the output of the previous stage, divide by two, and pass the value to the next stage. If the input bit is a 0, simply divide the output of the previous stage by two and pass along the resulting value.

Example 29.9

Find the output voltage for a 3-bit pipeline DAC for three cases: $D_A = 001$, $D_B = 110$, and $D_C = 101$. Show that the conversion time to perform all three conversions is five clock cycles using the pipeline approach. Assume that $V_{REF} = 5$ V.

The first stage operates on the LSBs of each word; the second stage operates on the middle bits; and the last stage, the MSBs. Based on the pipeline strategy, once the LSB of the first input word is performed and passed on, the LSB of the second word, D_B, can begin its conversion. Similarly, once the LSB of the second stage is completed and passed on, the LSB of the third word, D_C, can begin. The conversion cycle for all three input words produces the output shown in Fig. 29.20. The items that are in bold are associated with the first input word, D_A, whereas the italicized numbers represent the values associated with D_B and the underlined items, D_C.

IPR2023-00698
Theta EX2019

Clock Cycle	$v_{OUT(1)}$	$v_{OUT(2)}$	$v_{OUT(3)}$	D_0	D_1	D_2
1	**2.5**	0	0	**1**	0	0
2	*0*	**1.25**	0	*0*	**0**	0
3	2.5	*2.5*	**0.625**	1	*1*	**0**
4		1.25	*3.75*		0	*1*
5			3.125			1

Figure 29.20 Output from the 3-bit pipeline DAC used in Example 29.9.

The first output of the DAC is not valid until the end of the third clock cycle and should look familiar as the 3-bit DAC output for an input word of $D_2 D_1 D_0 = 001$. The following two clock cycles that produce outputs for $D_2 D_1 D_0$ equal 110 and 101, respectively. ■

29.2 ADC Architectures

A survey of the field of current A/D converter research reveals that a majority of effort has been directed to four different types of architectures: pipeline, flash-type, successive approximation, and oversampled ADCs. Each has benefits that are unique to that architecture and span the spectrum of high speed and resolution.

Since the ADC has a continuous, infinite-valued signal as its input, the important analog points on the transfer curve x-axis for an ADC are the ones that correspond to changes in the digital output word. These input transitions determine the amount of INL and DNL associated with the converter.

29.2.1 Flash

Flash or parallel converters have the highest speed of any type of ADC. As seen in Fig. 29.21, they use one comparator per quantization level ($2^N - 1$) and 2^N resistors (a resistor-string DAC). The reference voltage is divided into 2^N values, each of which is fed into a comparator. The input voltage is compared with each reference value and results in a thermometer code at the output of the comparators. A thermometer code exhibits all zeros for each resistor level if the value of v_{IN} is less than the value on the resistor string, and ones if v_{IN} is greater than or equal to voltage on the resistor string. A simple $2^N - 1{:}N$ digital thermometer decoder circuit converts the compared data into an N-bit digital word. The obvious advantage of this converter is the speed with which one conversion can take place. Each clock pulse generates an output digital word. The advantage of having high speed, however, is counterbalanced by the doubling of area with each bit of increased resolution. For example, an 8-bit converter requires 255 comparators, but a 9-bit ADC requires 511! Flash converters have traditionally been limited to 8-bit resolution with conversion speeds of 10–40 Ms/s using CMOS technology. The disadvantages of the Flash ADC are the area and power requirements of the $2^N - 1$ comparators. The speed is limited by the switching of the comparators and the digital logic.

IPR2023-00698
Theta EX2019

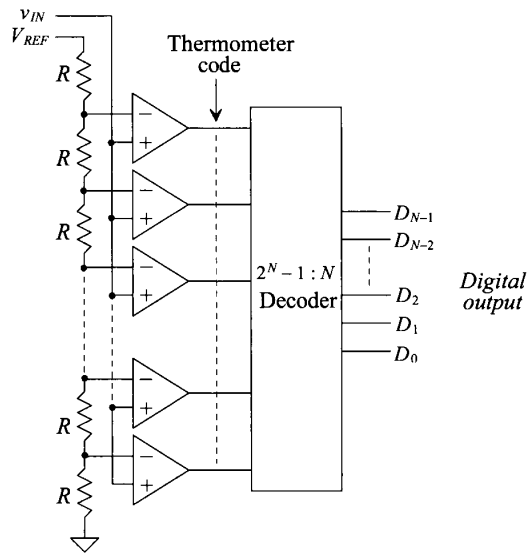


Figure 29.21 Block diagram of a Flash ADC.

Example 29.10

Design a 3-bit Flash converter, listing the values of the voltages at each resistor tap, and draw the transfer curve for v_{IN} = 0 to 5 V. Assume V_{REF} = 5 V. Construct a table listing the values of the thermometer code and the output of the decoder for v_{IN} = 1.5, 3.0, and 4.5 V.

The 3-bit converter can be seen in Fig. 29.22. As the values of all the resistors are equal, the voltage of each resistor tap, V_i, will be $V_i = V_{REF}\left(\frac{i}{8}\right)$ where i is the number of the resistor in the string for i = 1 to 7. Obviously, V_1 = 0.625 V, V_2 = 1.25 V, V_3 = 1.875 V, V_4 = 2.5 V, V_5 = 3.125, V_6 = 3.75 V, V_7 = 4.375 V. Therefore, when v_{IN} first becomes equal or greater than each of these values, a transition will occur in the transfer curve. The transfer curve can be seen in Fig. 29.23 and should look similar to those seen in Ch. 28. The quantization levels and their corresponding thermometer codes are summarized in Fig. 29.24.

The transfer curve of this ADC corresponds to the ADC with quantization error centered about +½ LSB, as discussed in Ch. 28 (Fig. 28.20). To shift the curve by ½ LSB so that the code transitions occur around the LSB values and the quantization error is centered around 0 LSB, the value of the last resistor in the string would have to be adjusted to $\frac{R}{2}$ and the value of the MSB resistor, closest to the reference voltage, would have to be made 1.5R. Then the first code transition would occur at v_{IN} = 0.3125 V, and the last code transition would occur at v_{IN} = 4.0625, and so the transfer curve would exactly match that of Fig. 28.20.

IPR2023-00698
Theta EX2019

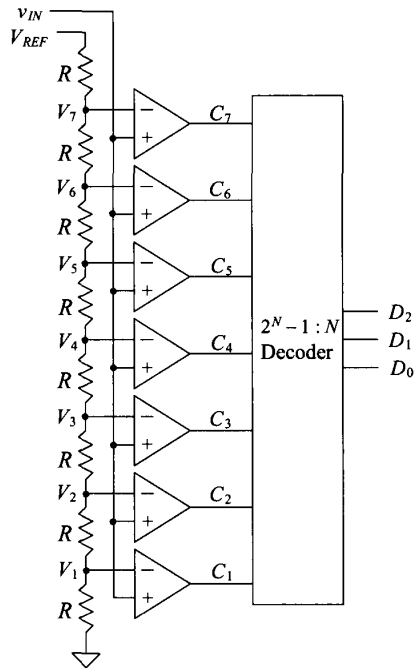


Figure 29.22 Three-bit Flash A/D converter to be used in Ex. 29.10.

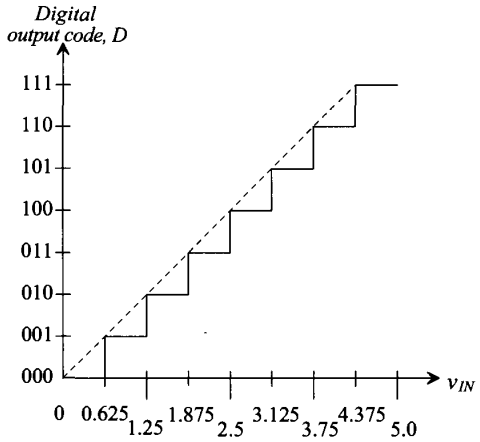


Figure 29.23 Transfer curve for the 3-bit Flash converter in Example 29.10.

IPR2023-00698
Theta EX2019

v_{IN}	$C_7C_6C_5C_4C_3C_2C_1$	$D_2D_1D_0$
$0 \le v_{IN} < 0.625$ V	0000000	000
0.625 V $\le v_{IN} < 1.25$ V	0000001	001
1.25 V $\le v_{IN} < 1.875$ V	0000011	010
1.875 V $\le v_{IN} < 2.5$ V	0000111	011
2.5 V $\le v_{IN} < 3.125$ V	0001111	100
3.125 V $\le v_{IN} < 3.75$ V	0011111	101
3.75 V $\le v_{IN} < 4.375$ V	0111111	110
$4.375 \le v_{IN}$	1111111	111

Figure 29.24 Code transitions for the Flash ADC used in Ex. 29.10.

Based on Fig. 29.24, when $v_{IN} = 1.5$ V, only comparators C_1 and C_2 will have outputs of 1, since both V_1 and V_2 are less than 1.5 V. The remaining comparator outputs will be 0 since V_3 through V_8 will be greater than 1.5 V, thus generating the thermometer code, 0000011. The encoder must then convert this into a 3-bit digital word, resulting in 010. The same reasoning can be used to construct the data shown in Fig. 29.25. It should be obvious that if the polarity of the comparators were reversed, the thermometer code would be inverted. ■

v_{IN}	$C_7C_6C_5C_4C_3C_2C_1$	$D_2D_1D_0$
1.5	0000011	010
3.0	0001111	100
4.5	1111111	111

Figure 29.25 Output for the Flash ADC used in Ex. 29.10.

Accuracy Issues for the Flash ADC

Accuracy depends on the matching of the resistor string and the input offset voltage of the comparators. From our discussions earlier, we know that an ideal comparator should switch at the point at which the two inputs, v_+ and v_-, are the same potential. However, the offset voltage, V_{os}, prohibits this from occurring as the comparator output switches states as follows:

$$v_o = 1 \qquad \text{when } v_+ \ge v_- + V_{os} \tag{29.41}$$

$$v_o = 0 \qquad \text{when } v_+ < v_- + V_{os} \tag{29.42}$$

The resistor-string DAC was analyzed and presented in Sec. 29.1.2; the voltage on the i-th tap of the resistor string was found to be

$$V_i = V_{i,ideal} + \frac{V_{REF}}{2^N} \cdot \sum_{k=1}^{i} \frac{\Delta R_k}{R} \tag{29.43}$$

IPR2023-00698
Theta EX2019

where $V_{i,ideal}$ is the voltage at the i-th tap if all the resistors had an ideal value of R. The term, ΔR_k, is the value of the resistance error (difference from ideal) due to the mismatch. Note that for the resistor-string DAC, the sum of the mismatch terms plays an important factor in the overall voltage at each tap.

The switching point for the i-th comparator, $V_{sw,i}$, then becomes

$$V_{sw,i} = V_i + V_{os,i} \tag{29.44}$$

where $V_{os,i}$ is the input referred offset voltage of the i-th comparator. The INL for the converter can then be described as

$$INL = V_{sw,i} - V_{sw,ideal} = V_{sw,i} - V_{i,ideal} \tag{29.45}$$

which becomes

$$INL = \frac{V_{REF}}{2^N} \cdot \sum_{k=1}^{i} \frac{\Delta R_k}{R} + V_{os,i} \tag{29.46}$$

The worst-case INL will occur at the middle of the string ($i = 2^{N-1}$), as described in Sec. 29.1.2 and Eq. (29.10). Including the offset voltage, the maximum INL will be

$$|INL|_{max} = \frac{V_{REF}}{2^N} \cdot \sum_{k=1}^{2^{N-1}} \frac{\Delta R_k}{R} + |V_{os,i}|_{max} = V_{REF} \cdot \frac{2^{N-1}}{2^N R} \cdot |\Delta R_k|_{max} + |V_{os,i}|_{max} \tag{29.47}$$

which can be rewritten as

$$|INL|_{max} = \frac{V_{REF}}{2} \cdot \left|\frac{\Delta R_k}{R}\right|_{max} + |V_{os,i}|_{max} \tag{29.48}$$

where it is assumed that the maximum positive mismatch occurs in all the resistors in the lower half of the string and the maximum negative mismatch occurs in the upper half (or vice versa) and that the comparator at the i-th tap contains the maximum offset voltage, $|V_{os,i}|_{max}$. Notice that the offset contributes directly to the maximum value for the INL. This explains another limitation to using Flash converters at high resolutions. The offset voltage alone can make the INL greater than ½ LSB.

Example 29.11

If a 10-bit Flash converter is designed, determine the maximum offset voltage of the comparators which will make the INL less than ½ LSB. Assume that the resistor string is perfectly matched and $V_{REF} = 5$ V.

Equation (29.48) requires that the offset voltage equal ½ LSB. Therefore,

$$|V_{os}|_{max} = \frac{5}{2^{11}} = 2.44 \text{ mV} \quad \blacksquare$$

The DNL calculation for the Flash converter is also attained using the analysis first presented in Sec. 29.1.2. Using the definition of DNL,

$$DNL = V_{sw,i} - V_{sw,i-1} - 1 \text{ LSB (in volts)} \tag{29.49}$$

Plugging in Eq. (29.44),

$$DNL = V_i + V_{os,i} - V_{i-1} - V_{os,i-1} - 1 \text{ LSB} \tag{29.50}$$

IPR2023-00698
Theta EX2019

which can be written by using Eq. (29.6) as

$$DNL = V_{i,ideal} - V_{i-1,ideal} + \frac{V_{REF}}{2^N} \cdot \frac{\Delta R_i}{R} + V_{os,i} - V_{os,i-1} - 1\ LSB \qquad (29.51)$$

which becomes

$$DNL = \frac{V_{REF}}{2^N} \cdot \frac{\Delta R_i}{R} + V_{os,i} - V_{os,i-1} \qquad (29.52)$$

The maximum DNL will occur, assuming ΔR_i is at its maximum, $V_{os,i}$ is at its maximum positive value, and $V_{os,i-1}$ is at its maximum negative voltage. Thus,

$$|DNL|_{max} = \frac{V_{REF}}{2^N} \cdot \left|\frac{\Delta R_i}{R}\right|_{max} + 2|V_{os}|_{max} \qquad (29.53)$$

which assumes that the maximum offset voltage in the positive and negative directions are symmetrical. Therefore, both resistor-string matching and offset voltage affect the DNL of the converter.

29.2.2 The Two-Step Flash ADC

Another type of Flash converter is called the two-step Flash converter or the parallel, feed-forward ADC. The basic block diagram of a two-step converter is seen in Fig. 29.26. The converter is separated into two complete Flash ADCs with feed-forward circuitry. The first converter generates a rough estimate of the value of the input, and the second converter performs a fine conversion. The advantages of this architecture are that the number of comparators is greatly reduced from that of the Flash converter—from $2^N - 1$ comparators to $2(2^{N/2} - 1)$ comparators. For example, an 8-bit Flash converter requires 255 comparators, while the two-step Flash requires only 30. The trade-off is that the conversion process takes two steps instead of one, with the speed limited by the bandwidth and settling time required by the residue amplifier and the summer. The conversion process is as follows:

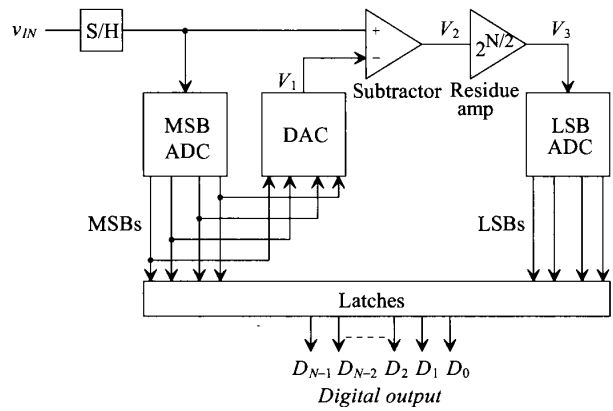


Figure 29.26 Block diagram of a two-step Flash ADC.

IPR2023-00698
Theta EX2019

1. After the input is sampled, the most significant bits (MSBs) are converted by the first Flash ADC.
2. The result is then converted back to an analog voltage with the DAC and subtracted with the original input.
3. The result of the subtraction, known as the *residue*, is then multiplied by $2^{N/2}$ and input into the second ADC. The multiplication not only allows the two ADCs to be identical, but also increases the quantum level of the signal input into the second ADC.
4. The second ADC produces the least significant bits through a Flash conversion.

Some architectures use the same set of comparators in order to perform both steps. The multiplication mentioned in step 3 can be eliminated if the second converter is designed to handle very small input signals. The accuracy of the two-step ADC depends primarily on the linearity of the first ADC.

Figure 29.27 illustrates the two-step nature of the converter. A more intuitive approach can be explained with this picture. The first conversion identifies the segment in which the analog voltage resides. This is also known as a *coarse conversion* of the MSBs. The results of the coarse conversion are then multiplied by $2^{N/2}$ so that the segment within which V_{IN} resides will be scaled to the same reference as the first conversion. The second conversion is known as the *fine conversion* and will generate the final LSBs using the same Flash approach. One can see why the accuracy of the first converter is so important. If the input value is close to the boundary between two coarse segments and the first ADC is unable to choose the correct coarse segment, then the second conversion will be completely erroneous. The following example further illustrates the two-step algorithm.

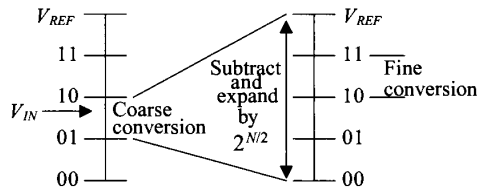


Figure 29.27 Coarse and fine conversions using a two-step ADC.

Example 29.12

Assume that the two-step ADC shown in Fig. 29.26 has four bits of resolution. Make a table listing the MSBs, V_1, V_2, V_3, and the LSBs for V_{IN} = 2, 4, 9, and 15 V assuming that V_{REF} = 16 V.

Since V_{REF} was conveniently made 16 V, each LSB will be 1 V. If V_{IN} = 2 V, the output of the first 2-bit Flash converter will be 00 since V_{REF} = 16 V and each resistor drops 4 V. The output of the 2-bit DAC, V_1, will therefore be 0, resulting in V_2 = 2 V. The multiplication of V_2 by the 4 results in V_3 = 8 V. Remember that

IPR2023-00698
Theta EX2019

each 2-bit Flash converter resembles that of Fig. 29.21. The thermometer code from the second Flash converter will be 0011, which results in 10 as the LSBs. The other values can be calculated as seen in Fig. 29.28. ■

V_{IN}	D_3D_2 (MSBs)	V_1	V_2	V_3	D_1D_0 (LSBs)
2	00	0	2	8	10
4	01	4	0	0	00
9	10	8	1	4	01
15	11	12	3	12	11

Figure 29.28 Output for the Flash ADC used in Ex. 29.12.

Accuracy Issues Related to the Two-Step Flash Converters

As stated previously, the overall accuracy of the converter depends on the first ADC. The second Flash must have only the accuracy of a stand-alone Flash converter. This means that if an 8-bit, two-step Flash converter contains two 4-bit Flash converters, the second Flash needs only to have the resolution of a 4-bit Flash, which is not difficult to achieve. However, the first 4-bit Flash must have the accuracy of an 8-bit Flash, meaning that the worst-case INL and DNL for the first 4-bit Flash must be less than ±½ LSB for an 8-bit ADC. Thus, the resistor matching and comparators contained in the first ADC must possess the accuracy of the overall converter. Refer to Sec. 29.2.1 for derivations on INL and DNL for a Flash. The DAC must also be accurate to within the resolution of the ADC.

Accuracy Issues Related to the Operational Amplifiers

With the addition of the summer and the amplifier, other sources of accuracy errors are present in this converter. The summer and the amplifier must add and amplify the signal to within ±½ LSB of the ideal value. It is difficult to implement standard operational amplifiers within high-resolution data converters because of these accuracy requirements. The nonideal characteristics of the op-amp are well known and in many cases alone limit the accuracy of the data converter. In this case, the amplifier is required to multiply the residue signal by some factor of two. Although this may not seem difficult at first glance, a closer examination will reveal a dependency on the open-loop gain.

Suppose that the amplifier were being used in a 12-bit, two-step data converter. Remember that in order for a data converter to be N-bit accurate, the INL and DNL need to be kept below ±½ LSB and one-half of an LSB can be defined as

$$0.5 \text{ LSB} = \frac{V_{REF}}{2^{N+1}} \tag{29.54}$$

Since the output of the amplifier gets quantized to 6 bits, the amplifier would need to be 6-bit accurate to within ±½ LSB, resulting in an accuracy of

$$\text{Accuracy} = \frac{0.5\ LSB}{\text{Full scale range } (V_{REF})} = \frac{1}{2^{6+1}} = \frac{1}{128} = 0.0078 = 0.78\% \tag{29.55}$$

And suppose that a feedback amplifier with a gain of 64, or $2^{N/2}$, is used as the residue amplifier. The gain would need to be within the following range:

IPR2023-00698
Theta EX2019

$$63.5 \text{ V/V} < A_{CL} < 64.5 \text{ V/V} \tag{29.56}$$

where A_{CL} is the closed-loop gain of the amplifier. Already, one can see the limitations of using operational amplifiers with feedback in high-accuracy applications. Designing an op-amp based amplifier with a high degree of gain accuracy can be difficult.

Generalizing this concept for an N-bit application requires knowledge of feedback theory discussed in Ch. 24. The closed-loop gain of the amplifier is expressed as

$$A_{CL} = \frac{v_o}{v_i} = \frac{A_{OL}}{1 + A_{OL}\beta} \tag{29.57}$$

where A_{OL} is the open-loop gain of the amplifier and β is the feedback factor. Also, from Ch. 24, it is known that as A_{OL} increases in value, the closed-loop gain, A_{CL}, approaches the value of $1/\beta$. Therefore, if it is assumed the closed-loop gain of the amplifier equals the ideal value of $1/\beta$ minus some maximum deviation from the ideal, ΔA, then,

$$A_{CL} = \frac{v_o}{v_i} = \frac{A_{OL}}{1 + A_{OL}\beta} = \frac{1}{\beta} - \Delta A \tag{29.58}$$

where $1/\beta$ is the desired value of the closed-loop gain (usually some factor of 2^N) and ΔA is the required accuracy ($\pm$½ LSB) of the gain (i.e., $(1/\beta) \cdot (1/2^{N+1})$). The right two terms of Eq. (29.58) can be solved for the open-loop gain of the amplifier,

$$|A_{OL}| = \frac{1}{\beta}(2^{N+1} - 1) \approx \frac{2^{N+1}}{\beta} \tag{29.59}$$

If the op-amp is used as a gain of 64 ($1/\beta$) and is required to amplify signals with 6-bit accuracy, then the open-loop gain of the amplifier must be at least $|A_{OL}| \geq 128 \cdot 64 = 8,192$ V/V. This is certainly an achievable specification. However, notice that for every bit increase in resolution, the open-loop gain requirement doubles. This is one reason two-step Flash converters are limited in resolution to approximately 12 bits (or less in a nanometer CMOS process).

The unity-gain frequency, f_{un} , required of an op-amp used in or with a data converter for a specific settling time t, (where $t < T_{clk}/2 = 1/2f_{clk}$) can be estimated assuming linear settling, and requiring the output of the op-amp be ½ LSB accurate, by

$$v_{out} = V_{outfinal}\left(1 - \frac{1}{2^{N+1}}\right) = V_{outfinal}(1 - e^{-t/\tau}) \text{ or } f_{un} \geq \frac{f_{clk} \cdot \ln 2^{N+1}}{\pi \cdot \beta} \tag{29.60}$$

This equation can be used to determine the minimum op-amp gain-bandwidth product (= f_{un}) needed to achieve a specific settling time provided the op-amp slew-rate doesn't come into play and the op-amp can be modeled as a first-order system.

Linearity of the amplifier is another aspect of amplifier performance that must be considered when designing ADCs. The amplifier must be able to linearly amplify the input signal over an input voltage range to within ½ LSB of the number bits that its output is quantized. If the amplifier is not designed correctly, nonlinearity is introduced as devices in the amplifier go into nonsaturation. Harmonic distortion occurs, resulting in an error within the ADC. Linearity is typically measured in terms of total harmonic distortion, or THD, (refer to Sec. 21.3.3). However, the transfer curve illustrates the

IPR2023-00698
Theta EX2019

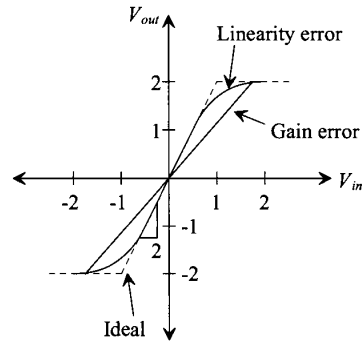


Figure 29.29 An op-amp transfer curve that distinguishes between gain error and linearity error.

limitation more effectively. Figure 29.29 shows a transfer curve of an op-amp with a gain of two. The ideal transfer curve is shown if the input range is known to be between –1 and 1 V. The actual transfer curve shows nonlinearity introduced at both ends of the input range. In order for the amplifier to be *N*-bit accurate, the slope of the actual transfer curve may not vary from the ideal by more than the accuracy required at the output of the amplifier. Note also in Fig. 29.29 the subtle difference between a gain error and nonlinearity. However, a gain error is much less harmful to an ADC's performance than harmonic distortion.

29.2.3 The Pipeline ADC

After examining the two-step ADC, one might wonder whether there is such a converter as a three-step or four-step ADC. In actuality, one could divide the number of conversions into many steps. The pipeline ADC is an *N*-step converter, with 1 bit being converted per stage. Able to achieve high resolution (10–13 bits) at relatively fast speeds, the pipeline ADC consists of *N* stages connected in series (Fig. 29.30). Each stage contains a 1-bit ADC (a comparator), a sample-and-hold, a summer, and a gain of two amplifier. Each stage of the converter performs the following operation:

1. After the input signal has been sampled, compare it to $\frac{V_{REF}}{2}$. The output of each comparator is the bit conversion for that stage.
2. If $v_{IN} > \frac{V_{REF}}{2}$ (comparator output is 1), $\frac{V_{REF}}{2}$ is subtracted from the held signal and pass the result to the amplifier. If $v_{IN} < \frac{V_{REF}}{2}$ (comparator output is 0), then pass the original input signal to the amplifier. The output of each stage in the converter is referred to as the *residue*.
3. Multiply the result of the summation by 2 and pass the result to the sample-and-hold of the next stage.

A main advantage of the pipeline converter is its high throughput. After an initial *latency* of *N* clock cycles, one conversion will be completed per clock cycle. While the residue of the first stage is being operated on by the second stage, the first stage is free to

IPR2023-00698
Theta EX2019

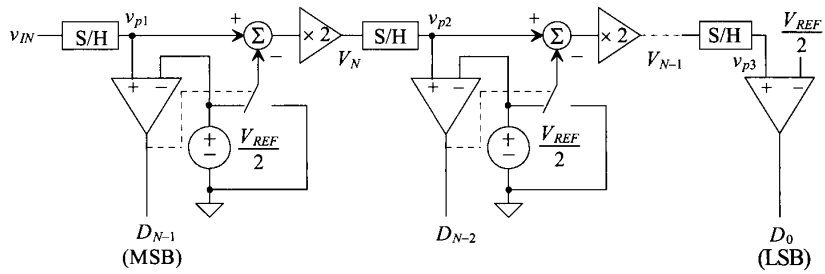


Figure 29.30 Block diagram of a pipeline ADC.

operate on the next samples. Each stage operates on the residue passed down from the previous stage, thereby allowing for fast conversions. The disadvantage is having the initial N clock cycle delay before the first digital output appears. The severity of this disadvantage depends, of course, on the application.

One interesting aspect of this converter is its dependency on the most significant stages for accuracy. A slight error in the first stage propagates through the converter and results in a much larger error at the end of the conversion. Each succeeding stage requires less accuracy than the one before, so special care must be taken when considering the first several stages.

Example 29.13

Assume that the pipeline converter shown in Fig. 29.30 is a 3-bit converter. Analyze the conversion process by making a table of the following variables: D_2, D_1, D_0, V_3, V_2, for $v_{IN} = 2$, 3, and 4.5 V. Assume that $V_{REF} = 5$ V, V_3 is the residue voltage out of the first stage, and V_2 is the residue voltage out of the second stage.

The output of the first comparator, $D_2 = 0$, since $v_{IN} < 2.5$ V. Since $D_2 = 0$, $V_3 = 2(2) = 4$ V. Passing this voltage down the pipeline, since $V_3 > 2.5$ V, $D_1 = 1$ and V_2 becomes

$$V_2 = \left(V_3 - \frac{V_{REF}}{2}\right) \times 2 = 3 \text{ V}$$

The LSB, $D_0 = 1$, since $V_2 > 2.5$ V, and the digital output corresponding to $V_{IN} = 2$ V, is $D_2D_1D_0 = 011$. The actual digital outputs are simply the comparator outputs, and the data can be completed as seen in Fig. 29.31. ■

v_{IN}	V_3 (V)	V_2 (V)	Digital Out ($D_2D_1D_0$)
2.0	4.0	3.0	011
3.0	1.0	2.0	100
4.5	4.0	3.0	111

Figure 29.31 Output for the pipeline ADC used in Ex. 29.13.

IPR2023-00698
Theta EX2019

Accuracy Issues Related to the Pipeline Converter

The 1-bit per stage ADC can be analyzed by examining the switching point of each comparator for the ideal and nonideal case. Using Fig. 29.30, and assuming that all of the components are ideal, let $v_{IN,1}$ represent the value of the input voltage when the first comparator switches. This occurs when

$$v_{IN,1} = \frac{1}{2}V_{REF} \tag{29.61}$$

The positive input voltage on the second comparator, v_{p2}, can be written in terms of the previous stage, or

$$v_{p2} = [v_{IN} - \tfrac{1}{2} \cdot D_{N-1} \cdot V_{REF}] \cdot 2 \tag{29.62}$$

where D_{N-1} is the MSB output from the first comparator and is either a 1 or a 0. The second comparator switches when $v_{p2} = \frac{1}{2}V_{REF}$. The value of v_{IN} at this point, denoted as $v_{IN,2}$, is

$$v_{IN,2} = \tfrac{1}{2} \cdot D_{N-1} \cdot V_{REF} + \tfrac{1}{4}V_{REF} \tag{29.63}$$

Continuing on in a similar manner, we can write the value of the voltage on the positive input of the third comparator in terms of the previous two stages as

$$v_{p3} = \left[[v_{IN} - \frac{1}{2} \cdot D_{N-1} \cdot V_{REF}] \cdot 2 - [\frac{1}{2} \cdot D_{N-2} \cdot V_{REF}] \right] \cdot 2 \tag{29.64}$$

and the third comparator will switch when $v_{p3} = \frac{1}{2}V_{REF}$, which corresponds to the point at which v_{IN} becomes

$$v_{IN,3} = \tfrac{1}{2} \cdot D_{N-1} \cdot V_{REF} + \tfrac{1}{4} \cdot D_{N-2} \cdot V_{REF} + \tfrac{1}{8}V_{REF} \tag{29.65}$$

By now, a general trend can be recognized and the value of v_{IN} can be derived for the point at which the comparator of the N-th stage switches. This expression can be written as

$$v_{IN,N} = \tfrac{1}{2} \cdot D_{N-1} \cdot V_{REF} + \tfrac{1}{4} \cdot D_{N-2} \cdot V_{REF} + \tfrac{1}{8} \cdot D_{N-3} \cdot V_{REF} + ... + \tfrac{1}{2^{N-1}} \cdot D_1 \cdot V_{REF} + \tfrac{1}{2^N} \cdot V_{REF} \tag{29.66}$$

Notice that the preceding equation does not include D_0. This is because D_0 is the output of the N-th stage comparator.

Now that we have derived the switching points for the ideal case, the nonideal case can be considered. Only the major sources of error will be included in the analysis so as not to overwhelm the reader. These include the comparator offset voltage, $V_{COS,x}$, and the sample-and-hold offset voltage, $V_{SOS,x}$. The variable, x, represents the number of the stage for which each of the errors is associated, and the "prime" notation will be used to distinguish between the ideal and nonideal case. The reader should also be aware that the offset voltages can be of either polarity. It will be assumed that all of the residue amplifiers have the same gain, denoted as A.

The positive input to the first nonideal comparator, v'_{p1}, will include the offset from the first sample-and-hold, such that

IPR2023-00698
Theta EX2019

$$v'_{p1} = v_{IN} + V_{SOS,1} \tag{29.67}$$

Now the first comparator will not switch until the voltage on the positive input overcomes the comparator offset as well. This occurs when

$$v'_{p1} = \tfrac{1}{2}V_{REF} + V_{COS,1} \tag{29.68}$$

Thus, equating Eqs. (29.67) and (29.68) and solving for the value of the input voltage when the switching occurs for the first comparator yields

$$v'_{IN,1} = \tfrac{1}{2}V_{REF} + V_{COS,1} - V_{SOS,1} \tag{29.69}$$

The input to the second comparator, v'_{p2}, can be written as

$$v'_{p2} = [v_{IN} + V_{SOS,1} - \tfrac{1}{2} \cdot D_{N-1} \cdot V_{REF}] \cdot A + V_{SOS,2} \tag{29.70}$$

and the value of input voltage at the point which the second comparator switches occurs when

$$v'_{IN,2} = \frac{1}{2} \cdot D_{N-1} \cdot V_{REF} + \frac{1}{2}\frac{V_{REF}}{A} - V_{SOS,1} - \frac{1}{A}(V_{SOS,2} - V_{COS,2}) \tag{29.71}$$

Continuing in the same manner, we can write the value of the input voltage that causes the third comparator to switch as

$$v'_{IN,3} = \tfrac{1}{2} \cdot D_{N-1} \cdot V_{REF} + \tfrac{1}{2} \cdot D_{N-2} \cdot \tfrac{V_{REF}}{A} - V_{SOS,1} - \tfrac{1}{A}V_{SOS,2} - \tfrac{1}{A^2}V_{SOS,3} - \tfrac{1}{A^2}\left[V_{COS,3} - \tfrac{1}{2}V_{REF}\right] \tag{29.72}$$

which can be generalized to the N-th switching point as

$$v'_{IN,N} = \tfrac{1}{2} \cdot D_{N-1} \cdot V_{REF} + \tfrac{1}{2} \cdot D_{N-2} \cdot \tfrac{V_{REF}}{A} + \ldots + \tfrac{1}{2} \cdot D_1 \cdot \tfrac{V_{REF}}{A^{N-2}} + \tfrac{1}{2} \cdot \tfrac{V_{REF}}{A^{N-1}} + \tfrac{V_{COS,N}}{A^{N-1}} - \sum_{k=1}^{N} \tfrac{V_{SOS,k}}{A^{k-1}} \tag{29.73}$$

The INL can be calculated by subtracting switching point between the nonideal and ideal case. Therefore, the INL of the first stage is found by subtracting Eqs. (29.69) and (29.61).

$$INL_1 = v'_{IN,1} - v_{IN,1} = V_{COS,1} - V_{SOS,1} \tag{29.74}$$

The second stage INL is

$$INL_2 = v'_{IN,2} - v_{IN,2} = \frac{V_{REF}}{2}\left(\frac{1}{A} - \frac{1}{2}\right) - V_{SOS,1} - \frac{V_{SOS,2}}{A} + \frac{V_{COS,2}}{A} \tag{29.75}$$

and the INL for the N-th stage is

$$INL_N = \frac{1}{2} \cdot D_{N-2} \cdot V_{REF} \cdot \left(\frac{1}{A} - \frac{1}{2}\right) + \frac{1}{2} \cdot D_{N-3} \cdot V_{REF} \cdot \left(\frac{1}{A^2} - \frac{1}{4}\right) + \ldots$$
$$+\frac{1}{2} \cdot D_1 \cdot V_{REF} \cdot \left(\frac{1}{A^{N-2}} - \frac{1}{2^{N-2}}\right) + \frac{1}{2} \cdot V_{REF} \cdot \left(\frac{1}{A^{N-1}} - \frac{1}{2^{N-1}}\right) + \frac{V_{COS,N}}{A^{N-1}} - \sum_{k=1}^{N} \frac{V_{SOS,K}}{A^{k-1}} \tag{29.76}$$

IPR2023-00698
Theta EX2019

Equations (29.74)–(29.76) are very important to understanding the limitations of the pipeline ADC. Notice the importance of the comparator and summer offsets in Eq. (29.74). The worst-case addition of the offsets must be less than ½ LSB to keep the ADC N-bit accurate. The second stage is more dependent on the gain of the residue amplifier as seen in Eq. (29.75). The gain error discussed in the previous section plays an important role in determining the overall accuracy of the converter. Now examine the effects of the offsets on the INL of the N-th stage. In Eq. (29.76), both the comparator and summer offsets of the N-th stage (when $k = N$) are divided by a large gain. Therefore, the latter stages in a pipeline ADC are not as critical to the accuracy as the first stages, and die area and power can be reduced by using less accurate designs for the least significant stages. The summation term in Eq. (29.76) also reveals that the summer offset of the first stage ($k = 1$) has a large effect on the N-th stage. However, this point is inconsequential since $V_{SOS,1}$ must be minimized to achieve N-bit accuracy for the first stage anyway. Typically, if the INL and DNL specifications can be made N-bit accurate in the first few stages, the latter stages will not adversely affect overall accuracy.

The DNL can be found by calculating the difference between the worst-case switching points and subtracting the ideal value for an LSB. As defined earlier, the worst case will occur at midscale when the output switches from 0111...111 to 1000...000 as v_{IN} increases. Thus, the DNL is

$$DNL_{max} = v'_{IN,1} - v'_{IN,N} - \frac{V_{REF}}{2^N} \tag{29.77}$$

where $v'_{IN,N}$ is calculated using Eq. (29.73) and assuming that D_{N-1} is a zero and that all of the other bits are ones. Plugging in Eqs. (29.69) and (29.73) into Eq. (29.77) yields

$$DNL_{max} = \frac{1}{2}V_{REF}\left(1-\sum_{k=1}^{N-1}\frac{1}{A^k}\right) + V_{COS,1} - \frac{V_{COS,N}}{A^{N-1}} + \sum_{k=2}^{N}\frac{V_{SOS,K}}{A^{k-1}} - \frac{V_{REF}}{2^N} \tag{29.78}$$

Again, the term that dominates this expression is the comparator offset associated with the first stage and the summer offset of the second stage. The entire expression in Eq. (29.78) must be less than ½ LSB for the ADC to have N-bit resolution.

29.2.4 Integrating ADCs

Another type of ADC performs the conversion by integrating the input signal and correlating the integration time with a digital counter. Known as single- and dual-slope ADCs, these types of converters are used in high-resolution applications but have relatively slow conversions. However, they are very inexpensive to produce and are commonly found in slow-speed, cost-conscious applications.

Single-Slope Architecture

Figure 29.32 illustrates the single-slope converter in block level form. A counter determines the number of clock pulses that are required before the integrated value of a reference voltage is equal to the sampled input signal. The number of clock pulses is proportional to the actual value of the input, and the output of the counter is the actual digital representation of the analog voltage.

Since the reference is a DC voltage, the output of the integrator should start at zero and linearly increase with a slope that depends on the gain of the integrator. Notice that the reference voltage is defined as negative so that the output of the inverting

IPR2023-00698
Theta EX2019

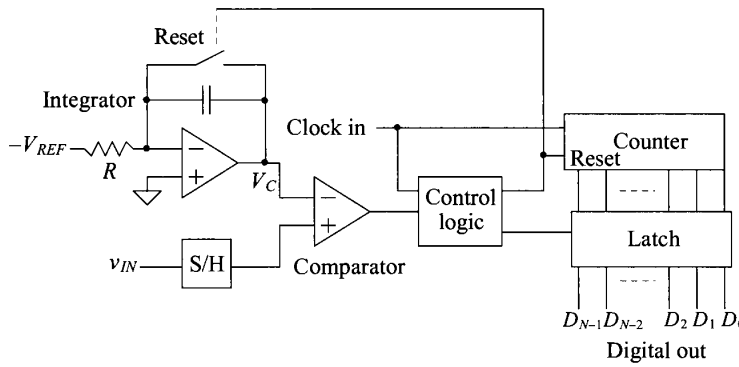


Figure 29.32 Block diagram of a single-slope ADC.

integrator is positive. At the time when the output of the integrator surpasses the value of the S/H output, the comparator switches states, thus triggering the control logic to latch the value of the counter. The control logic also resets the system for the next sample. Figure 29.33 illustrates the behavior of the integrator output and the clock.

Note that if the input voltage is very small, the conversion time is very short, as the counter has to increment only a few times before the comparator latches the data. However, if the input voltage is at its full-scale value, the counter must increment to its maximum value of 2^N clock cycles. Thus, the clock frequency must be many times faster than the bandwidth of the input signal. The conversion time, t_c, depends on the value of the input signal and can be described as

$$t_c = \frac{v_{IN}}{V_{REF}} \cdot 2^N \cdot T_{CLK} \tag{29.79}$$

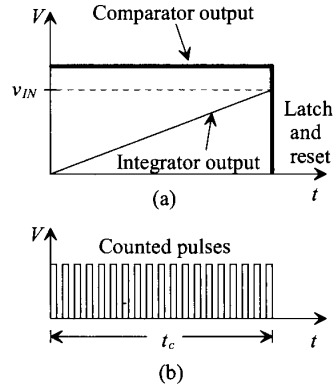


Figure 29.33 Single-slope ADC timing diagrams for (a) the comparator inputs and outputs and (b) the resulting counted pulses.

IPR2023-00698
Theta EX2019

where T_{CLK} is the period of the clock. The sampling rate is inversely proportional to the conversion time and can be written as

$$f_{sample} = \frac{V_{REF}}{V_{IN} \cdot 2^N} \cdot f_{CLK} \tag{29.80}$$

Example 29.14

Determine the clock frequency needed to form an 8-bit, single-slope converter, if the analog signal bandwidth is 20 kHz.

Since the sampling rate required is 40 kHz, then the worst-case situation would occur for a full-scale input, in which event the integrator output would have to climb to its maximum value and the counter would increment 2^N times during the corresponding 25 μs period between samples. Therefore, the clock frequency would need to be 2^N times faster than the sampling rate or 10.24 MHz. ■

Accuracy Issues Related to the Single-Slope ADC

Obviously, many potential error sources abound in this architecture. At the end of the conversion, the voltage across the integrating capacitor, V_C, assuming no initial condition, will be

$$V_C = \frac{1}{C}\int_0^{t_c} \frac{V_{REF}}{R} dt = \frac{V_{REF} \cdot t_c}{RC} \tag{29.81}$$

where t_c is the conversion time. Plugging Eq. (29.79) into Eq. (29.81) yields

$$V_C = \frac{2^N \cdot T_{CLK} \cdot v_{IN}}{RC} = \frac{2^N \cdot v_{IN}}{f_{CLK} \cdot RC} \tag{29.82}$$

Equation (29.82) is a revealing one in that the final voltage on the integrator output depends not only on the value of the input voltage, which is to be expected, but also on the value of R, C, and f_{CLK}. Therefore, any nonideal effects affecting these values will have an influence on the accuracy of the integrator output from sample to sample. For example, if an integrated diffused-resistor is used, then the voltage coefficient of the resistor could limit the accuracy, since the resistor will be effectively nonlinear. Similarly, the capacitor may have charge leakage or aging effects associated with it. Also, any jitter in the clock will affect the overall accuracy. The integrator must have a linear slope to within the accuracy of the converter, which depends on the specifications of the op-amp (open- loop gain, settling time, offset, etc.) and must be considered accordingly.

Offset voltages on the comparator, the S/H, or the integrator result in additional or fewer clock pulses, depending on the polarity of the offset. A delay also exists from the time that the inputs to the comparator are equal and the time that the output of the counter is actually latched. The reference voltage must also stay constant to within the accuracy of the converter.

Dual-Slope Architecture

A slightly more sophisticated design known as the dual-slope, integrating ADC (Fig. 29.34) eliminates most of the problems encountered when using the single-slope converter. Here, two integrations are performed, one on the input signal and one on V_{REF}. The input voltage in this case is assumed to be negative, so that the output of the inverting

IPR2023-00698
Theta EX2019

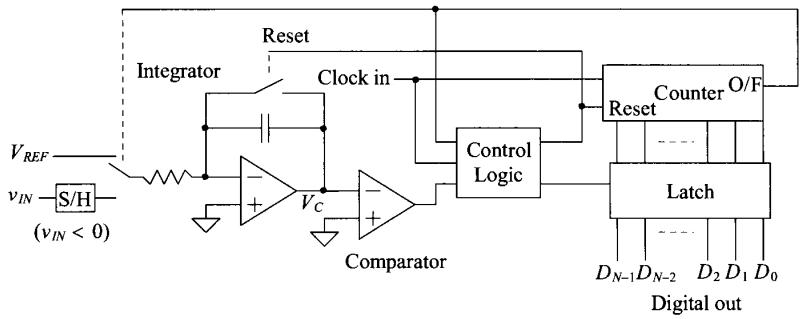


Figure 29.34 Block diagram of a dual-slope ADC.

integrator results in a positive slope during the first integration. Figure 29.35 illustrates the behavior for two separate samples. The first integration is of fixed length, dictated by the counter, in which the sample-and-held signal is integrated, resulting in the first slope. After the counter overflows and is reset, the reference voltage is connected to the input of the integrator. Since v_{IN} was negative and the reference voltage is positive, the inverting integrator output begins discharging back down to zero at a constant slope. A counter again measures the amount of time for the integrator to discharge, thus generating the digital output.

For Fig. 29.35, a 3-bit ADC is being used. Thus, the first integration period continues until the beginning of the eighth (2^3) clock pulse, which corresponds to the counter's overflow bit. Note that the integrator's output corresponding to V_B is twice the value of the output corresponding to V_A. Thus, it requires twice as many clock pulses for

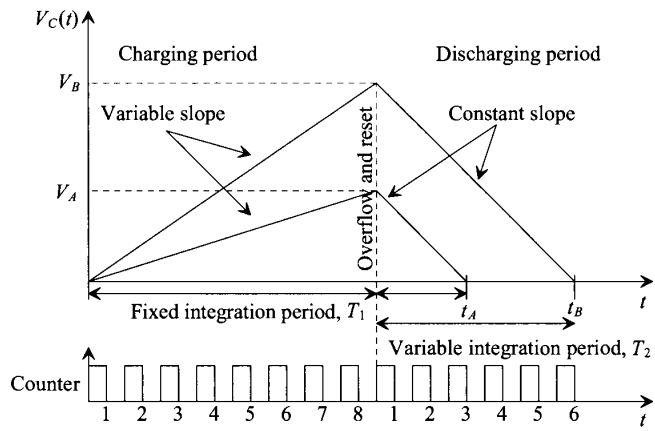


Figure 29.35 Integration periods and counter output for two separate samples of a 3-bit dual-slope ADC.

IPR2023-00698
Theta EX2019

the integrator to discharge back to zero from V_B than from V_A. The output of the counter at t_A is three or 011, while the counter output at t_B is twice that value or six (110) and the quantization is complete.

Notice that the first slope varies according to the value of the input signal, while the second slope, dependent only on V_{REF}, is constant. Similarly, the time required to generate the first slope is constant, since it is limited by the size of the counter. However, the discharging period is variable and results in the digital representation of the input voltage.

Accuracy Issues Related to the Dual-Slope ADC

One may wonder how the dual-slope converter is an improvement over the single-slope architecture, since a significantly longer conversion time is required. The first integration period requires a full 2^N clock cycle and cannot be decreased, because the second integration might require the full 2^N clock cycles to discharge if the maximum value of v_{IN} is being converted. However, the dual slope is the preferred architecture because the same integrator and clock are used to produce both slopes. Therefore, any nonidealities will essentially be canceled. For example, assuming that the S/H is ideal, the gain of the integrator at the end of the first integration period, T_1, becomes

$$V_C = -\frac{1}{C}\int_0^{T_1} \frac{v_{IN}}{R} dt = \frac{|v_{IN}| \cdot T_1}{RC} \tag{29.83}$$

The output at the end of T_1 is positive since the input voltage is considered to be negative and the integrator is inverting. After the clock has been reset, the discharging commences, with the initial condition defined by the value of the integrator output at the end of the charging period, or

$$V_C = \frac{|v_{IN}| \cdot T_1}{RC} - \frac{1}{C}\int_0^{T_2} \frac{V_{REF}}{R} dt \tag{29.84}$$

Once the value of the integrator output, V_C, reaches zero volts, Eq. (29.84) becomes

$$V_C = \frac{|v_{IN}| \cdot T_1}{RC} - \frac{V_{REF} \cdot T_2}{RC} = 0 \tag{29.85}$$

or,

$$|v_{IN}| \cdot T_1 = V_{REF} \cdot T_2 \tag{29.86}$$

At the end of the conversion, the dependencies on R and C have canceled out. Since we also know that the counter increments 2^N times at time, T_1, and the counter increments D times at time, T_2, Eq. (29.86) can be rewritten as

$$\frac{D}{2^N} = \frac{|v_{IN}|}{V_{REF}} \tag{29.87}$$

where D is the counter output that is actually the digital representation of the input voltage. Thus, it can be written that the ratio of the input voltage and the reference voltage is proportional to the ratio of the binary value of the digital word, D, and 2^N. Therefore, since the same clock pulse is responsible for the charging and discharging times, any irregularities will also cancel out.

IPR2023-00698
Theta EX2019

29.2.5 The Successive Approximation ADC

The successive approximation converter performs basically a binary search through all possible quantization levels before converging on the final digital answer. The block diagram is seen in Fig. 29.36. An N-bit register controls the timing of the conversion where N is the resolution of the ADC. V_{IN} is sampled and compared to the output of the DAC. The comparator output controls the direction of the binary search, and the output of the successive approximation register (SAR) is the actual digital conversion. The successive approximation algorithm is as follows.

1. A 1 is applied to the input of the shift register. For each bit converted, the 1 is shifted to the right 1-bit position. $B_{N-1} = 1$ and B_{N-2} through $B_0 = 0$.
2. The MSB of the SAR, D_{N-1}, is initially set to 1, while the remaining bits, D_{N-2} through D_0, are set to 0.
3. Since the SAR output controls the DAC and the SAR output is 100...0, the DAC output will be set to $\frac{V_{REF}}{2}$.
4. Next, v_{IN} is compared to $\frac{V_{REF}}{2}$. If $\frac{V_{REF}}{2}$ is greater than v_{IN}, then the comparator output is a 0 and the comparator resets D_{N-1} to 0. If $\frac{V_{REF}}{2}$ is less than v_{IN}, then the comparator output is a 1 and the D_{N-1} remains a 1. D_{N-1} is the actual MSB of the final digital output code.
5. The 1 applied to the shift register is then shifted by one position so that $B_{N-2} = 1$, while the remaining bits are all 0.
6. D_{N-2} is set to a 1, D_{N-3} through D_0 remain 0, while D_{N-1} remains the value from the MSB conversion. The output of the DAC will now either equal $\frac{V_{REF}}{4}$ (if $D_{N-1} = 0$) or $\frac{3V_{REF}}{4}$ (if $D_{N-1} = 1$).

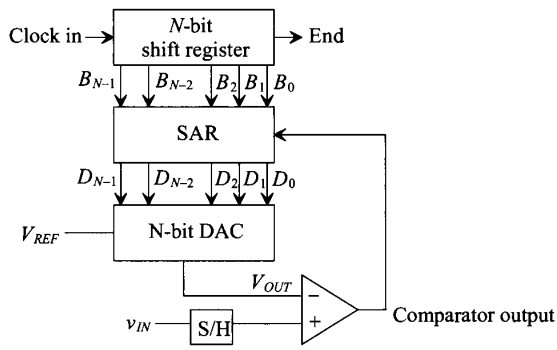


Figure 29.36 Block diagram of the successive approximation ADC.

IPR2023-00698
Theta EX2019

7. Next, v_{IN} is compared to the output of the DAC. If the DAC output is greater than v_{IN} then the comparator output drives D_{N-2} to 0. If the DAC output is less than v_{IN} then D_{N-2} remains a 1.

8. The process repeats until the output of the DAC converges to the value of v_{IN} within the resolution of the converter.

Figure 29.37 shows an example of the binary search nature of the converter. The bolded line shows the path of the conversion for 101, corresponding to $\frac{5}{8}V_{REF}$. All possible quantization levels are represented in the binary tree. With each bit decided, the search space decreases by one-half until the correct answer is converged upon.

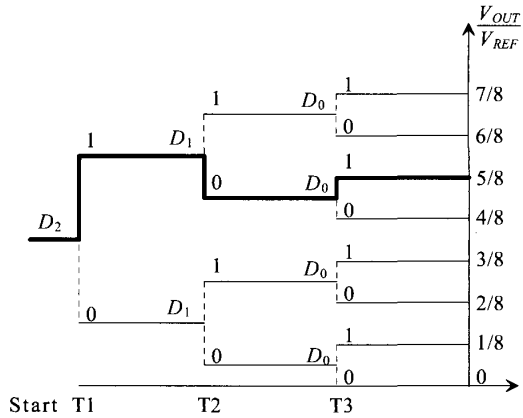


Figure 29.37 Binary search performed by a 3-bit successive approximation ADC for D=101.

Example 29.15

Perform the operation of a 3-bit successive approximation ADC similar to Fig. 29.36 with V_{REF} = 8. Make a table that consists of $D_2D_1D_0$, $B_2B_1B_0$, V_{OUT} (the output from the DAC) and the comparator output, which shows the binary search algorithm of the converter for v_{IN} = 5.5 V and 2.5 V.

We will designate $D_2D_1D_0$ as the initial output of the SAR before the comparator makes its decision. The final value is designated as $D_2D_1D_0$. Notice that if the comparator is a 1, $D_2D_1D_0$ differs from $D_2D_1D_0$, but if the comparator outputs a 0, then $D_2D_1D_0$ = $D_2D_1D_0$. The output of the shift register is designated as $B_2B_1B_0$.

Following the algorithm discussed previously, initially v_{IN} = 5.5 V and is compared with 4 V. Since the comparator output is 0, the MSB remains a 1. The next bit is examined, and the output of the DAC is now 6 V. Since $V_{OUT} > v_{IN}$, the comparator output is 1, which resets the current SAR bit, D_1, to a 0 at the end of period T2. Lastly, the LSB is examined, and v_{IN} is compared with 5 V. Since $v_{IN} > V_{OUT}$, the comparator output is a 0, and the current SAR bit, D_0, remains a 1. The

IPR2023-00698
Theta EX2019

results can be examined in Fig. 29.38a. The final value for $D_2D_1D_0$ is 101, which is what is expected considering that 101 in binary is equivalent to 5_{10}. Figure 29.38b shows the data for the ADC using v_{IN} = 2.5 V. The final value for v_{IN} = 2.5 is 010, which again is what is expected for 3-bit resolution. ■

Step	v_{IN}	$B_2B_1B_0$	$D_2D_1D_0$	V_{OUT}	Comp Out	$D_2D_1D_0$
T1	5.5	100	100	$1/2\ V_{REF} = 4$ V	0	100
T2	5.5	010	110	$(1/2+1/4)V_{REF} = 6$ V	1	100
T3	5.5	001	101	$(1/2+1/8)V_{REF} = 5$ V	0	101

(a)

Step	v_{IN}	$B_2B_1B_0$	$D'_2D'_1D'_0$	V_{OUT}	Comp Out	$D_2D_1D_0$
T1	2.5	100	100	$1/2\ V_{REF} = 4$ V	1	000
T2	2.5	010	010	$1/4\ V_{REF} = 2$ V	0	010
T3	2.5	001	011	$(1/4+1/8)V_{REF} = 3$ V	1	010

(b)

Figure 29.38 Results from the 3-bit successive approximation ADC using (a) v_{IN} = 5.5 and (b) 2.5 V.

The successive approximation ADC is one of the most popular architectures used today. The simplicity of the design allows for both high speed and high resolution while maintaining relatively small area. The limit to the ADC's accuracy depends mainly on the accuracy of the DAC. If the DAC does not produce the correct analog voltage with which to compare the input voltage, the entire converter output will contain an error. Referring again to Fig. 29.37, we can see that if a wrong decision is made early, a massive error will result as the converter attempts to search for the correct quantization level in the wrong half of the binary tree.

The Charge-Redistribution Successive Approximation ADC

One of the most popular types of successive approximation architectures uses the binary-weighted capacitor array (analyzed in Sec. 29.1.5) as its DAC. Called a charge-redistribution, successive-approximation ADC, this converter samples the input signal and then performs the binary search based on the amount of charge on each of the DAC capacitors. Figure 29.39 shows an N-bit architecture. A comparator has replaced the unity gain buffer used in the DAC architecture. The binary-weighted capacitor array also samples the input voltage, so no external sample-and-hold is needed.

The conversion process begins by discharging the capacitor array, via the reset switch. Although this may appear to be an insignificant action, the converter is also performing automatic offset cancellation. Once the reset switch is closed, the comparator acts as a unity gain buffer. Thus, the capacitor array charges to the offset voltage of the comparator. This requires that the comparator is designed to be unity-gain stable, which means that internal compensation may have to be switched in during the reset period. Next, the input voltage, v_{IN}, is sampled onto the capacitor array. The reset switch is still closed, for the top plate of the capacitor array needs to be connected to virtual ground of the unity gain buffer. The equivalent circuit is seen in Fig. 29.40a. The reset switch is

IPR2023-00698
Theta EX2019

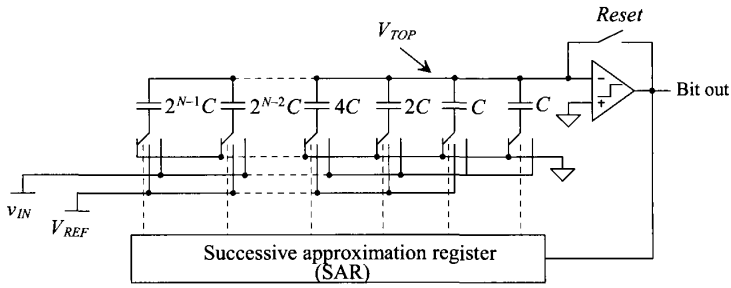


Figure 29.39 A charge redistribution ADC using a binary-weighted capacitor array DAC.

then opened, and the bottom plates of each capacitor in the array are switched to ground, so that the voltage appearing at the top plate of the array is now $V_{OS} - v_{IN}$ (Fig. 29.40b). The conversion process begins by switching the bottom plate of the MSB capacitor to V_{REF} (Fig. 29.40c). If the output of the comparator is high, the bottom plate of the MSB capacitor remains connected to V_{REF}. If the comparator output is low, the bottom plate of the MSB is connected back to ground. The output of the comparator is D_{N-1}. The voltage at the top of the capacitor array, V_{TOP}, is now

$$V_{TOP} = -v_{IN} + V_{OS} + D_{N-1} \cdot \frac{V_{REF}}{2} \tag{29.88}$$

The next largest capacitor is tested in the same manner as seen in Fig. 29.40d. The voltage at the top plate of the capacitor after the second capacitor is tested becomes

$$V_{TOP} = -v_{IN} + V_{OS} + D_{N-1} \cdot \frac{V_{REF}}{2} + D_{N-2} \cdot \frac{V_{REF}}{4} \tag{29.89}$$

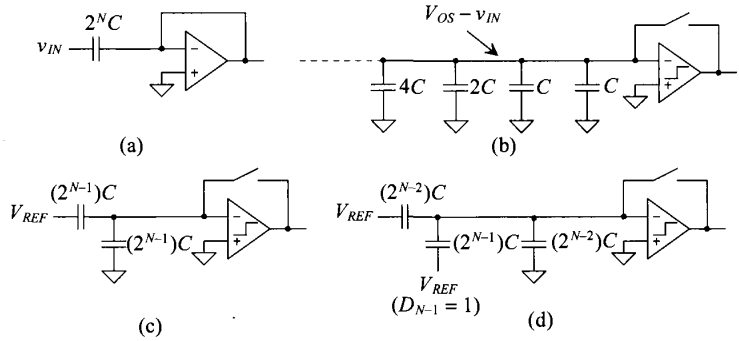


Figure 29.40 The charge redistribution process: (a) Sampling the input while autozeroing the offset, (b) the voltage at the top plate after sampling, (c) the equivalent circuit while converting the MSB, and (d) the equivalent circuit while converting the next largest capacitor with the MSB result equal to one.

IPR2023-00698
Theta EX2019

The conversion process continues on with the remaining capacitors so that the voltage on the top plate of the array, V_{TOP}, converges to the value of the offset voltage, V_{OS} (within the resolution of the converter), or

$$V_{TOP} = -v_{IN} + V_{OS} + D_{N-1} \cdot \frac{V_{REF}}{2} + D_{N-2} \cdot \frac{V_{REF}}{4} + ... + D_1 \cdot \frac{V_{REF}}{2^{N-1}} + D_0 \cdot \frac{V_{REF}}{2^N} \approx V_{OS} \quad (29.90)$$

Note that the initial charge stored on the capacitor array is now redistributed onto only those capacitors that have their bottom plates connected to V_{REF}.

Accuracy Issues Related to the Charge-Redistribution, Successive-Approximation ADC

Obviously, the limitation of this architecture is the capacitor matching. The mismatch is analyzed in the same manner as the binary-weighted current source array of Sec. 29.1.4. Thus, substituting the value of the unit capacitance, C, for the value of the unit current source, I, and using Eqs. (29.27) – (29.30),

$$|INL|_{max} = \frac{2^{N-1} \cdot V_{REF} \cdot (C + |\Delta C|_{max,INL})}{2^N \cdot C} - \frac{2^{N-1} \cdot V_{REF} \cdot C}{2^N \cdot C} = \frac{V_{REF}}{2} \cdot \frac{|\Delta C|_{max,INL}}{C} \quad (29.91)$$

where the maximum ΔC that will result in an INL that is ½ LSB is

$$|INL|_{max} = \frac{V_{REF}}{2} \cdot \frac{|\Delta C|_{max,INL}}{C} = \frac{V_{REF}}{2^{N+1}} = \frac{1}{2} \text{ LSB} \rightarrow |\Delta C|_{max,INL} = \frac{C}{2^N} \quad (29.92)$$

The DNL is defined by

$$DNL_{max} = \frac{(2^N - 1) \cdot V_{REF} |\Delta C|_{max,DNL}}{2^N \cdot C} \quad (29.93)$$

with the maximum ΔC, which results in a DNL less than ½ LSB:

$$DNL_{max} = \frac{(2^N - 1) \cdot V_{REF} |\Delta C|_{max,DNL}}{2^N \cdot C} = \frac{V_{REF}}{2^{N+1}} = \frac{1}{2} \text{ LSB} \rightarrow |\Delta C|_{max,DNL} = \frac{C}{2^{N+1} - 2} \quad (29.94)$$

29.2.6 The Oversampling ADC

ADCs can be separated into two categories depending on the rate of sampling. The first category samples the input at the Nyquist rate, or $f_N = 2F$ where F is the bandwidth of the signal and f_N is the sampling rate. The second type samples the signal at a rate much higher than the signal bandwidth. This type of converter is called an oversampling converter.

The oversampling ADC is able to achieve much higher resolution than the Nyquist rate converters. This is because digital signal processing techniques are used in place of complex and precise analog components. The accuracy of the converter does not depend on the component matching, precise sample-and-hold circuitry, or trimming, and only a small amount of analog circuitry is required. Switched-capacitor implementations are easily achieved, and, as a result of the high sampling rate, only simplistic anti-aliasing circuitry needs to be used. However, because of the amount of time required to sample the input signal, the throughput is considerably less than the Nyquist rate ADCs.

IPR2023-00698
Theta EX2019

Differences in Nyquist Rate and Oversampling ADCs

The typical process used in analog-to-digital conversion is seen in Fig. 29.41a, while the block diagram for the oversampling ADC is seen in Fig. 29.41b. After filtering the signal to help minimize aliasing effects, the signal is sampled, quantized, and encoded or decoded using simple digital logic to provide the digital data in the proper format. When using oversampling ADCs, little if any, anti-alias filtering is needed, no dedicated S/H is required, the quantization is performed with a modulator, and the encoding usually takes the form of a digital filter.

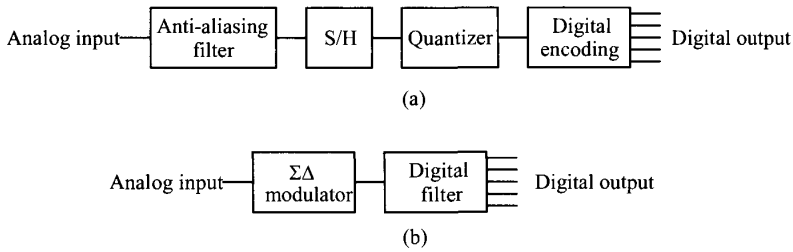


Figure 29.41 Typical block diagram for (a) Nyquist rate converters and (b) oversampling ADCs.

Since the oversampling converter samples the signal bandwidth at many times, aliasing is not a serious problem. A discussion of the frequency characteristics of aliasing was presented in Ch. 28. Figure 29.42a shows that when using Nyquist rate converters, a sampled signal in the frequency domain appears as a series of band-limited signals at multiples of the sampling frequency (see Fig. 28.26 for more details). As the sampling frequency decreases, the frequency spectra begin to overlap, and aliasing (Fig. 29.42b) occurs. Complex, "brickwall" filters are needed to correct the problem.

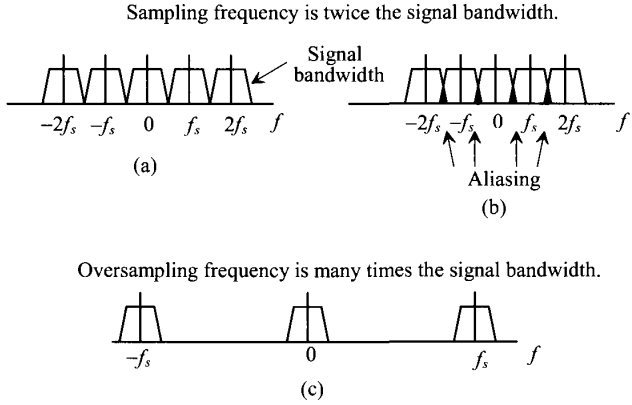


Figure 29.42 Frequency domain for (a) Nyquist rate converters, (b) the aliasing that occurs, and (c) an oversampling converter.

IPR2023-00698
Theta EX2019

For oversampled ADCs, aliasing becomes much less of a factor. Since the sampling rate is much greater than the bandwidth of the signal, the frequency domain representation shows that the spectra are widely spaced, as seen in Fig. 29.42c. Therefore, overlapping of the spectra, and thus aliasing, will not occur, and only simple, first-order filters are required.

Oversampling converters typically employ switched-capacitor circuits and therefore do not need sample-and-hold circuits. The output of the modulator is a pulse-density modulated signal that represents the average of the input signal. The modulator constructs these pulses in real time, and so it is not necessary to hold the input value and perform the conversion.

As stated previously, the modulator actually provides the quantization in the form of a pulse-density modulated signal. Referred to as sigma-delta ($\Sigma\Delta$) or delta-sigma ($\Delta\Sigma$) modulation, the density of the pulses represents the average value of the signal over a specific period. Figure 29.43 illustrates the output of the modulator for the positive half of a sine wave input. Note that for the peak of the sine wave, most of the pulses are high. As the sine wave decreases in value, the pulses become distributed between high and low according to the sine wave value.

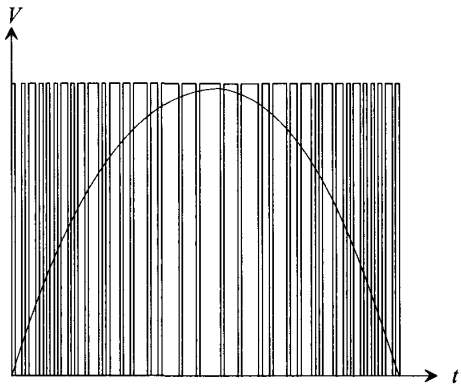


Figure 29.43 Pulse-density output from a sigma-delta modulator for a sine wave input.

If the frequency of the sine wave represented the highest frequency component of the input signal, a Nyquist rate converter would take only two samples. The oversampling converter, however, may take hundreds of samples over the same period to produce this pulse-density signal.

Digital signal processing is then used, which has two purposes: to filter any out-of-band quantization noise and to attenuate any spurious out-of-band signals. The output of the filter is then downsampled to the Nyquist rate so that the resulting output of the ADC is the digital data. This data represents the average value of the analog voltage over the oversampling period. The effective resolution of oversampling converters is determined by the values of signal-to-noise ratio and dynamic range obtained.

IPR2023-00698
Theta EX2019

The First-Order ΣΔ Modulator

Now that the basic function of the ΣΔ modulator has been described, it would be useful to examine its inner workings and determine why ΣΔ modulation is so beneficial for generating high-resolution data. A basic first-order ΣΔ modulator can be seen in Fig. 29.44. Here, an integrator and a 1-bit ADC are in the forward path, and a 1-bit DAC is in the feedback path of a single-feedback loop system. The variables labeled are in terms of time, T, which is the inverse of the sampling frequency and k, which is an integer. The 1-bit ADC is simply a comparator that converts an analog signal into either a high or a low. The 1-bit DAC uses the comparator output to determine if $+V_{REF}$ or $-V_{REF}$ is summed with the input.

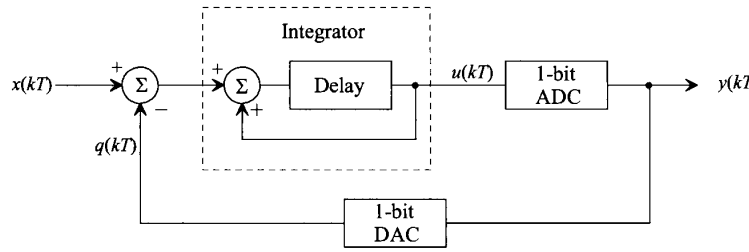


Figure 29.44 A first-order sigma-delta modulator.

While the benefits of ΣΔ modulation are not obvious, a simple derivation of the output, $y(kT)$, illuminates its distinct advantages. The output of the integrator, $u(kT)$, can be described as

$$u(kT) = x(kT - T) - q(kT - T) + u(kT - T) \tag{29.95}$$

where, $x(kT - T) - q(kT - T)$ is equal to the integrator's previous input, and $u(kT - T)$ is its previous output. The quantization error for the 1-bit ADC, as discussed in Ch. 28, is again defined as the difference between its output and input such that

$$Q_e(kT) = y(kT) - u(kT) \tag{29.96}$$

Plugging Eq. (29.95) into Eq. (29.96), the output response, $y(kT)$ is

$$y(kT) = Q_e(kT) + x(kT - T) - q(kT - T) + u(kT - T) \tag{29.97}$$

An ideal 1-bit DAC has the following characteristic: if the input, $y(kT) = 0$, the output, $q(kT) = -V_{REF}$, and if $y(kT) = 1$, then $q(kT) = V_{REF}$. In reality, a 1-bit DAC consists of a couple of switches connecting V_{REF} or $-V_{REF}$ to a common node, so it is not difficult to assume that the DAC is ideal. Therefore,

$$y(kT) = q(kT) \tag{29.98}$$

Utilizing Eq. (29.96) and Eq. (29.97), we find that Eq. (29.98) becomes

IPR2023-00698
Theta EX2019

$$y(kT) = x(kT - T) + Q_e(kT) - Q_e(kT - T) \tag{29.99}$$

Therefore, the output of the modulator consists of a quantized value of the input signal delayed by one sample period, plus a differencing of the quantization error between the present and previous values. Thus, the real power of ΣΔ modulation is that the quantization noise, Q_e, cancels itself out to the first order.

A frequency domain example further illuminates this important fact. Suppose that the first-order modulator can be modeled in the s domain, as seen in Fig. 29.45, with an ideal integrator represented with transfer function of $\frac{1}{s}$, the 1-bit ADC modeled as a simple error source, $Q_e(s)$, and again the DAC considered to be ideal, such that $y(s)$ is equal to $q(s)$. It is also assumed that the bandwidth of the input signal is much less than the bandwidth of the modulator. Therefore, using simple feedback theory, $v_{OUT}(s)$ becomes

$$v_{OUT}(s) = Q_e(s) + \frac{1}{s} \cdot [v_{IN}(s) - v_{OUT}(s)] \tag{29.100}$$

and solving for v_{OUT} yields,

$$v_{OUT}(s) = Q_e(s) \cdot \frac{s}{s+1} + v_{IN}(s)\frac{1}{s+1} \tag{29.101}$$

Note that the transfer function from v_{IN} to v_{OUT} follows that of a low-pass filter and that the transfer function of the quantization noise follows that of a high-pass filter. Plotted together in Fig. 29.46, it is seen that in the region where the signal is of interest, the noise has a small value while the signal has a high gain, and that at higher frequencies, beyond the bandwidth of the signal, the noise increases. The modulator has essentially pushed the power of the noise out of the bandwidth of the signal. This high-pass characteristic is known as *noise shaping* and is a powerful concept used within oversampling ADCs. Low-pass filtering is then performed by the digital filter in order to remove all of the out-of-band quantization noise, which then permits the signal to be downsampled to yield the final high-resolution output.

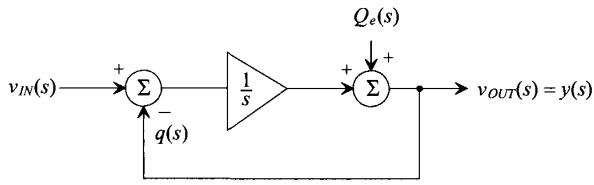


Figure 29.45 A frequency domain model for the first-order sigma-delta modulator.

As the ΣΔ modulator is generating the pulse-density modulated output, it is interesting to examine the mechanics occurring in the loop, which result in an average of the input. An actual ΣΔ modulator might resemble Fig. 29.47. A switched-capacitor integrator provides the summing as well as the delay needed. The 1-bit ADC is a simple comparator, and the 1-bit DAC is simply two voltage-controlled switches that select either V_{REF} or $-V_{REF}$ to be summed with the input. A latched comparator provides the necessary loop delay. Notice that the variables are voltage representations of the variables

IPR2023-00698
Theta EX2019

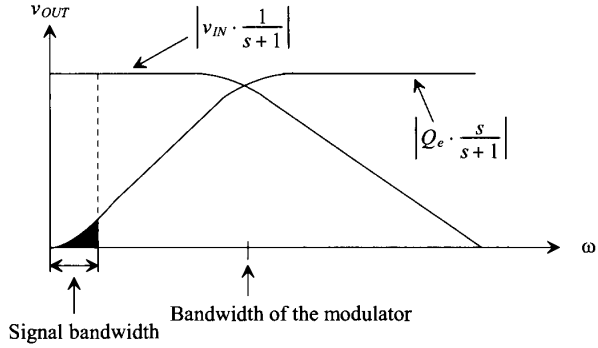


Figure 29.46 Frequency response of the first-order sigma-delta modulator.

used in Fig. 29.44. Remember that the function of the integrator is to accumulate differences between the input signal and the output of the DAC. If it is assumed that the input, $v_x(kT)$, is a positive DC voltage, then the output of the integrator should increase. However, the feedback mechanism is such that the 1-bit ADC (the comparator) has a low output if the integrator output, $v_u(kT)$, is positive. Thus, V_{REF} appears at the output of the DAC and is subtracted from the input, and the integrator output is driven back toward zero. The opposite occurs when $v_u(kT)$ is negative such that the integrator output is always driven toward zero by the feedback mechanism. An example will illustrate the operation of the modulator in more detail.

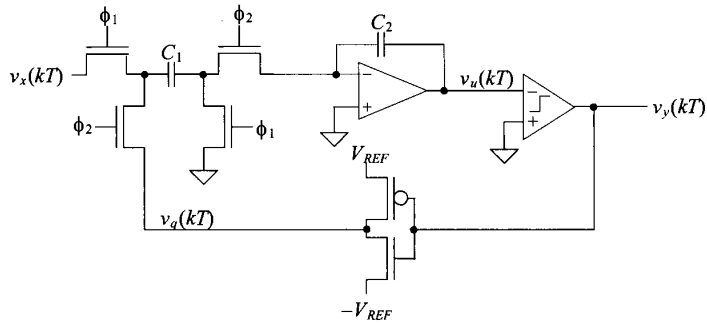


Figure 29.47 Implementation of a first-order sigma-delta modulator using a switched capacitor integrator.

Example 29.16

Using a general first-order ΣΔ modulator, assume that the input to the modulator, $v_x(kT)$ is a positive DC voltage of 0.4 V. Show the values of each variable around the ΣΔ modulator loop and prove that the overall average output of the DAC

IPR2023-00698
Theta EX2019

approaches 0.4 V after 10 cycles. Assume that the DAC output is ±1 V, and that the integrator output has a unity gain with an initial output voltage of 0.1 V, and that the comparator output is either ±1 V.

The present integrator output will be equal to the sum of the previous integrator output and the previous integrator input. Therefore, Eq. (29.95) becomes

$$v_u(kT) = v_u(kT - T) + v_a(kT - T) \tag{29.102}$$

where

$$v_a(kT) = v_x(kT) - v_q(kT) \tag{29.103}$$

and the quantizing error, $Q_e(kT)$, is defined by Eqs. (29.96) and (29.98) as

$$Q_e(kT) = v_q(kT) - v_u(kT) \tag{29.104}$$

The initial conditions define the values of the variable for $k = 0$. The output of the integrator is given to be 0.1 V. Thus, the ADC output is low, the DAC output is V_{REF}, and the output of the summer, $v_a(0)$, is $0.4 - V_{REF} = -0.6$ V.

The output for $k = 1$ begins again with the integrator output. Using Eq. (29.102), $v_u(kT)$ becomes

$$v_u(T) = 0.1 + (-0.6) = -0.5 \text{ V}$$

Since the output of the integrator is negative, the output of the comparator is positive and $-V_{REF}$ is subtracted from 0.4 to arrive at the value for $v_a(T)$.

Continuing in the same manner and using the previous equations, we note the voltages for each cycle in Fig. 29.48. After 10 cycles through the modulator, the average value of $v_q(kT)$ becomes,

$$\overline{v_q(kT)} = \frac{7-3}{10} = 0.4\ V$$

k	$v_a(kT)$	$v_u(kT)$	$v_q(kT) = v_y(kT)$	$Q_e(kT)$	$\overline{v_q(kT)}$
0	–0.6	0.1	1.0	0.9	1.0
1	1.4	–0.5	–1.0	–0.5	0
2	–0.6	0.9	1.0	0.1	0.333
3	–0.6	0.3	1.0	0.7	0.50
4	1.4	–0.3	–1.0	–0.7	0.20
5	–0.6	1.1	1.0	–0.1	0.333
6	–0.6	0.5	1.0	0.5	0.429
7	1.4	–0.1	–1.0	–0.9	0.25
8	–0.6	1.3	1.0	–0.3	0.333
9	–0.6	0.7	1.0	0.3	0.40

Figure 29.48 Data from the first-order ΣΔ modulator.

IPR2023-00698
Theta EX2019

Notice that the behavior of $\overline{v_q(kT)}$ swings around the desired value 0.4 V. If we were to continue computing values, as k increases, the amount that $\overline{v_q(kT)}$ differs from 0.4 V would decrease. Ideally, we could make the deviation of $\overline{v_q(kT)}$ as small as desired by allowing the modulator to take as many samples as necessary to meet that accuracy. ■

It is interesting to examine the effects of using a nonideal comparator. Suppose the integrator's output was smaller than the offset voltage of the comparator. A wrong decision would be made, causing $v_y(kT)$ to be the opposite of the desired value. However, as k increases, this error is averaged out, and the modulator still converges on the correct answer. Therefore, the comparator does not have to be very accurate in its ability to distinguish between two voltages, in contrast to Nyquist rate comparators.

The Higher Order ΣΔ Modulators

Higher order ΣΔ modulators exist which provide a greater amount of noise shaping. A second-order ΣΔ modulator can be seen in Fig. 29.49. A derivation of the second-order transfer function would reveal that the output contained a delayed version of the input plus a second-order differencing of the quantization noise, Q_e (see Problem 29.38). A third-order modulator would contain third-order differencing of the quantization and can be constructed by adding another integrator similar to integrator A into the system.

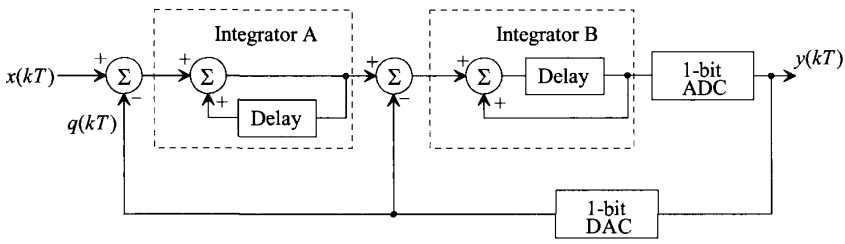


Figure 29.49 A second-order, sigma-delta modulator.

Figure 29.50 shows the noise-shaping functions of a first-, second-, and third-order modulator. The cross-hatched area under each of the curves represents the noise that remains in the signal bandwidth and is a magnified version of the blackened area of Fig. 29.46. As the order increases, notice that more of the noise is pushed out into the higher frequencies, thus decreasing the noise in the signal bandwidth. It should be reiterated that ΣΔ modulators do not attenuate noise at all. In fact, they add quantization noise that is very large at high frequencies. But because almost all of the noise is out of the signal bandwidth, it can easily be filtered, leaving only a small portion within the signal bandwidth. This point is important because the ΣΔ modulator should not be construed as a filtering circuit.

The resolution also increases as the order of the ΣΔ modulator and the oversampling ratio increases, as seen in Fig. 29.51. Using a first-order modulator, one can expect an increase in dynamic range of 9 dB with every doubling of the oversampling

IPR2023-00698
Theta EX2019

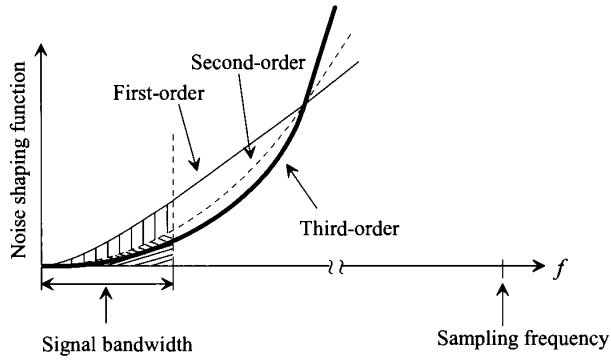


Figure 29.50 Noise shaping comparison of a first-, second- and third-order modulator.

ratio. This correlates to an approximate increase in resolution of 1.5 bits according to Eq. (28.28). The higher-order modulators have even greater gains in resolution as a 2.5-bit increase is attained with each doubling of the oversampling ratio using a second-order modulator, while the third-order modulator increases 3.5 bits.

One could essentially construct a high-order ΣΔ modulator with many integrators. However, as with any system employing feedback, stability becomes a critical issue. The same holds true for the high-order ΣΔ modulators. Several other topologies have been developed which can implement modulators in a cascaded fashion and are guaranteed to be stable. However, considerable matching requirements need to be overcome.

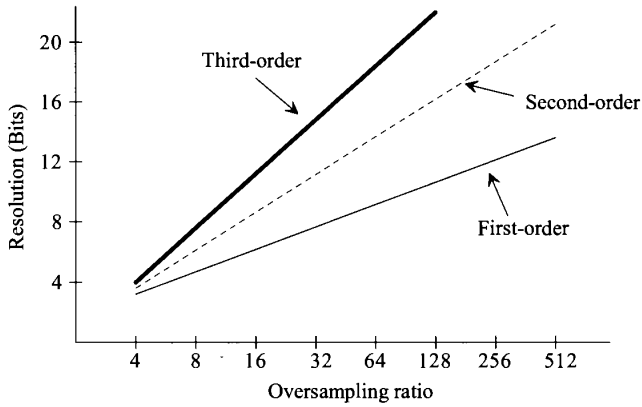


Figure 29.51 Comparison of first-, second-, and third-order modulators versus oversampling ratio and resolution.

IPR2023-00698
Theta EX2019

ADDITIONAL READING

[1] R. J. Baker, *CMOS: Mixed-Signal Circuit Design, Second Edition*, Wiley-IEEE Press, 2009.

[2] R. Schreier and G. C. Temes, *Understanding Delta-Sigma Data Converters*, Wiley-IEEE Press, 2005. ISBN 978-0471465850

[3] R. J. van de Plassche, *CMOS Integrated Analog-to-Digital and Digital-to-Analog Converters*, Second Edition, Springer, 2003.

[4] J. W. Bruce, "Nyquist-rate digital to analog converter architectures," *IEEE Potentials*, vol. 20, no. 3, pp. 24–28, August 2001. Good overview of digital to analog converters.

[5] M. Gustavsson, J. J. Wikner, and N. Tan, *CMOS Data Converters for Communications*, Springer, 2000. ISBN 978-0792377801

[6] J. W. Bruce, "Meeting the analog world challenge: Nyquist rate analog to digital converter architectures," *IEEE Potentials,* vol. 17, no. 5, pp. 36–39, January 1999. Very good introduction to ADCs.

[7] S. R. Norsworthy, R. Schreier, and G. C. Temes (eds.), *Delta-Sigma Data Converters: Theory, Design, and Simulation*, Wiley-IEEE Press, 1997. ISBN 978-0780310452

[8] M. Ismail and T. Fiez, *Analog VLSI Signal and Information Processing*, McGraw-Hill, 1994.

[9] M. J. M. Pelgrom, et. al, "25-Ms/s 8-bit CMOS A/D Converter for Embedded Application," *IEEE Journal of Solid-State Circuits*, vol. 29, no. 8, pp. 879–886, August 1994.

[10] D. Choi, et. al, "Analog Front-End Signal Processor for a 64 Mbits/s PRML Hard-Disk Drive Channel," *IEEE Journal of Solid-State Circuits*, vol. 29, no. 12, pp. 1596–1605, December 1994.

[11] M. Yotsuyanagi, T. Etoh, and K. Hirata, "A 10 Bit 50 MHz Pipelined CMOS A/D Converter with S/H," *IEEE Journal of Solid State Circuits*, vol. 28, no. 3, pp. 292–300, March 1993.

[12] P. Vorenkamp and J. P. M. Verdaasdonk, "A 10 b 50 Ms/s Pipelined ADC," *IEEE ISSCC Digest of Technical Papers*, pp. 34–35, February 1992.

[13] B. Razavi and B. A. Wooley, "A 12-b, 5-MSample/s Two-Step CMOS A/D Converter," *IEEE Journal of Solid State Circuits*, vol. 27, no. 12, pp. 1667–1678, December 1992.

[14] N. Shiwaku, "A Rail-to-Rail Video-band Full Nyquist 8-bit A/D Converter," *Proceedings of the 1991 Custom Integrated Circuits Conference.*

[15] B. P. Brandt, *Oversampled Analog-to-Digital Conversion,* Integrated Circuits Laboratory, Technical Report No. ICL91-009, Stanford University, 1991.

[16] R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI - Design Techniques for Analog and Digital Circuits*, McGraw-Hill Publishing Co., 1990.

IPR2023-00698
Theta EX2019

[17] B. S. Song, S. H. Lee, and M. F. Tompsett, "A 10-bit 15 MHz CMOS Recycling Two-Step A/D Converter," *IEEE Journal of Solid State Circuits*, vol. 25, no. 12, pp. 1328–1338, December 1990.

[18] T. Shimizu, et al., "A 10-bit, 20 MHz Two-Step Parallel A/D Converter with Internal S/H," *IEEE Journal of Solid State Circuits*, vol. 24, no. 1, pp. 13–20, February 1989.

[19] J. Dornberg, P. R. Gray, and D. A. Hodges, "A 10-bit, 5-Msample/s CMOS Two-Step Flash ADC," *IEEE Journal of Solid State Circuits*, vol. 24, no. 2, pp. 241–249, April 1989.

[20] B. E. Boser, "Design and Implementation of Oversampled Analog-to-Digital Converters," Ph.D. Dissertation, Stanford University, 1988.

[21] S. Sutarja and P. R. Gray, "A Pipelined 13-bit, 250-ks/s, 5-V Analog-to-Digital Converter," *IEEE Journal of Solid State Circuits*, vol. 23, no. 6, pp. 1316–1323, December 1988.

[22] K. Uchimura et al, "Oversampling A-to-D and D-to-A Converters with Multistage Noise Shaping Modulators," *IEEE Transactions on Acoustics, Speech and Signal Processing*, pp. 1899–1905, December 1988.

[23] B. S. Song, M. F. Tompsett, and K. R. Lakshmikumar, "A 12-bit, 1-MSample/s Capacitor Error-Averaging Pipelined A/D Converter," *IEEE Journal of Solid State Circuits*, vol. 23, no. 6, pp. 1324–1333, December 1988.

[24] Y. Matsuya, K. Uchimura, et al, "A 16-bit Oversampling A/D Conversion Technology Using Triple Integration Noise Shaping," *IEEE Journal of Solid State Circuits*, vol. 22, no. 6, pp. 921–929, December 1987.

[25] S. H. Lewis and P. R. Gray, "A Pipelined 5-Msample/s 9 bit Analog-to-Digital Converter," *IEEE Journal of Solid State Circuits*, vol. 22, no. 6, pp. 954–961, December 1987.

[26] K. Bacrania, "A 12 Bit Successive-Approximation ADC with Digital Error Correction," *IEEE Journal of Solid State Circuits*, vol. 21, no. 6, pp. 1016–1025, December 1986.

[27] J. Shyu, G. C. Temes, and F. Krummenacher, "Random Errors in MOS Capacitors and Current Sources," *IEEE Journal of Solid State Circuits*, vol. 16, no. 6, pp. 948–955, December 1984.

[28] P. R. Gray, D. A. Hodges, and R. W. Brodersen (eds.), *Analog MOS Integrated Circuits*, Wiley-IEEE, 1980. ISBN 0-471-08964-8

[29] J. L. McCreary and P. R. Gray, "All-MOS Charge Redistribution Analog-to-Digital Conversion Techniques - Part I," *IEEE Journal of Solid State Circuits*, vol. 10, no. 6, pp. 371–379, December 1975.

[30] R. E. Suarez, P. R. Gray, and D. A. Hodges, "All-MOS Charge Redistribution Analog-to-Digital Conversion Techniques - Part II," *IEEE Journal of Solid State Circuits*, vol. 10, no. 6, pp. 379–385, December 1975.

IPR2023-00698
Theta EX2019

PROBLEMS

29.1 A 3-bit, resistor-string DAC similar to the one shown in Fig. 29.2a was designed with a desired resistor of 500 Ω. After fabrication, mismatch caused the actual value of the resistors to be

$R_1 = 500, R_2 = 480, R_3 = 470, R_4 = 520, R_5 = 510, R_6 = 490, R_7 = 530, R_8 = 500$

Determine the maximum INL and DNL for the DAC assuming $V_{REF} = 5$ V.

29.2 An 8-bit resistor string DAC similar to the one shown in Fig. 29.2b was fabricated with a nominal resistor value of 1 kΩ. If the process was able to provide matching of resistors to within 1%, find the effective resolution of the converter. What is the maximum INL and DNL of the converter? Assume that $V_{REF} = 5$ V.

29.3 Compare the digital input codes necessary to generate all eight output values for a 3-bit resistor string DAC similar to those shown in Fig. 29.2a and b. Design a digital circuit that will allow a 3-bit binary digital input code to be used for the DAC in Fig. 29.2a. Discuss the advantages and disadvantages of both architectures.

29.4 Plot the transfer curve of a 3-bit *R*-2*R* DAC if all *R*s = 1.1 kΩ and 2*R*s = 2 kΩ. What is the maximum INL and DNL for the converter? Assume all of the switches to be ideal and $V_{REF} = 5$ V.

29.5 Suppose that a 3-bit *R*-2*R* DAC contained resistors that were perfectly matched and that $R = 1$ kΩ and $V_{REF} = 5$ V. Determine the maximum switch resistance that can be tolerated for which the converter will still have 3-bit resolution. What are the values of INL and DNL?

29.6 The circuit illustrated in Fig. 29.5 is known as a current-mode *R*-2*R* DAC, since the output voltage is defined by the current through R_F. Shown in Fig. 29.52 is an *N*-bit voltage-mode *R*-2*R* DAC. Design a 3-bit voltage mode DAC and determine the output voltage for each of the eight input codes. Label each node voltage for each input. Assume that $R = 1$ kΩ and that $R_2 = R_1 = 10$ kΩ and $V_{REF} = 5$ V.

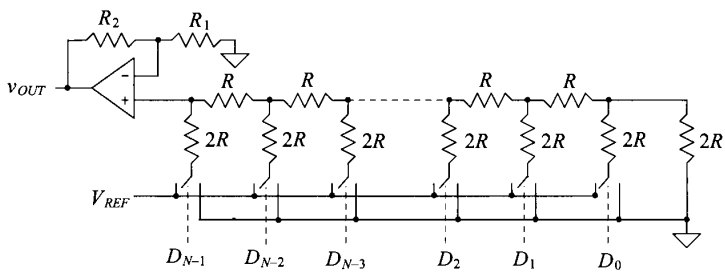


Figure 29.52 DAC used in Problem 29.6.

IPR2023-00698
Theta EX2019

29.7 Design a 3-bit, current-steering DAC using the generic current-steering DAC shown in Fig. 29.9. Assume that each current source, I, is 5 μA, and find the total output current for each input code.

29.8 A certain process is able to fabricate matched current sources to within 0.05%. Determine the maximum resolution that a current-steering (nonbinary-weighted) DAC can attain using this process.

29.9 Design an 8-bit current-steering DAC using binary-weighted current sources. Assume that the smallest current source will have a value of 1 μA. What is the range of values that the current source corresponding to the MSB can have while maintaining an INL of ½ LSB? Repeat for a DNL less than or equal to ½ LSB.

29.10 Prove that the 3-bit charge-scaling DAC used in Ex. 29.6 has the same output voltage increments as the R-$2R$ DAC in Ex. 29.3 for $V_{REF} = 5$ V and $C = 0.5$ pF.

29.11 Determine the output of the 6-bit, charge-scaling DAC used in Ex. 29.7 for each of the following inputs: $D = 000010$, 000100, 001000, and 010000.

29.12 Design a 4-bit, charge-scaling DAC using a split array. Assume that $V_{REF} = 5$ V and that $C = 0.5$ pF. Draw the equivalent circuit for each of the following input words and determine the value of the output voltage: $D = 0001$, 0010, 0100, 1000. Assuming the capacitor associated with the MSB had a mismatch of 4 percent, calculate the INL and DNL.

29.13 For the cyclic converter shown in Fig. 29.17, determine the gain error for a 3-bit conversion if the feedback amplifier had a gain of 0.45 V/V. Assume that $V_{REF} = 5$ V.

29.14 Repeat Problem 29.13 assuming that the output of the summer was always 0.2 V greater than the ideal and that the amplifier in the feedback path had a perfect gain of 0.5 V/V.

29.15 Repeat Problem 29.13 assuming that the output of the summer was always 0.2 V greater than ideal and that the amplifier in the feedback path had a gain of 0.45 V/V.

29.16 Design a 3-bit pipeline DAC using $V_{REF} = 5$ V. (a) Determine the maximum and minimum gain values for the first-stage amplifier for the DAC to have less than ±½ LSBs of DNL assuming that the rest of the circuit is ideal. (b) Repeat for the second-stage amplifier. (c) Repeat for the last-stage amplifier.

29.17 Using the same DAC designed in Problem 29.16, (a) determine the overall error (offset, DNL, and INL) for the DAC if the S/H amplifier in the first stage produces an offset at its output of 0.25 V. Assume that all of the remaining components are ideal. (b) Repeat for the second-stage S/H. (c) Repeat for the last-stage S/H.

29.18 Design a 3-bit Flash ADC with its quantization error centered about zero LSBs. Determine the worst-case DNL and INL if resistor matching is known to be 5%. Assume that $V_{REF} = 5$ V.

IPR2023-00698
Theta EX2019

29.19 Using the ADC designed in Problem 29.18, determine the maximum offset that can be tolerated if all of the comparators have the same magnitude of offset, but with different polarities, to attain a DNL of less than or equal to ±½ LSB.

29.20 A 4-bit Flash ADC converter has a resistor string with mismatch as shown in Table 29.1. Determine the DNL and INL of the converter. How many bits of resolution does this converter possess? V_{REF} = 5 V.

Resistor	Mismatch (%)
1	2
2	1.5
3	0
4	−1
5	−0.5
6	1
7	1.5
8	2
9	2.5
10	1
11	−0.5
12	−1.5
13	−2
14	0
15	1
16	1

Table 29.1 Mismatch in resistors used in Problem 29.20

29.21 Determine the open-loop gain required for the residue amplifier of a two-step ADC necessary to keep the converter to within ½ LSB of accuracy with resolutions of (a) 4 bits, (b) 8 bits, and (c) 10 bits.

29.22 Assume that a 4-bit, two-step Flash ADC uses two separate Flash converters for the MSB and LSB ADCs. Assuming that all other components are ideal, show that the first Flash converter needs to be more accurate than the second converter. Assume that V_{REF} = 5 V.

29.23 Repeat Ex. 29.12 for V_{IN} = 3, 5, 7.5, 14.75 V.

29.24 Repeat Ex. 29.13 for V_{IN} = 1, 4, 6, 7 V and V_{REF} = 8 V.

29.25 Assume that an 8-bit pipeline ADC was fabricated and that all the amplifiers had a gain of 2.1 V/V instead of 2 V/V. If V_{IN} = 3 V and V_{REF} = 5 V, what would be the resulting digital output if the remaining components were considered to be ideal? What are the DNL and INL for this converter?

IPR2023-00698
Theta EX2019

29.26 Show that the first-stage accuracy is the most critical for a 3-bit, 1-bit per stage pipeline ADC by generating a transfer curve and determining DNL and INL for the ADC for two cases: (1) The gain of the first-stage residue amplifier set equal to 2.2 V/V and (2) the second-stage residue amplifier set equal to 2.2 V/V. For each case, assume that the remaining components are ideal. Assume that the V_{REF} = 5 V.

29.27 An 8-bit single-slope ADC with a 5 V reference is used to convert a slow-moving analog signal. What is the maximum conversion time assuming that the clock frequency is 1 MHz? What is the maximum frequency of the analog signal? What is the maximum value of the analog signal which can be converted?

29.28 An 8-bit single slope ADC with a 5 V reference uses a clock frequency of 1 MHz. Assuming that all of the other components are ideal, what is the limitation on the value of RC? What is the tolerance of the clock frequency which will ensure less than 0.5 LSB of INL?

29.29 An 8-bit dual slope ADC with a 5 V reference is used to convert the same analog signal in Problem 29.27. What is the maximum conversion time assuming that the clock frequency is 1 MHz? What is the minimum conversion time that can be attained? If the analog signal is 2.5 V, what will be the total conversion time?

29.30 Discuss the advantages and disadvantages of using a dual-slope versus a single slope ADC architecture.

29.31 Repeat Ex. 29.15 for a 4-bit successive approximation ADC using V_{REF} = 5 V for v_{IN} = 1, 3, and full-scale.

29.32 Assume that v_{IN} = 2.49 V for the ADC used in Problem 29.31 and that the comparator, because of its offset, makes the wrong decision for the MSB conversion. What will be the final digital output? Repeat for v_{IN} = 0.3025, assuming that the comparator makes the wrong decision on the LSB.

29.33 Design a 3-bit, charge-redistribution ADC similar to that shown in Fig. 29.39 and determine the voltage on the top plate of the capacitor array throughout the conversion process for v_{IN} = 2, 3, and 4 V, assuming that V_{REF} = 5 V. Assume that all components are ideal. Draw the equivalent circuit for each bit decision.

29.34 Determine the maximum INL and maximum DNL of the ADC designed in Problem 29.33 assuming that the capacitor array matching is 1%. Assume that the remaining components are ideal and that the unit capacitance, C, is 1 pF.

29.35 Show that the charge redistribution ADC used in Problems 29.32 and 29.33 is immune to comparator offset by assuming an initial offset voltage of 0.3 and determining the conversion for v_{IN} = 2 V.

29.36 Discuss the differences between Nyquist rate ADCs and oversampling ADCs.

29.37 Write a simple computer program or use a math program to perform the analysis shown in Ex. 29.16. Run the program for k = 200 clock cycles and show that the average value of $v_q(kT)$ converges to the correct answer. How many clock cycles will it take to obtain an average value if $v_q(kT)$ stays within 8-bit accuracy of the ideal value of 0.4 V? 12-bit accuracy? 16-bit accuracy?

IPR2023-00698
Theta EX2019

29.38 Prove that the output of the second-order ΣΔ modulator shown in Fig. 29.49 is,

$$y(kT) = x(kT - T) + Q_e(kT) - 2Q_e(kT - T) + Q_e(kT - 2T)$$

29.39 Assume that a first order ΣΔ ADC used on a satellite in a low earth orbit experiences radiation in which an energetic particle causes a noise spike resulting in the comparator making the wrong decision on the 10th clock period. Using the program written in Problem 29.37, determine the number of clock cycles required before the average value of $v_q(kT)$ is within 12-bit accuracy of the ideal value of 0.4 V. How many extra clock cycles were required for this case versus the ideal conversion used in Prob. 37?

IPR2023-00698
Theta EX2019

Chapter

30

Implementing Data Converters

CMOS technology continues to scale towards smaller dimensions. This feature size reduction is driven mainly by the desire to implement digital systems of increased complexity in a smaller area. This natural trend in feature size reduction, with accompanying reduction in supply voltage and poorer matching, can present challenges for the CMOS circuit designer. The accompanying lower supply voltage, for example, results in an inherent reduction in dynamic range, decrease in *SNR*, and increasing challenges when implementing analog circuitry with little, ideally zero, voltage overhead.

This chapter presents and discusses implementation methods and trade-offs for designing data converters in nanometer CMOS. For DAC design, we focus on converters implemented with both resistors using *R*-2*R* networks and current sources. The benefit of, and reason we are focusing on, using *R*-2*R* networks and current sources over other methods for DAC implementation, such as charge redistribution DACs, is the absence of good poly-poly capacitors in nanometer digital CMOS processes. *R*-2*R*-based DACs can be laid out in a small area while achieving resolutions in excess of 10-bits without calibrations or trimming. Charge-scaling DACs require linear capacitors. The layout area needed for these capacitors can often be very large and practically limit both the resolution and accuracy of the DAC.

The first section of this chapter discusses *R*-2*R* and current-steering DACs. The second section of the chapter discusses the use of op-amps in data converters, while the third section presents an overview of general ADC implementations in nanometer CMOS processes. In this last section we concentrate our discussion on the implementation of pipeline analog-to-digital data converters.

It's important to note that the goal in this chapter is not to provide an exhaustive overview of data converter design but rather to provide discussions and practical insight helpful when implementing any type of data converter in CMOS technology. We assume that the reader is familiar with data converter fundamentals, Ch. 28, and data converter architectures, Ch. 29. For example, the reader knows the difference between differential nonlinearity (DNL) and integral nonlinearity (INL) or the difference between a two-step flash ADC and a pipeline ADC.

IPR2023-00698
Theta EX2019

30.1 *R*-2*R* Topologies for DACs

We begin this section by discussing *R*-2*R* DAC topologies. The problems encountered in the traditional *R*-2*R* topologies with low-voltage overhead are illustrated. Also, concerns related to the performance of the op-amps used in data converters (both ADCs and DACs) are discussed. Finally, matching and accuracy concerns are presented along with techniques to remove these imperfections using calibration.

30.1.1 The Current-Mode *R*-2*R* DAC

The *R*-2*R* DAC can be classified into two categories: voltage-mode and current-mode. A current-mode *R*-2*R* DAC is shown in Fig. 30.1. The branch currents flowing through the 2*R* resistors are of a binary-weighted relationship caused by the voltage division of the *R*-2*R* ladder network and are diverted either to the inverting input of the op-amp (actually the feedback resistor) or the noninverting input of the op-amp (actually V_{REF-}). The voltage on the *R*-2*R* resistor string at the Xth tap (where X ranges from 0 to $N-1$), in Fig. 30.1, can be written as

$$V_{TAPX} = \frac{2^X}{2^N} \cdot (V_{REF+} - V_{REF-}) + V_{REF-} \tag{30.1}$$

where V_{REF+} and V_{REF-} are the N-bit DAC's reference voltages. The current that flows through the 2*R* resistor at the Xth tap is then

$$I_{TAPX} = \frac{V_{TAPX} - V_{REF-}}{2R} = \frac{1}{2R} \cdot \frac{2^X}{2^N}(V_{REF+} - V_{REF-}) \tag{30.2}$$

This current is summed at the inverting input of the op-amp and flows through the feedback resistor to the DAC output, V_{out}. The output voltage of the DAC can be written as

$$V_{out} = V_{REF-} - R \cdot \sum_{X=0}^{N-1} (b_X \cdot I_{TAPX}) \text{ for } V_{REF+} > V_{REF-} \tag{30.3}$$

where b_X is either a 1 or 0, or

$$V_{out} = V_{REF-} + R \cdot \sum_{X=0}^{N-1} (b_X \cdot I_{TAPX}) \text{ for } V_{REF+} < V_{REF-} \tag{30.4}$$

Using these equations, we can see the main problem with the basic current mode topology of Fig. 30.1 in a nanometer CMOS process using low-power supply voltages, namely, limited output swing. If V_{REF-} is set to 0 V, with $V_{REF+} > 0$, then the output of the DAC must be negative, which, of course, can't happen when the only power supply voltage is (positive) *VDD*. If V_{REF-} is set to *VDD*, then we can see from Eq. (30.4) that this would require $V_{out} > VDD$. After reviewing Eqs. (30.1)-(30.4), we see that the range of output voltages associated with the current mode *R*-2*R* DAC is limited to *VDD*/2, e.g., 0 to *VDD*/2, *VDD*/2 to *VDD* or 0.25*VDD* to 0.75*VDD*, etc. Giving up half of the power-supply range in a DAC and correspondingly reducing the dynamic range, is usually not desirable.

By removing the requirement that the noninverting input of the op-amp be tied to V_{REF-} and that the feedback resistor be *R* (the same value used in the *R*-2*R* string), we can increase the output range of the DAC. The output of the op-amp is level-shifted by the voltage on the noninverting input of the op-amp and by the increased value of closed-loop gain of the op-amp. Similarly, we could add a gain stage to the output of the DAC (two

IPR2023-00698
Theta EX2019

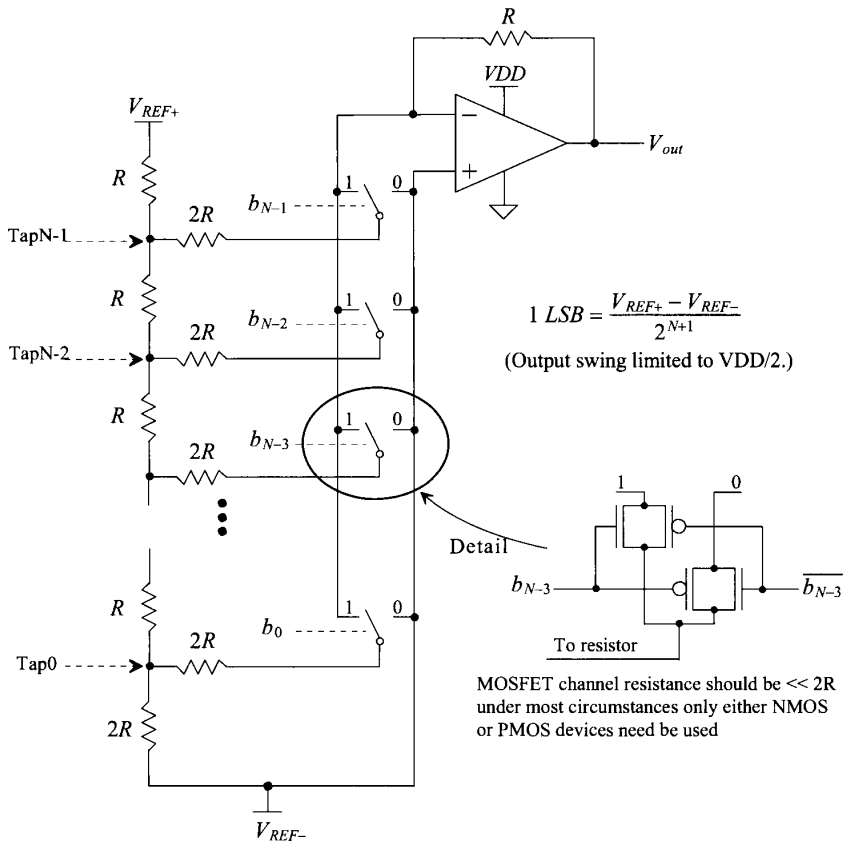


Figure 30.1 Traditional current-mode R-2R DAC.

op-amps would then be used) to achieve wider DAC output swing. We don't cover these options any further here because they either put more demand on the op-amp design, such as increased op-amp open-loop gain/speed, increased power consumption, or won't, in a practical implementation, result in a rail-to-rail output swing.

30.1.2 The Voltage-Mode *R-2R* DAC

Figure 30.2 shows a schematic of a voltage-mode DAC. The voltage on the non-inverting input of the op-amp can be written as

$$V_+ = \frac{b_{N-1} \cdot V_{REF+} + \overline{b_{N-1}} \cdot V_{REF-}}{2^1} + \frac{b_{N-2} \cdot V_{REF+} + \overline{b_{N-2}} \cdot V_{REF-}}{2^2} + \ldots + \frac{V_{REF-}}{2^N} \qquad (30.5)$$

IPR2023-00698
Theta EX2019

or, in general terms,

$$V_{+} = \sum_{k=1}^{N} \frac{b_{N-k} \cdot V_{REF+} + \overline{b_{N-k}} \cdot V_{REF-}}{2^k} + \frac{V_{REF-}}{2^N} \tag{30.6}$$

The output of the N-bit voltage-mode DAC can be written as

$$V_{out} = \left[1 + \frac{R_F}{R_I}\right] \cdot \left[\sum_{k=1}^{N} \frac{b_{N-k} \cdot V_{REF+} + \overline{b_{N-k}} \cdot V_{REF-}}{2^k} + \frac{V_{REF-}}{2^N}\right] \tag{30.7}$$

If the input code is all zeroes, with $V_{REF-} = 0$, $V_{REF+} = VDD$, and the op-amp in the follower configuration, then $V_{out} = V_{REF-}$. If the input code is all ones, then the output of the DAC is $V_{REF+} - 1$ LSB.

By using the voltage-mode DAC, we would seem to have solved the problem of the limited output swing associated with the current-mode DAC of Fig. 30.1. However, consider how the finite common-mode rejection ratio (*CMRR*) of the op-amp in Fig. 30.2 can affect the linearity of the overall DAC design. We know the effects of finite *CMRR* can be modeled as a variable offset voltage, ΔV_{OS} (see Ch. 24), in series with the noninverting input of the op-amp that is a function of the change in the op-amp common-mode voltage, ΔV_C, or

$$\Delta V_{OS} = \frac{\Delta V_C}{CMRR} \tag{30.8}$$

We should see the problem at this point: that is, ΔV_{OS} is in series with the *R*-2*R* resistor string and will ultimately limit the linearity of the DAC. To further illustrate the problem, let's assume the *CMRR* of the op-amp in Fig. 30.2 is 20 dB at 1 MHz. Since the common-mode voltage on the input of the op-amp, again assuming $V_{REF+} = VDD$, $V_{REF-} = 0$, and the op-amp in the unity follower configuration can range from zero to approximately *VDD*, the change in the offset voltage used to model finite *CMRR* when the DAC's inputs are changing at 1 MHz is 10% of *VDD*. At first glance we might simply consider the resulting offset as a nonlinear gain error affecting only the large-signal linearity (INL). However, it is unlikely in any practical op-amp design that the *CMRR* will vary linearly with changes in the input common-mode voltage and so the small-signal linearity (DNL) will be affected as well. Since, for this example, 1 LSB = $VDD/2^N$, the resolution of the DAC, because of the finite *CMRR* and assuming 1 LSB > ΔV_{OS}, is limited to 4 bits! Performing DC or audio-frequency tests on the voltage-mode DAC made with an op-amp with a *CMRR* of, for example, 120 dB at DC results in no practical resolution limit (indicating that if DAC speed isn't a concern, the voltage-mode configuration may still be used for high resolutions). Note how *CMRR* isn't a concern with the current-mode *R*-2*R* DAC (assuming no secondary effects, such as common mode substrate noise, are present on the input of the op-amp). *For precision, high-speed data converter design we must use an inverting op-amp topology where the inputs of the op-amp remain at a fixed voltage.*

30.1.3 A Wide-Swing Current-Mode R-2R DAC

We've shown that it is desirable to have a wide output swing, as is provided by the voltage-mode *R*-2*R* DAC, while at the same time having a fixed input common mode voltage, as is provided by the current-mode *R*-2*R* DAC. Figure 30.3 shows a wide-swing

IPR2023-00698
Theta EX2019

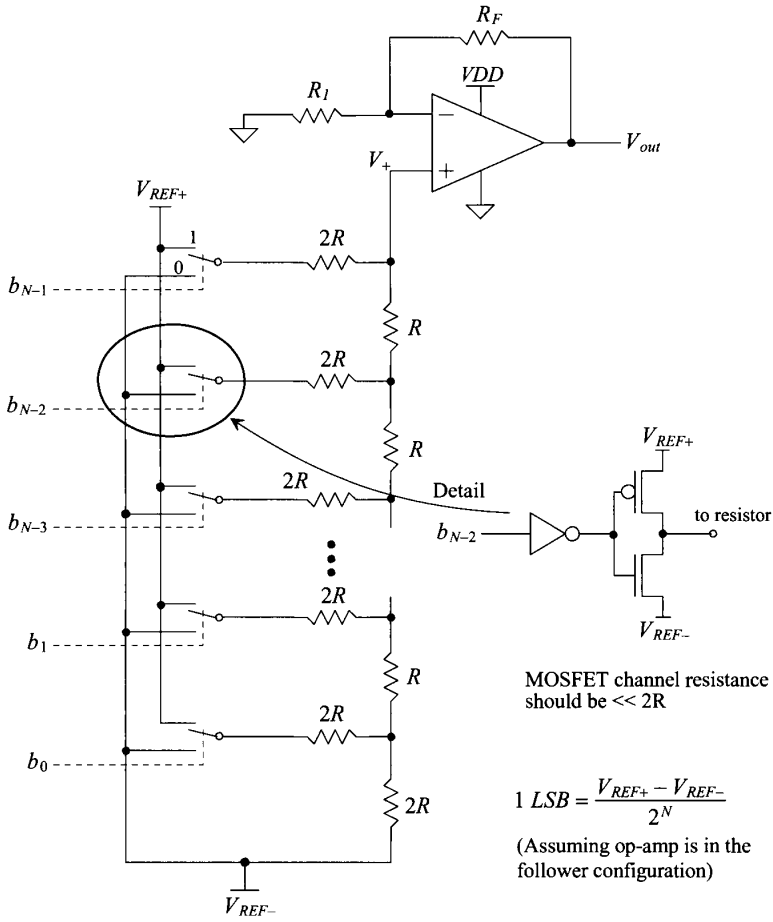


Figure 30.2 Traditional voltage-mode R-2R DAC.

current-mode *R*-2*R* DAC configuration that has a rail-to-rail output swing while keeping the input common-mode voltage of the op-amp fixed at the common mode voltage, V_{CM}, or $(V_{REF+} + V_{REF-})/2$.

Like traditional current-mode *R*-2*R* DACs, the DAC scheme shown in Fig. 30.3 operates on currents. Using superposition and assuming V_{REF-} is the reference for calculations, the current flowing in the feedback resistor, R_F, is given by

$$I_F = -\frac{V_{REF+} - V_{REF-}}{2R} + \frac{V_{REF+} - V_{REF-}}{2R} \cdot \left[1 \cdot \overline{b_{N-1}} + \frac{1}{2} \cdot \overline{b_{N-2}} + \ldots + \frac{1}{2^{N-1}} \cdot \overline{b_0} \right] \quad (30.9)$$

IPR2023-00698
Theta EX2019

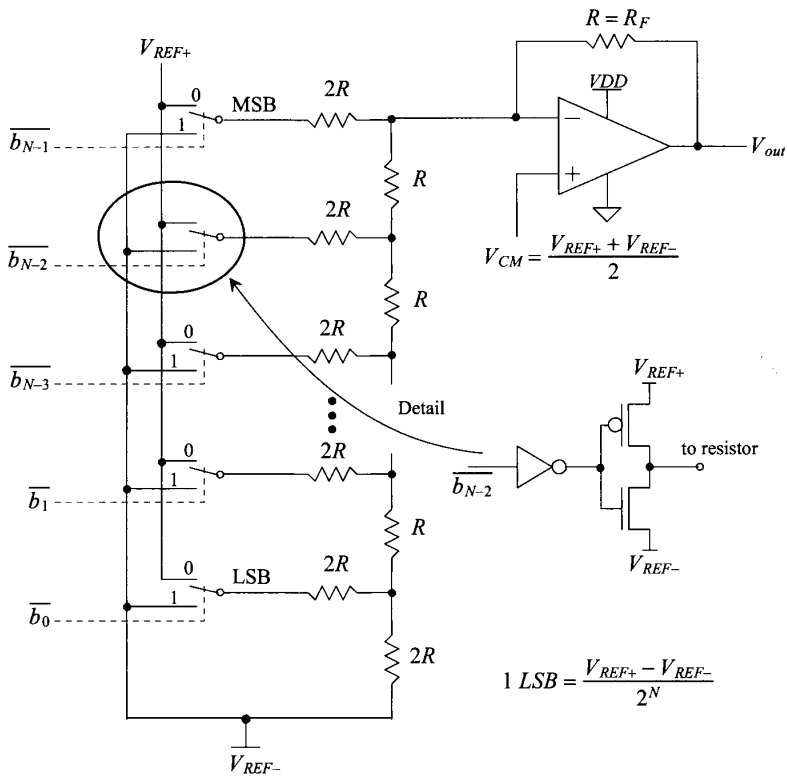


Figure 30.3 Wide-swing current-mode R-2R DAC.

noting the inversion used in the control logic seen in Fig. 30.3. The output voltage of the DAC is then given, assuming $R = R_F$, by

$$V_{out} = V_{REF-} + \frac{V_{REF+} - V_{REF-}}{2} + I_F \cdot R \tag{30.10}$$

or

$$V_{out} = V_{REF-} + (V_{REF+} - V_{REF-}) \cdot \left[1 - \left(\frac{1}{2} \cdot \overline{b_{N-1}} + \frac{1}{4} \cdot \overline{b_{N-2}} + \ldots + \frac{1}{2^N} \cdot \overline{b_0} \right) \right] \tag{30.11}$$

From this equation, we see that as the digital input code is sequenced through 000... to 111... the output of the DAC changes in steps of $(V_{REF+} - V_{REF-})/2^N$ (= 1 LSB) from V_{REF+} (when the input code is 111...) to V_{REF-} + 1 LSB (when the input code is 000...). Setting V_{REF-} to ground and V_{REF+} to *VDD* allows the DAC output to swing from rail to rail.

IPR2023-00698
Theta EX2019

In practice, since any rail-to-rail output op-amp has high nonlinearity close to its power-supply rails, a slightly "shrunk" output range from power rails is often desired. For example, we can set $V_{REF+} = 0.9 \cdot VDD$ and $V_{REF-} = 0.1 \cdot VDD$. The output will change between 10 and 90% of VDD centered at $VDD/2$. Another way to shrink the output range is to make the feedback resistance R_F smaller than R (as seen in Eq. [30.10]) by either trimming or programming the value of the feedback resistor R_F.

The matching between the resistors of the R-$2R$ ladder is one of the most important and limiting factors that determine the linearity (e.g., DNL and INL) of the entire DAC. It is helpful, when designing any type of resistor string DAC, if we can estimate the resistor matching requirements based on a desired resolution.

DNL Analysis

It was shown in Ch. 29 that for a binary-weighted DAC the worst case DNL condition tends to occur at midscale when the code transitions from 01...11 to 10...00. Let's assume in a worst-case scenario the $2R$ resistance of the MSB input in Fig. 30.3 has a maximum positive mismatch of ΔR, and all other resistors have a maximum negative mismatch of $-\Delta R$. *In this case, the current provided by the MSB has to match the sum of currents provided by all other lower input bits plus one LSB.* Again using the superposition principle we can verify that the step error of the current flowing through the feedback resistor R_F, caused by the resistor mismatch at the midscale transition, is approximately equal to

$$\Delta I = \frac{V_{REF+} - V_{REF-}}{2(R - \Delta R)} \cdot \left[1 - \frac{1}{2^{N-1}}\right] - \frac{V_{REF+} - V_{REF-}}{2(R + \Delta R)} \qquad (30.12)$$

Assuming $R_F = R$, the final output step error (DNL) is approximately

$$DNL = \Delta I \cdot R \approx (V_{REF+} - V_{REF-}) \cdot \left[\frac{\Delta R}{R} - \frac{1}{2^N}\right] \qquad (30.13)$$

For the DNL to be within 1 LSB (1 LSB equals to $[V_{REF+} - V_{REF-}]/2^N$) the matching required of the resistors is

$$\text{Resistor mismatch} = \left|\frac{\Delta R}{R}\right| \leq \frac{1}{2^{N-1}} \qquad (30.14)$$

For a 10-bit data converter to have a DNL of less than 1 LSB requires the MSB resistor to match within 0.2% (= $\Delta R/R$) of the lower resistors (which were assumed to have the same value, i.e., the maximum mismatch from the MSB resistor) in the R-$2R$ string. Equation (30.14) results in a pessimistic estimate for the matching required of the resistors because the variation in resistance along the string does not vary abruptly at the MSB resistor but rather, in most cases, varies linearly from LSB to MSB. As we'll see in the experimental results discussed in the next section, the matching requirements result in a practical limit of 10-bits for an R-$2R$-based converter with no special layout or circuit techniques (for example, averaging variations using multiple resistor strings or using segmentation).

INL Analysis

Since any change of the $2R$ resistance in the MSB has the largest influence on the ladder output current among that of all the branch resistors ($2R$), the worst case INL tends to occur when the input code is 01...11. (The gain error is nulled from the INL calculation

IPR2023-00698
Theta EX2019

here, and therefore there is no INL error, but a gain error instead, if all the resistors have a maximum mismatch.) Assuming that the $2R$ resistance of the MSB has a maximum positive mismatch of $\Delta R/R$, the error in the current flowing through R_F from its ideal value caused by the resistance mismatch is

$$\Delta I = \frac{V_{REF+} - V_{REF-}}{2(R+\Delta R)} - \frac{V_{REF+} - V_{REF-}}{2R} \approx -\frac{V_{REF+} - V_{REF-}}{2} \cdot \frac{\Delta R}{(R+\Delta R)\cdot R} \quad (30.15)$$

The worst-case INL tends to occur, assuming $R_F = R$, when

$$INL = -\Delta I \cdot R \approx \frac{V_{REF+} - V_{REF-}}{2} \cdot \frac{\Delta R}{R+\Delta R} \quad (30.16)$$

For the INL to be within 1 LSB, this also approximately yields

$$\text{Resistor mismatch} = \left|\frac{\Delta R}{R}\right| \le \frac{1}{2^{N-1}} \quad (30.17)$$

Again, as was mentioned in the DNL analysis, this is a pessimistic estimate if the sheet resistance varies linearly with distance. Equations (30.14) and (30.17) indicate that a resistance matching to within $1/2^N$ is required for less than ½ LSB of DNL and INL for the DAC scheme in Fig. 30.3. Layout of R-$2R$ resistors was discussed in Ch. 5.

Switches

The switches (MOSFETs) used in the R-$2R$ DAC should have an effective switching resistance much less than the resistors used in the R-$2R$ ladder. The inherent switching time of the switches is extremely fast (speeds comparable to logic gate delays). Since the switches are in series with the branch resistances of the R-$2R$ ladder, the R-$2R$ relationship is broken if the switch resistance is not negligible, and this affects both the INL and the DNL. Also note that we can try to compensate for the switch-effective resistance by making the length of the $2R$ resistor slightly shorter than the length of the R resistor. However, if not careful, this may lead to problems over the process corners and temperature.

Experimental Results

The wide-swing, current-mode R-$2R$ DAC, based on the scheme in Fig. 30.3, was fabricated in a 210 nm, VDD = 1.8 V, CMOS process with resolutions of 8, 10, and 12 bits. The cell dimensions of the 12-bit DAC are 150 μm by 300 μm. The goal of the experimental results was to verify that the topology of Fig. 30.3 would indeed perform as predicted by Eqs. (30.14) and (30.17) and to generate a low-power, small-area DAC cell for general-purpose, mixed-signal circuit designs. Unsilicided n+ poly was used for the R-$2R$ resistances as discussed earlier (see Table 4.1). The mismatch indicated in Table 4.1 for an unsilicided n+ poly resistor is 0.005 (= $\Delta R/R$). Using Eqs. (30.14) and (30.17), we would estimate that our resolution is limited to 8.6 bits if we want both INL and DNL less than 1 LSB. The results in Table 30.1, however, show that the resolution is better than estimated. This may be because of our pessimistic assumption of how the resistor values change with position as discussed in the derivation of these equations. The nominal resistor value used in these experimental DACs is 10k. To enhance the resistance matching, dummy resistors are implemented at both ends of the R-$2R$ ladder (see Fig. 5.28). The output range of the DAC is programmable by choosing the value of the feedback resistance or by the setting of the reference voltages V_{REF+} and V_{REF-}.

IPR2023-00698
Theta EX2019

Table 30.1 Summary of experimental results.

	8-bit	10-bit	12-bit
DNL (LSB)	0.150	0.450	2.000
INL (LSB)	0.200	1.000	3.000
Settling time	200 ns		
Power	3.88 mW (driving a 1k load)		
Area (mm^2)	0.045		
$f_{clk,max}$	4 MHz		
Output swing	$0 < V_{out} < VDD$ (= 1.8 V)		

The measured INL and DNL profiles of the three DACs with resolutions of 8, 10, and 12 bits are shown in Fig. 30.4. The outputs of the DACs are configured to swing to both rails ($V_{REF+} = VDD$ and $V_{REF-} = 0$). The first several points, adjacent to the two rails, are not shown in Fig. 30.4 due to the high nonlinearity of the op-amp in those regions. Major performance results are maximum DNLs of 0.15 LSB, 0.45 LSB, and 2 LSB for

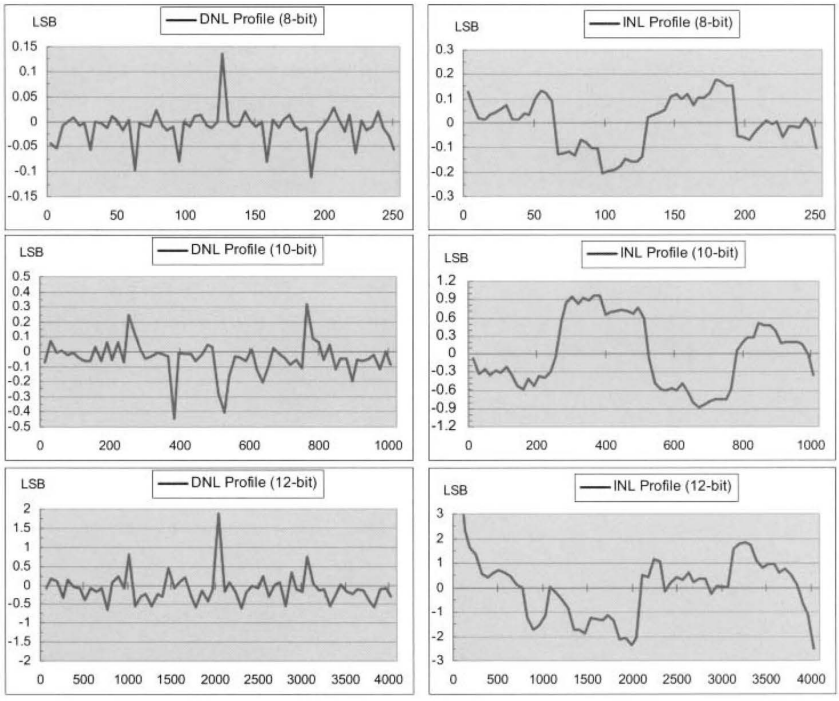


Figure 30.4 Experimental results for the wide-swing DAC of Fig. 30.3.

IPR2023-00698
Theta EX2019

8-bit, 10-bit, and 12-bit resolutions, respectively, with no special circuit techniques (laid out as shown in Fig. 5.28) or trimming (adjustments). The corresponding maximum DAC INLs are 0.2 LSB, 1 LSB, and 3 LSB, respectively. Notice that the LSB of the 8-, 10-, and 12-bit DACs are 7.03 mV, 1.75 mV, and 439 μV, respectively. The DNL/INL can be written in terms of a voltage as 1.05 mV/1.4 mV for the 8-bit DAC, 0.788 mV/1.75 mV for the 10-bit DAC, and 0.878 mV/1.31 mV for the 12-bit DAC. The measurements were taken while the DAC was driving a 1k load. The power dissipated by the DAC, with 1.8 V output, while driving a 1k resistor is 3.88 mW. The unloaded power dissipation of the DAC is approximately 500 μW. The DACs were designed using op-amps with simulated unity gain frequencies of 10 MHz. The measured DAC settling time was approximately 200 ns.

Improving DNL (Segmentation)

After reviewing the DNL plots in Fig. 30.4, we see that the worst-case DNL occurs when the input code transitions from 01111... to 10000... (midscale) where the current in the top 2R should be 1 LSB (equivalent in current) greater than the sum of all of the currents contributed by the lower resistors. As an example, consider the 12-bit *R*-2*R* ladder in Fig. 30.5 where we have used 1 μA to indicate an LSB of current contribution to the feedback

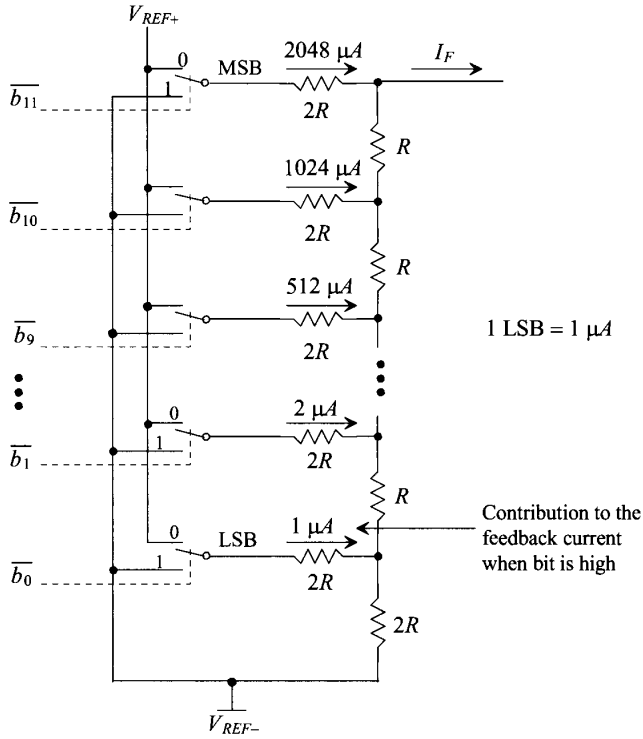


Figure 30.5 Showing how currents sum into the feedback current.

IPR2023-00698
Theta EX2019

path. When the input digital code is 0111 1111 1111, the feedback current is 2047 μA. When the code changes to 1000 0000 0000, the feedback current becomes (ideally) 2048 μA. If a 1/2 LSB error (0.5 μA) is the maximum error allowable, then the accuracy required of the currents when transitioning is 0.5/2048 or 0.0244%. If we use fewer bits, say eight, then the accuracy required when transitioning from 255 μA to 256 μA is 0.5/256 or 0.2%.

Let's consider segmenting the upper four bits in Fig. 30.5 so that the four bits control 16 segments each contributing 256-μA to the feedback current. This segmentation makes attaining good DNL with less accurate components possible. A segmented wide-swing DAC is shown in Fig. 30.6. In this figure we've taken the upper four bits and segmented their current contributions to the feedback resistor. If we use the numbers from Fig. 30.5, then when the code 0000 1111 1111 (255 μA) transitions to 0001 0000 0000 (256 μA) the 1 output of the decoder goes high and the bottom resistor connected to the

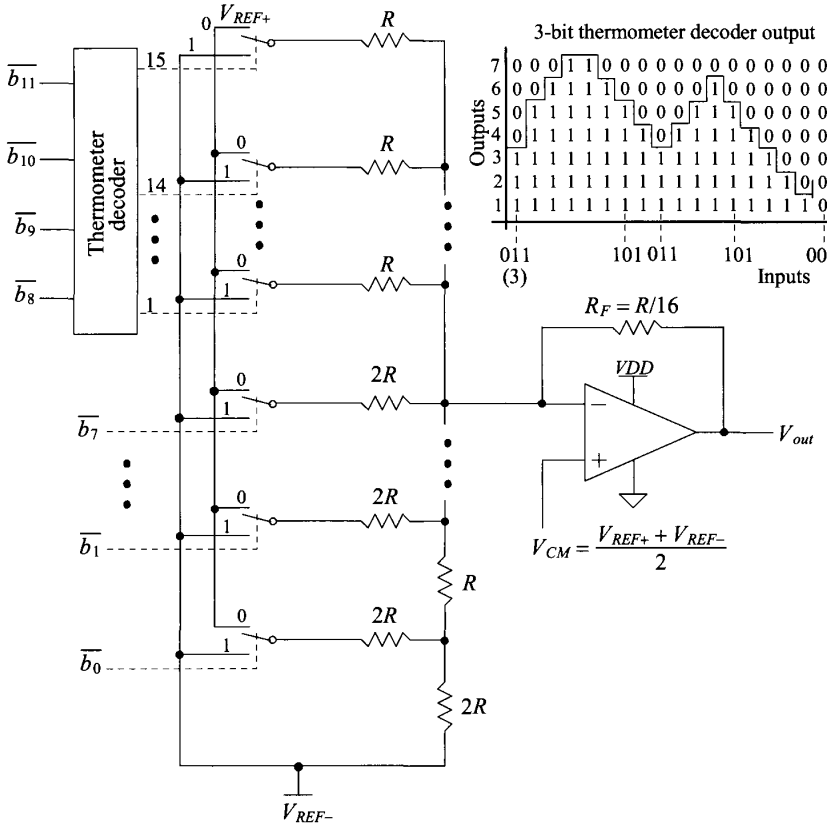


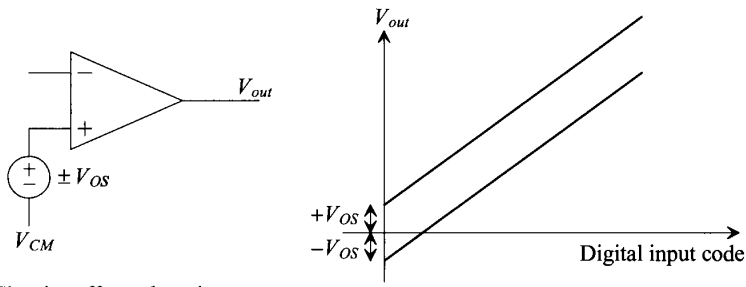
Figure 30.6 Segmentation in a wide-swing R-2R DAC.

IPR2023-00698
Theta EX2019

output of the decoder contributes 256 μA to the feedback path. When the code changes from 0001 1111 1111 (511 μA) to 0010 0000 0000 (512 μA), both lower outputs (1 and 2) of the thermometer decoder are high. Since the 1 decoder output continues to contribute to the output current, the step height is set by the difference between the 2 decoder outputs and the contributions from the lower eight bits. This makes the accuracy requirements for 1/2 LSB DNL in a 12-bit converter set by 8-bit matching. Note that while segmentation reduces DNL error, it does nothing for INL. Segmentation can also be used to reduce the glitch area associated with the changing DAC output.

Trimming DAC Offset

Figure 30.7 shows how the op-amp's offset voltage shifts the DAC's output. It may be desirable in some situations to trim or remove this offset. The offset may be the result of an inherent systematic offset in the op-amp or the result of random variations in the characteristics of the MOSFETs used in the op-amp. An offset may also result because of the voltage dependence of the resistors used to generate the common-mode voltage, V_{CM}.



(a) Showing offset voltage in an op-amp.

(b) DAC transfer curves showing offset.

Figure 30.7 Showing how an op-amp offset affects the DACs transfer curves.

Figure 30.8 shows one possible method to generate a common-mode voltage that is adjustable with a digital code. Here we are assuming that V_{CM} is ideally 0.5 V. We should recognize the *R*-2*R* ladder from Fig. 30.2. The output voltage of this ladder, as seen in Eq. (30.6), is an analog voltage related to the digital input word (assuming the voltage divider made up of R_{big} and the two *R* resistors connected to the output in Fig. 30.8 doesn't load the circuit). Figure 30.9 shows the output of this circuit for all possible digital input words when *R* is 10k and R_{big} is 100k. This figure also shows that the adjustability of the output is approximately 1 mV. To decrease this value, we can either increase R_{big} (resulting in a decrease in the output swing) or increase the number of bits in the *R*-2*R* DAC. The value, *R*, of the resistors on the output can be decreased, but this can result in an increase in power dissipation.

Note that the accuracy required of the 5-bit DAC can be very loose. N-well resistors can be used to implement the offset trimming circuit to reduce area and power. The main concerns are considering the possibility of substrate noise injection and making sure that the same resistive material is used for the entire circuit. We wouldn't want the temperature behavior of an n-well resistor used in a circuit with a poly resistor because

IPR2023-00698
Theta EX2019

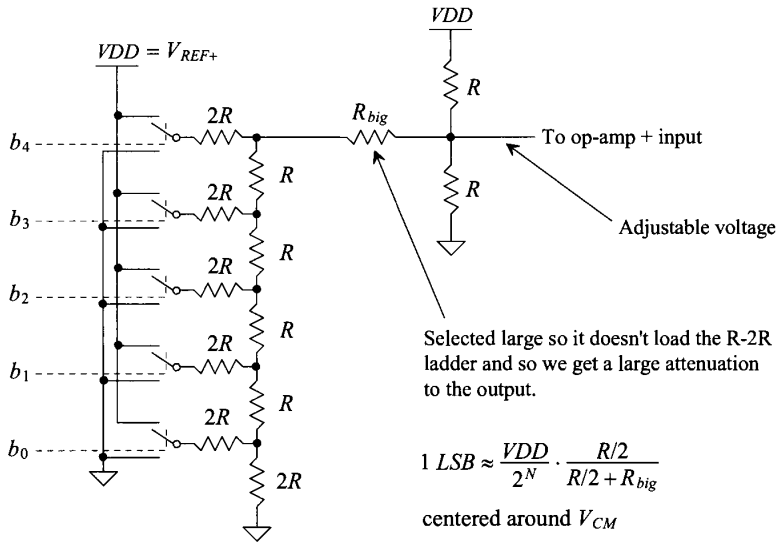


Figure 30.8 Trimming circuit for DAC offset.

the temperature dependencies are different (the offset trimming would only be effective at the temperature it was performed). Finally, note that in a practical circuit it is a good idea to add capacitors from the output of the circuit to both *VDD* and ground to ensure that the + op-amp input is connected to a good AC ground.

Trimming or calibrating out the offset can be performed at a time prior to packaging the chip, or it can be performed with some autocalibration sequence after the chip has been fabricated where the output of the DAC is compared to a known voltage reference. The concern, as with any calibration, is to adjust only one known error at a time (known as orthogonal tuning in filter design). For example, the DAC may not have any offset but may have an INL error for a given input code, Fig. 30.10a. If we were only

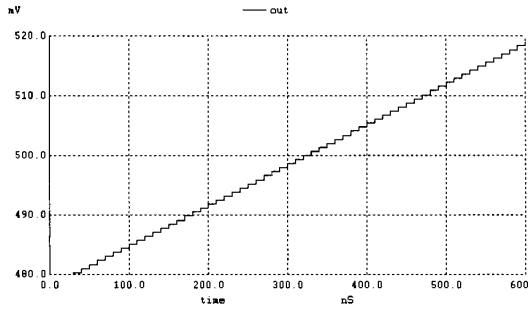
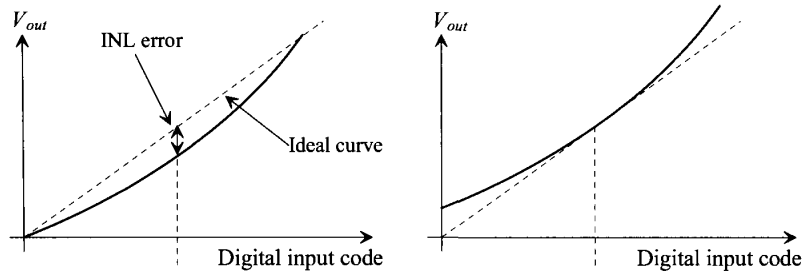


Figure 30.9 Output of the circuit in Fig. 30.8 for all possible digital codes.

IPR2023-00698
Theta EX2019



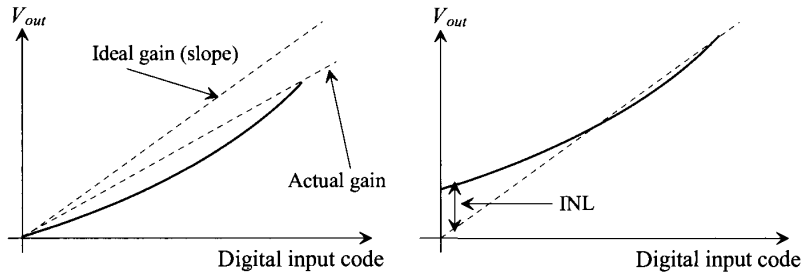
(a) DAC transfer curves before calibration. (b) DAC transfer curves after offset calibration

Figure 30.10 Showing how INL can be seen as an offset error.

to look at this one input code, say 10000... (V_{CM} in binary offset), we wouldn't know if the error is an INL error or an offset error. After the offset is calibrated out, Fig. 30.10b, we would then perform an INL calibration to pull the end-points of the transfer curve back to the ideal straight line transfer curve.

Trimming DAC Gain

We assumed in Fig. 30.10 that the gain of the DAC was one, in other words, there wasn't any gain error in the DAC's transfer function. If there is a gain error, the offset calibration can lead to poorer INL. Consider Fig. 30.11a showing gain and INL errors without any offset. Performing an offset calibration, Fig. 30.11b, can result in significant INL error. We can avoid this situation by calibrating out the gain error by trimming the op-amp's feedback resistor prior to offset calibration. A reference voltage close to the ends of the transfer curve is used while adjusting the gain of the op-amp used in the DAC. If V_{REF+} is less than *VDD* (to avoid op-amp saturation as its output approaches the supply rails), then it can be compared directly to the output of the DAC (keeping in mind the maximum output of the DAC may be $V_{REF+} - 1\ LSB$). Having gone through all of this discussion, it



(a) DAC transfer curves with gain error.

(b) DAC transfer curves after offset calibration with gain error.

Figure 30.11 Showing gain error and how it can cause problems in an offset calibration.

IPR2023-00698
Theta EX2019

still would be nicer if we could simply perform two calibrations, offset calibration and INL calibration, effectively using the INL calibration to remove the gain error. The drawback of this two calibration method is the requirement that an INL calibration circuit be capable of removing very large INL errors.

Improving INL by Calibration

We can calibrate out errors in our wide-swing DAC in two basic ways as seen in Fig. 30.12. The method shown in part (a) adds or subtracts a current from the feedback path to adjust the DAC output to the correct value. In part (b) the noninverting input of the op-amp is varied to force the DAC output to the correct value. The offset calibration described earlier uses the method shown in part (b). Note that the resistance looking from the inverting op-amp terminal back through the ladder to AC ground is simply R, so using the method in part (b) results in a noninverting op-amp configuration with a gain of two. (A variation of 1 mV on the + op-amp terminal causes an output variation of 2 mV.) Because we already have a circuit, Fig. 30.8, to make adjustments to the DAC output and the topology of part (b) doesn't provide any DC load to the calibrating voltage source and provides the least interaction with the main R-$2R$ ladder, we will use this topology to illustrate how we can calibrate out INL errors.

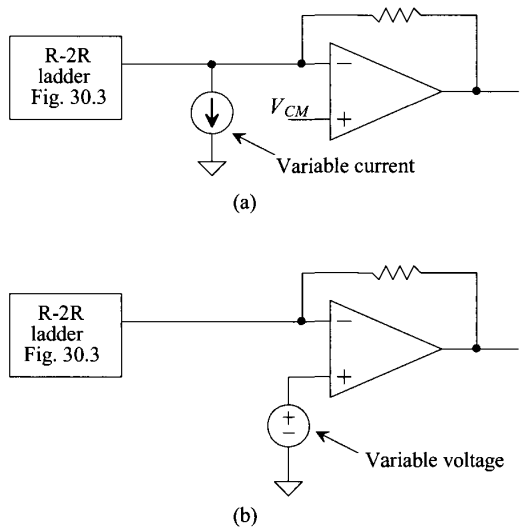


Figure 30.12 Trimming the output of the DAC using (a) current and (b) voltage.

Consider the calibration circuit shown in Fig. 30.13. In this figure the five most significant bits of a 12-bit DAC, that is, b_{11}, b_{10}, b_9, b_8, and b_7 are applied to the 12-bit DAC and to the address input of a 32-to-1 MUX with 5-bit input and output words. The MUX drives the R-$2R$ circuit of Fig. 30.8. The 5-bit register feeding each MUX input is used to store the calibration values. Again the calibration can be performed after the DAC

IPR2023-00698
Theta EX2019

is manufactured or during its use by employing a self-calibration sequence. In this scheme the DC offset calibration shown in Fig. 30.10 is simply one case of the 32 calibrations performed. (The top five bits of the input word are 10000 for the DC calibration.) This technique can be used to more precisely set the linearity of the DAC, perhaps up to 16-bits. Note that segmentation must still be employed to keep the DNL small. Also note that adding a large capacitance, C, to the noninverting input of the op-amp can result in a long time constant $\left(\approx \frac{R}{2} \cdot C\right)$, slowing the settling of the circuit and affecting the high-frequency *SNDR*.

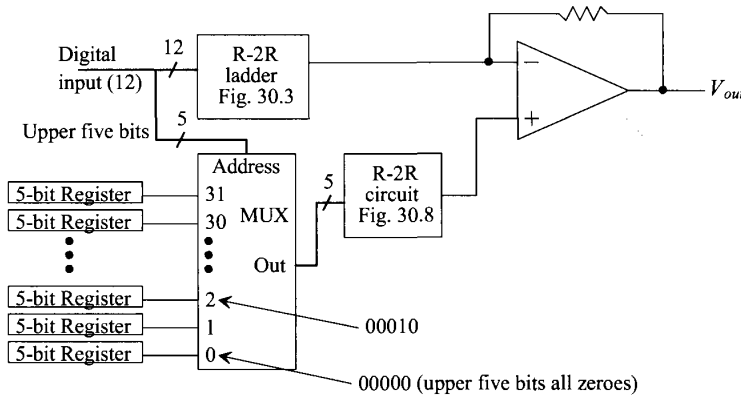


Figure 30.13 Calibration scheme for 12-bit DAC.

30.1.4 Topologies Without an Op-Amp

We discuss the requirements of op-amps used in data converters in detail in the next section. In this section we present an overview of topologies using voltage and current mode DACs based on both *R-2R* and *W-2W* topologies. The benefit of using a DAC with an op-amp is that an arbitrary load impedance can be connected to the DAC's output. Without the op-amp, the load impedance must either be known, capacitive, or very large, since it will load the DAC's output and affect INL, DNL, and, ultimately, the *SNR*. The benefits of not using an op-amp are faster-speed (with light loads) and guaranteed stability.

The Voltage-Mode DAC

The simplest voltage-mode DAC is the *R-2R* string shown in Fig. 30.14. We should recognize this circuit from both Figs. 30.2 and 30.8. For the moment we assume the load is purely capacitive so that errors resulting from sourcing a DC current are not present in the DAC. Here, through several examples, we discuss settling time, resistor voltage coefficient, matching, and the effects of a DC load. While the fact that this DAC can't drive an arbitrary load resistance may be an issue in a particular design, the simplicity and small size of the topology make it very useful in many situations.

IPR2023-00698
Theta EX2019

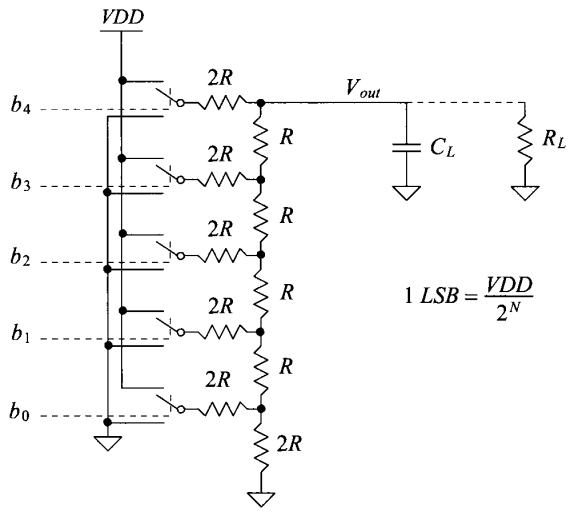


Figure 30.14 Voltage-mode (5-bit) DAC without an op-amp.

Example 30.1

Suppose a 10-bit, voltage-mode DAC with the topology seen in Fig. 30.14 is implemented where R = 10k and C_L = 10 pF. Estimate the maximum clocking frequency that can be used to clock the register supplying the input words to the DAC. Verify your answer using SPICE.

For complete settling the DAC must be 10-bit accurate to within 0.5 LSBs over its full-scale range

$$\text{Accuracy} = \frac{0.5\ LSB}{\text{Full scale range } (VDD)} = \frac{VDD/2^{N+1}}{VDD} = \frac{1}{2^{11}} = 0.04883\%$$

The time constant associated with the DAC and capacitive load is

$$RC_L = 10k \cdot 10p = 100\,ns$$

This time constant can be related to the final ideal output voltage, $V_{outfinal}$, and the actual output voltage, V_{out}, using

$$V_{out} = V_{outfinal}(1 - e^{-t/RC_L})$$

or, relating this to the required accuracy,

$$\frac{1}{2^{N+1}} = 1 - \frac{V_{out}}{V_{outfinal}} = e^{-t_{settling}/RC_L}$$

The required settling time is then

$$t_{settling} = RC_L \cdot \ln 2^{N+1} \qquad (30.18)$$

IPR2023-00698
Theta EX2019

Using the numbers from this example results in $t_{settling}$ = 762 ns. The SPICE simulation results are shown in Fig. 30.15. The maximum clock frequency is then estimated as

$$f_{clk,\max} = \frac{1}{t_{settling}} = \frac{1}{RC_L \cdot \ln 2^{N+1}} \tag{30.19}$$

For this example, $f_{clk,\max}$= 1.3 MHz. Note that the fundamental way to decrease the settling time is to decrease the resistance in the R-$2R$ ladder (assuming we have no control over the load capacitance). The practical problem then becomes implementing the switches (MOSFETs) with a resistance small compared to R. ■

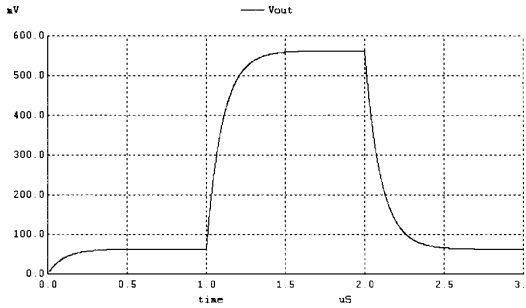


Figure 30.15 Example output for the 10-bit DAC in Ex. 30.1 showing settling time limitations.

Example 30.2

Suppose that the $2R$ MSB resistor in the DAC described in Ex. 30.1 experiences a 0.5% mismatch. Estimate the resulting DAC's INL and DNL. Use SPICE to verify your answer.

The 0.5% mismatch ($\Delta R/R$ or 1 σ [standard deviation]) is the mismatch specified for the unsilicided n+ polysilicon resistors in Table 4.1. Again it is desirable to use poly resistors because they sit above the substrate, on the field oxide, and are more immune to substrate noise. Note that the voltage coefficient can (will) also cause nonlinearities. However, instead of the worst-case situation of an abrupt mismatch between the lower resistors and the MSB $2R$ resistor, as used in this example, a first-order voltage coefficient error will cause a linear variation of the resistor values from the LSB resistor up to the MSB resistor and so the effects of the voltage coefficient, for reasonably small values, are generally not significant compared to the random mismatch effects.

Rewriting Eqs. (30.14) and (30.15) to estimate the maximum number of bits possible with 1 LSB INL or DNL results in

$$N = 1 - 3.3 \cdot \log\left(\frac{\Delta R}{R}\right) \tag{30.20}$$

Using this equation with $\Delta R/R = 0.005$ results, again, in N = 8.6 bits. For a 10-bit DAC, we would estimate both the INL and DNL as 2.4 bits.

IPR2023-00698
Theta EX2019

To verify these results using SPICE, let's input a code of 01 1111 1111 (ideally 499 mV) and then step the input code to 10 0000 0000 (ideally, 500 mV). With the MSB $2R$ resistor changed to 20.1k (a 0.5% mismatch from its ideal 20k value) the simulation results are shown in Fig. 30.16. With this mismatch the output of the DAC is 500.3 mV when the input is 01 1111 1111. The INL with this input code is 1.25 LSBs (roughly 1.5 mV). The INL when the input digital code is 10 0000 0000 is −1.25 LSBs. The DNL at this worst-case point is −2.5 LSBs. Note that the DAC is nonmonotonic (DNL < −1 LSB). An increase in the digital input code results in a decrease in the output voltage. Nonmonotonic DACs can result in circuits that don't function properly (an example being a successive approximation ADC). A DNL of −1 LSB would indicate the output voltage of the DAC doesn't change when the input code changes.

To improve the DNL, the upper bits of the DAC must be segmented as seen in Fig. 30.6. Improving the INL relies on calibrating out the mismatch errors (see Figs. 30.12 and 30.13). Also note, again, that mismatch can be improved by layout techniques (e.g., common-centroid) and by averaging the outputs of multiple resistor strings. ■

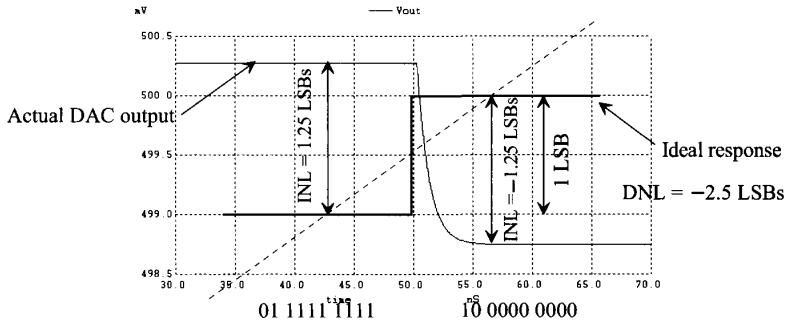


Figure 30.16 Output if MSB resistor in Fig. 30.14 experiences a 0.5% mismatch.

We can characterize the effects of a DC load resistance, R_L, as seen in Fig. 30.14, by noticing that R_L forms a divider with the R-$2R$ ladder. The LSB with a load can be written as

$$1\ \text{LSB} = \frac{VDD}{2^N} \cdot \frac{R_L}{R + R_L} \tag{30.21}$$

Notice that if $R_L \to \infty$, this equation reduces to the LSB value given in Fig. 30.14. The time constant associated with driving an output capacitance can now be written as

$$\tau = R \| R_L \cdot C_L \tag{30.22}$$

Two Important Notes Concerning Glitches

Note that we have assumed that the RC delay through the resistors used in the R-2R ladder is negligible. This may not be the case in many practical situations (especially if diffused or implanted resistors are used), resulting in a DAC output glitch. Also, we have

IPR2023-00698
Theta EX2019

been simulating with perfectly aligned digital signals, that is, signals that change at the exact same moment. When the digital signals are slightly misaligned, a significant glitch can occur in the DAC's output. *This means that the inputs to the DAC should be provided by the same digital hold register.* Using segmentation with the required thermometer decoder can result in the digital signals driving the *R*-2*R* ladder seeing differing delays. Care must be exercised when designing the DAC input clocking circuit (e.g., add small dummy delays).

Example 30.3

Repeat Ex. 30.2 if a 200 ps skew is experienced by the lower nine bits in the digital inputs with relation to the MSB.

The simulation results are shown in Fig. 30.17. When comparing this result to Fig. 30.16, the magnitude of the glitch in relation to the much smaller final step in the output voltage should be obvious. The small 200 ps skew in the digital inputs causes a code of 00 0000 0000 to be applied to the DAC for 200 ps. For this very short period of time the output begins to discharge from 500 mV down to ground. Note in this figure and in Fig. 30.16 the load capacitance was reduced to 0.1 pF to decrease the settling time. ■

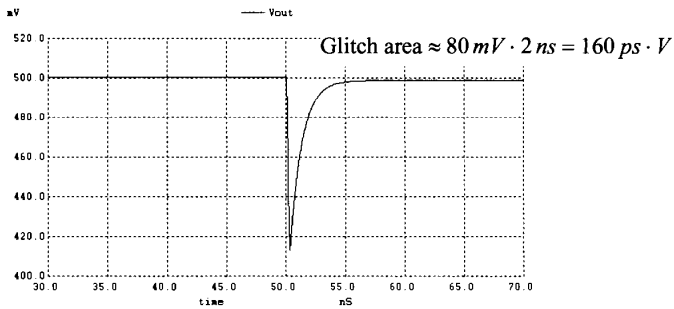
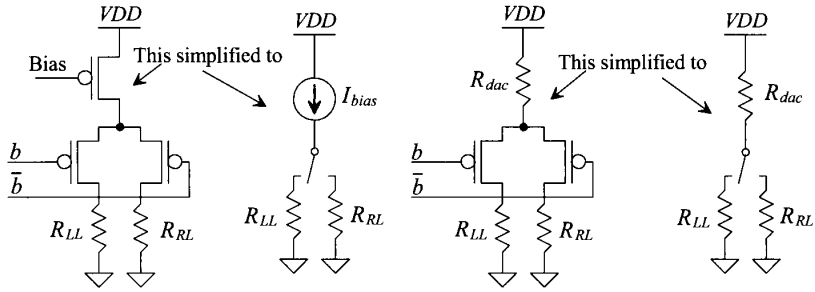


Figure 30.17 Showing glitch if the lower 9-bits are skewed by 200 ps in Ex. 30.2.

The Current-Mode (Current Steering) DAC

Figure 30.18 shows the two basic cells used in the implementation of a current-mode DAC. In this section we focus on the use of the current-source based cell. The advantage of the current-source cell is the fact that the value of the current can be adjusted, via the bias voltage, to compensate for process variations, while the value of the resistor, in the resistor-based cell is fixed. The advantages of the resistor-based cell are wider output swing (the MOSFET current source must remain in the saturation region), better voltage coefficient (no channel length modulation or other finite MOSFET output resistance effects), and better substrate noise immunity (assuming the resistors are integrated on the top of the field oxide and not down in the substrate with the MOSFETs). Notice, in this figure, that we've implemented the cells with two complementary outputs. Having complementary outputs is useful, for example, in a DAC used with a video monitor (driving two complementary 75 Ω loads). The load resistors are labeled the left load resistor, R_{LL}, and the right load resistor, R_{RL}.

IPR2023-00698
Theta EX2019



Current-source-based cell. Resistor-based cell.

Figure 30.18 Basic cell used in a current-mode DAC.

Figure 30.19 shows the block diagram implementation of a current-mode DAC. The output voltages depend on both I_{REF} and the load resistors. The current sources are implemented using the PMOS cell in Fig. 30.18 and a PMOS *W*-2*W* mirror (see Fig. 30.20 for the NMOS version of the circuit) to improve layout area. The combination of binary-weighted current sources must be used together with the required segmentation of the upper bits to reduce DNL. Although not drawn so in Fig. 30.18, the current sources can be cascoded to increase their output resistance (decrease their voltage coefficient). Again, the switches connected to the loads should be controlled by signals from the same register to avoid significant glitches in the outputs.

While the matching requirements of current-mode (current steering) DACs were discussed in Ch. 29, we should comment on ways to improve matching. The layout of the *W*-2*W* ladder should follow the basic techniques discussed in Ch. 20, i.e., devices oriented the same way, use of dummy poly and source/drain implants, attempt to keep the source-drain voltages constant, and use of long *L* devices. The layout of the *W*-2*W* mirror should look similar to the layout of the *R*-2*R* string in Fig. 5.28, but it can also employ two or more *W*-2*W* segments to average variations. The upper segmented bits can be laid out adjacent to the *W*-2*W* and can also benefit from averaging the outputs of several DAC layouts. In some cases the number of bits used in the *W*-2*W* ladder equals the number of bits used for the upper segments. For example, a 10-bit DAC would use a 5-bit *W*-2*W* ladder (whose MSB is $I_{REF}/2$) and 31 segments with values of I_{REF}. To improve INL separate layouts can be connected together to average the random mismatch effects and increase the linearity of the DAC. Of course, the current output levels increase for the same reference biasing levels.

Finally, as mentioned earlier, we can use a *W*-2*W* current mirror to generate the binary weighted currents used in a current-steering DAC, Fig. 30.20. What we are going to do, with the binary weighted current mirror, is utilize the fact that MOSFETs in series and parallel can be combined, as seen in this figure, into single MOSFETs. Two MOSFETs placed in series with equal widths are equivalent to a single MOSFET with the sum of the two MOSFET's lengths. Two MOSFETs in parallel with the same length sum to form a single MOSFET with the sum of the two MOSFET's widths. The benefit of this topology is the removal of the effects of oxide encroachment and lateral diffusion.

IPR2023-00698
Theta EX2019

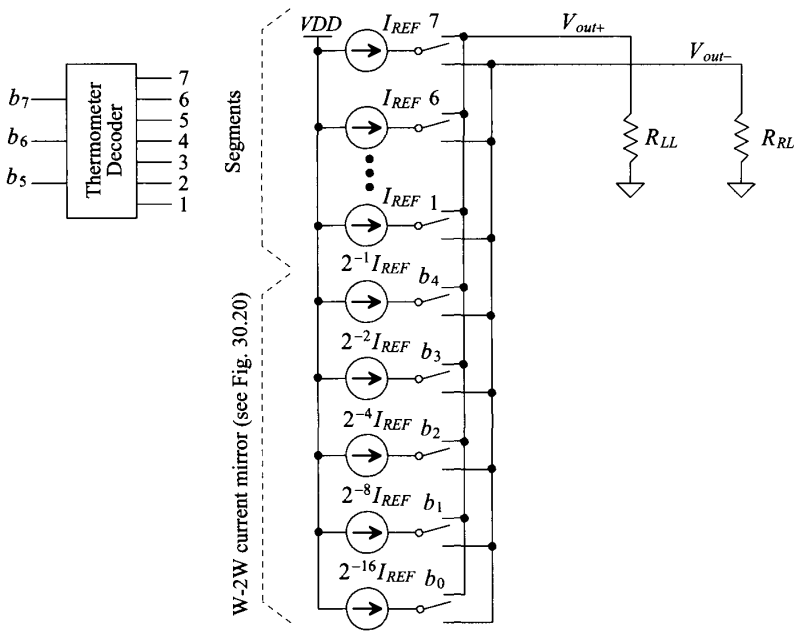


Figure 30.19 Implementation of a current-mode DAC.

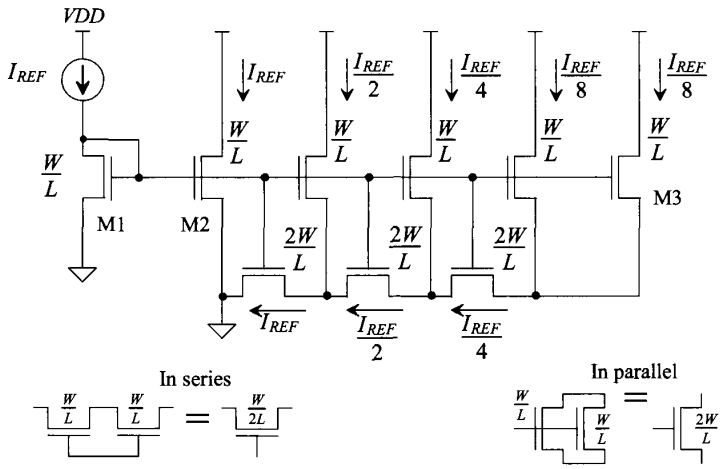


Figure 30.20 W-2W current mirror.

IPR2023-00698
Theta EX2019

30.2 Op-Amps in Data Converters

The open-loop magnitude and phase responses of a typical op-amp are shown in Fig. 30.21. In this section we discuss the gain and bandwidth requirements of op-amps used in either a DAC or an ADC. *We assume* that the op-amp is designed to have a phase margin of 90 degrees under full load conditions and over process variations. (We should point out that this assumption is easily met using an OTA that is compensated by a load capacitance as discussed in Ch. 24.) It's important to understand why having a 90-degree phase margin is important, namely, to avoid a second-order step response with the associated ringing. If the phase margin is 90 degrees, we get an RC-like settling response shape as seen in Fig. 30.15. Figure 26.60 shows the step response of a mixed-signal op-amp. The phase margin of this op-amp was less than 90 degrees and thus the step response shows ringing. Decreasing the phase margin increases the peak amplitude of the ringing and can lengthen the settling time (the time it takes the op-amp's output to settle to within 1/2 LSB of the ideal final value).

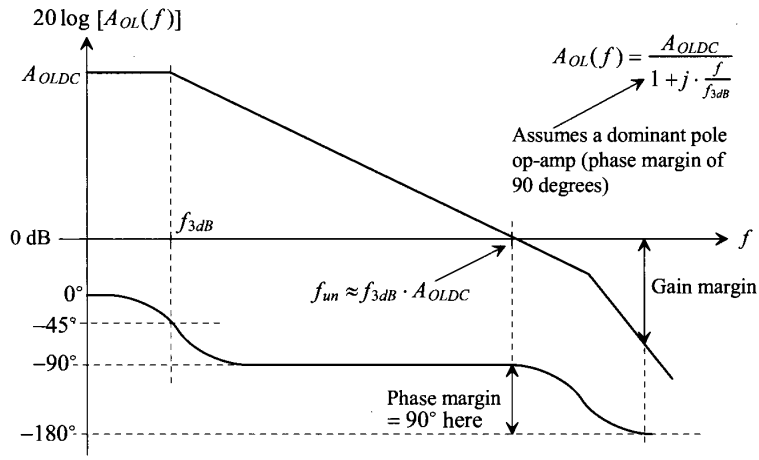


Figure 30.21 Magnitude and phase responses of an op-amp.

Gain Bandwidth Product of the Noninverting Op-Amp Topology

Figure 30.22 shows the basic topology of a noninverting op-amp amplifier. The voltage on the inverting op-amp input can be written as

$$v_- = v_{out} \cdot \overbrace{\frac{R_1}{R_1 + R_2}}^{\beta} \tag{30.23}$$

where β is the feedback factor for this *series-shunt* feedback amplifier (see Ch. 31, the ideal closed-loop gain, A_{CL}, is 1/β or 1 + R_2/R_1). The output of the amplifier is

$$v_{out} = (v_{in} - v_-) \cdot A_{OL}(f) \tag{30.24}$$

Solving these equations for the closed-loop bandwidth of the amplifier, $f_{CL,3dB}$, gives

IPR2023-00698
Theta EX2019

$$f_{CL,3dB} \approx \beta \cdot A_{OLDC} \cdot f_{3dB} = \beta \cdot f_{un} \tag{30.25}$$

The gain bandwidth product of the noninverting amplifier is then

$$\text{Gain} \cdot \text{bandwidth} = f_{un} \tag{30.26}$$

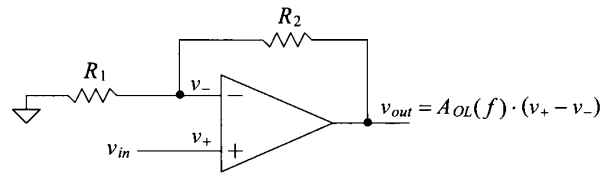


Figure 30.22 Noninverting op-amp topology.

Gain Bandwidth Product of the Inverting Op-Amp Topology

Figure 30.23 shows the schematic diagram of an inverting op-amp topology. Summing the currents at the inverting input node gives

$$\frac{v_{in} - v_-}{R_1} = \frac{v_- - v_{out}}{R_2} \tag{30.27}$$

The output of the amplifier is related to the op-amp's input terminals using

$$v_{out} = (-v_-) \cdot A_{OL}(f) \tag{30.28}$$

Solving these two equations for the closed-loop bandwidth once again results in Eq. (30.25) with β defined as indicated in Eq. (30.23). This can be confusing because the feedback factor, β, for the inverting amplifier is not the same as for the noninverting amplifier. The inverting op-amp is an example of a *shunt-shunt* amplifier (current input and voltage output). The feedback factor for this amplifier is $-1/R_2$ (= β) where

$$\frac{v_{out}}{i_{in}} = -R_2 \tag{30.29}$$

If we assume the input current source has a source resistance of R_1 so that $v_{in} = i_{in} \cdot R_1$ then

$$|A_{CL}| = \left|\frac{v_{out}}{v_{in}}\right| = \frac{R_2}{R_1} \tag{30.30}$$

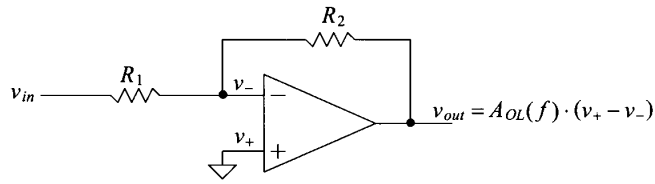


Figure 30.23 Inverting op-amp topology.

IPR2023-00698
Theta EX2019

Keeping in mind that the closed-loop bandwidth of the inverting amplifier is still, from Eq. (30.25),

$$f_{CL,3dB} \approx \frac{R_1}{R_1 + R_2} \cdot f_{un} \tag{30.31}$$

we can write

$$\text{Gain} \cdot \text{bandwidth} = \frac{R_2}{R_1 + R_2} \cdot f_{un} \tag{30.32}$$

Example 30.4

Compare the bandwidth of a +1 gain amplifier implemented using a noninverting op-amp topology (Fig. 30.22) to the bandwidth of a −1 gain amplifier using the inverting op-amp topology (Fig. 30.23).

Using Eq. (30.26), the bandwidth of the +1 gain amplifier is f_{un}. This amplifier is commonly known as a unity voltage follower and has $R_1 = \infty$ (an open) and $R_2 = 0$ (a short). The bandwidth of the inverting, −1, gain amplifier can be determined using Eq. (30.32) with $R_1 = R_2$ and is $0.5f_{un}$. *This result is important because it shows that for the fastest speed the noninverting op-amp topology offers the best choice.* Practically, however, the nonlinearities related to the finite *CMRR* (see Eq. [30.8]) force the use of inverting op-amp topologies. As discussed earlier, the *input common-mode voltage must remain constant* in any precision application. Note that a fully-differential op-amp topology is also a *shunt-shunt* amplifier with a gain bandwidth product given by Eq. (30.32). ■

Example 30.5

Comment on the derivations of Eqs. (29.59)-(29.60) in the last chapter.

The equations are still valid; however, if the op-amp topology is designed to keep the common-mode input voltage constant we would need to halve the value of β used in the equations. For example, for an ideal amplifier $\beta = 1/(2A_{CL})$. ■

30.2.1 Op-Amp Gain

In this section we answer the question of how large the DC open-loop gain of the op-amp, A_{OLDC}, must be in a data converter with a resolution of N bits. We know that the op-amp must amplify signals to within 1/2 LSB of the ideal value. Further we know that the closed-loop gain of an amplifier can be written as

$$A_{CL} = \frac{A_{OL}(f)}{1 + \beta \cdot A_{OL}(f)} \tag{30.33}$$

The feedback factor can be written, after reviewing Figs. 30.23 or 30.24, as

$$\beta = \frac{R_1}{R_1 + R_2} \text{ or } \frac{C_F}{C_F + C_I} \left[= \frac{1/j\omega C_I}{1/j\omega C_I + 1/j\omega C_F} \right] \tag{30.34}$$

As discussed in Ch. 29 the output of the amplifier will be equal to its ideal value minus some maximum deviation, ΔA. We can write the gain of the amplifier in Fig. 30.24

$$|A_{CL}| = \frac{C_I}{C_F} \tag{30.35}$$

IPR2023-00698
Theta EX2019

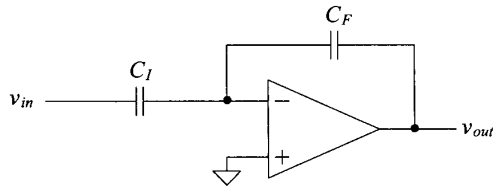


Figure 30.24 Inverting op-amp topology.

Next, let's write

$$|A_{CL}| = \frac{C_I}{C_F} - \Delta A = \frac{A_{OLDC}}{1 + A_{OLDC} \cdot \frac{C_F}{C_F + C_I}} \tag{30.36}$$

If the maximum value of ΔA is at most 1/2 LSB of the ideal gain, or,

$$\Delta A = \frac{C_I}{C_F} \cdot \frac{1/2\ LSB}{\text{Full scale output}} = \frac{C_I}{C_F} \cdot \frac{1/2 \cdot (V_{REF+} - V_{REF-})/2^N}{(V_{REF+} - V_{REF-})} = \frac{C_I}{C_F} \cdot \frac{1}{2^{N+1}} \tag{30.37}$$

then we can estimate the minimum required DC open-loop gain as

$$|A_{OLDC}| \geq \frac{1}{\beta} \cdot 2^{N+1} \tag{30.38}$$

If $\beta = 1/2$, as in the R-$2R$ DAC of Fig. 30.3 or when $C_I = C_F$, then

$$|A_{OLDC}| \geq 2^{N+2} \tag{30.39}$$

A 12-bit ADC or DAC requires the use of an op-amp with a gain greater than 16k while a 16-bit converter must have $|A_{OLDC}| \geq 256k$. Clearly, this estimate can present a real design concern. Note that Eq. (30.38) is optimistic. For a general design, an error of 1/2 LSB due just to op-amp gain is not desirable (so a larger value of A_{OLDC} must be used).

30.2.2 Op-Amp Unity Gain Frequency

The speed of a data converter is mainly limited by the op-amp used. In general, the minimum op-amp gain-bandwidth product, f_{un}, required for a specific settling time t (where t is less than $1/2f_{clk} = T_{clk}/2$, within a dead band of ±1/2 LSB) can be estimated, assuming no slew-rate limitations (see also Eqs. [30.18] and [30.19]), by

$$v_{out} = v_{outfinal}\left(1 - \frac{1}{2^{N+1}}\right) = v_{outfinal}(1 - e^{-t/\tau}) \tag{30.40}$$

where, once again,

$$\tau = \frac{1}{2\pi \cdot \beta \cdot f_{un}} \tag{30.41}$$

The minimum required op-amp unity gain frequency is then given by

$$f_{un} \geq \frac{f_{clk} \cdot \ln 2^{N+1}}{\pi \cdot \beta} \tag{30.42}$$

or, again assuming $\beta = 1/2$,

$$f_{un} \geq 0.44 \cdot (N+1) \cdot f_{clk} \tag{30.43}$$

IPR2023-00698
Theta EX2019

If we design a 12-bit ADC that is clocked at 100 MHz, we need to use op-amps with unity gain frequencies, f_{un}, of 572 MHz (and a DC gain of at least 16k). Again, this estimate for the unity gain frequency is optimistic. A good design would use a larger f_{un} than what is specified by Eq. (30.43).

30.2.3 Op-Amp Offset

A critical characteristic of any op-amp used in a data converter is its offset voltage. We introduced the concept of reducing the offset voltage of an op-amp back in Ch. 25. Here we provide additional comments and possibilities for offset reduction.

Adding an Auxiliary Input Port

A simple method of nulling the offset voltage of an op-amp is shown in Fig. 30.25. In this figure the added MOSFETs, M1 and M2 (which operate in the triode region) are essentially used to balance the current flowing in the current mirror load. We can think of the added MOSFETs as providing an auxiliary input port for offset calibration.

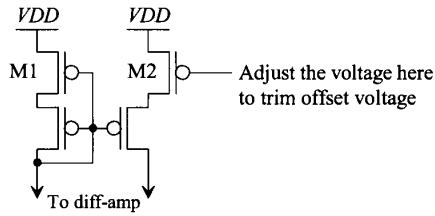


Figure 30.25 Trimming offset using an auxiliary input port.

Figure 30.26 shows how we would use the auxiliary input port to remove (lower) the offset. When zeroing out the offset, the op-amp is removed from the circuit by opening S1 (and possibly a switch [not shown] in series with the op-amp's output). This is followed by closing S2 and S3 so that a control voltage is stored on C. Note that we have assumed that the op-amp is used in an inverting configuration (that is, the noninverting input of the op-amp, +, is tied to V_{CM}). The offset removal is dynamic and will have to be performed periodically. We could also use a simple R-$2R$ DAC with a topology similar to what is seen in Fig. 30.8 to calibrate out the offset (eliminating the dynamic nature of the method). The output of the DAC would be connected to the auxiliary input port. Note that an increase in voltage on the auxiliary input port must result in a decrease in output

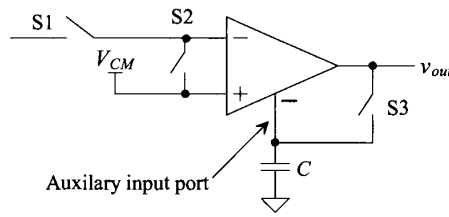


Figure 30.26 Using an auxiliary input port to lower offset.

IPR2023-00698
Theta EX2019

voltage. (There must be negative feedback when connecting the output of the op-amp to the input port.)

The practical problem with the topology of Fig. 30.26 is the charge injection and capacitive feedthrough resulting from shutting off (opening) S3. This "glitch" of charge causes a change in the auxiliary port's input voltage and can place a significant limitation on the minimum possible offset voltage attainable after calibration. The amplitude of the glitch can be reduced by increasing C or by increasing the length of the MOSFET used in the op-amp (M2 in Fig. 30.25). Increasing the length results in a decrease in the MOSFET's transconductance (keeping in mind that the MOSFET is operating in the triode region) making the amplitude of the glitch less harmful. The drawback of increasing the MOSFET's length is that the range of offset voltages we can remove is reduced.

Example 30.6

Suppose perfect switches are available for the circuit of Fig. 30.26. Estimate the residual offset voltage in terms of the op-amp's gain, A_G, from the auxiliary port to the op-amp output.

If the offset voltage before reduction is V_{OS}, then the offset voltage after reduction is V_{OS}/A_G. For reasonable values of A_G the final inherent offset voltage is negligible. The point of this example is that the charge injection and capacitive feedthrough from the switches is the dominant source of offset error using this technique. ■

We've seen the problem of charge injection and capacitive feedthrough before. The most common technique for reducing its effect is to use a fully-differential topology. Figure 30.27 shows a modification of Figs. 30.25 and 30.26 to compensate for charge injection. The idea is that when S4 and S3 turn off (open) the variation in voltages on the gates of M1 and M2 are equal resulting in a common change in each MOSFET's resistance. Ideally then the current will remain balanced in the diff-amp. Note that while we've shown the use of triode-operating MOSFETs M1/M2 in Figs. 30.25 and 30.27 in series with the load of a diff-amp on the input of an op-amp, we could also use this concept in later stages of the op-amp.

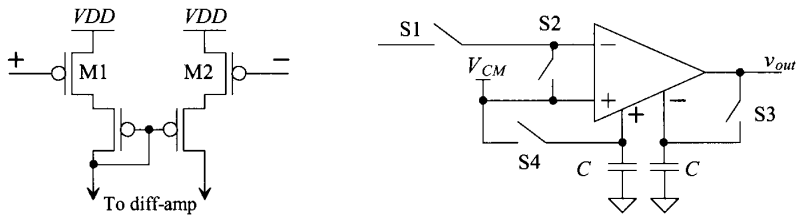


Figure 30.27 Using an auxiliary input port to lower offset (two terminals).

Figure 30.28 shows another possible topology for offset removal using an auxiliary input. An additional diff-amp is added in parallel to the main input diff-amp stage of an op-amp to balance the currents and zero out the offset voltage. Again, long

IPR2023-00698
Theta EX2019

length MOSFETs are used in the added input so that the glitches resulting from the imperfections in the MOSFET switches (S4 and S3 in Fig. 30.27) have the least effect on the operation of the circuit.

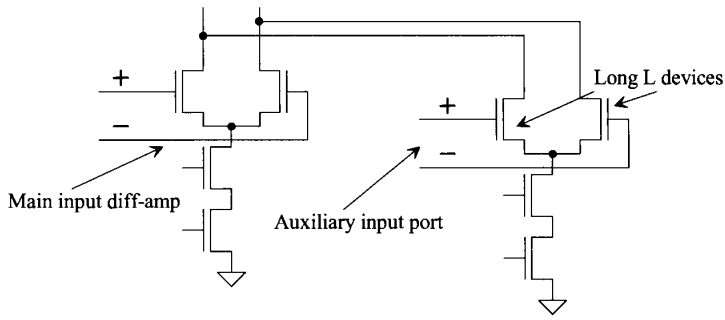


Figure 30.28 Using an auxiliary diff-amp for balancing current in an op-amp's input.

We can estimate the maximum offset voltage we can zero out using the technique of Fig. 30.28 by writing the imbalance in the main diff-amp's currents because of its offset voltage as

$$g_m \cdot V_{OS,\max} = i_d \tag{30.44}$$

The auxiliary input must sum the opposite of this current in the main diff-amp's load to balance the currents in the main diff-amp (and hence eliminate the offset voltage). If we label the transconductance of the diff-amp used in the auxiliary input g_{maux} and the maximum allowable differential voltage on the auxiliary input for linear operation $V_{aux,\max}$ then we can write

$$g_m \cdot V_{OS,\max} = g_{maux} \cdot V_{aux,\max} \tag{30.45}$$

Because we are using long length devices in the auxiliary input $g_{maux} << g_m$ for the same biasing current levels. If $V_{aux,\max} = 200$ mV (a differential voltage of ± 200 mV will cause all of the diff-amp tail current to flow through one side of the diff-amp) and $g_m/g_{maux} = 10$, then we can zero out at most 20 mV of op-amp offset.

The offset storage technique shown in Fig. 30.27 relies on the removal of the op-amp from the circuit while autozeroing the offset. The scheme in Fig. 30.29 shows how the technique can be extended to remove the offset while leaving the main op-amp, O1, in the circuit at all times. When S2, S3, and S4 are closed, the offset of O2 is zeroed out. At this time switches S1 and S5 are open. After O2's offset is stored, S2, S3, and S4 are then opened. Next S1 and S5 close. O2 is used to precisely set the inverting input of O1 to V_{CM} through the feedback around O1 (not shown). When O2 goes back to zeroing out its own offset (S2-S4 close) the capacitor connected to the auxiliary port of O1 retains the charge, and thus voltage, needed to keep O1's offset nulled out to zero. Again this capacitor should be large to avoid problems from the imperfections of S5.

IPR2023-00698
Theta EX2019

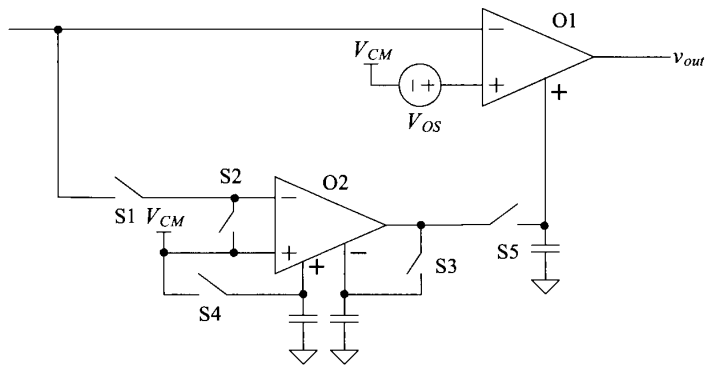


Figure 30.29 Continuous-time offset removal.

30.3 Implementing ADCs

In this section we continue to discuss implementing data converters with design concerns for S/Hs, cyclic ADCs, and pipeline ADCs.

30.3.1 Implementing the S/H

We assume that the reader is familiar with the fundamental implementation of a CMOS S/H discussed in Ch. 25 and in particular bottom-plate sampling. Figure 30.30 shows the more general implementation of a S/H. Note that if C_I goes to 0 (an open) this topology reduces to the basic S/H given in Ch. 25 (repeated in Fig. 30.31 for convenience).

We can determine the relationship between the input of the S/H and its output by writing the charge stored on C_I and C_F when the ϕ_1 and ϕ_2 switches are closed (the ϕ_3 switches are open) as

$$Q_{I,F}^{\phi_1} = C_{I,F} \cdot (v_{in} - V_{CM} \pm V_{OS}) \tag{30.46}$$

where V_{OS} is the offset voltage of the op-amp and the input (and output) voltages are referenced to ground (v_{in} [v_{out}] varies from 0 to $2V_{CM}$ [= VDD here]). Note that the reason why the ϕ_2 switches turn off slightly after the ϕ_1 switches was discussed in Sec. 25.2.1 (bottom plate sampling, an important technique used in the circuits we present here). When ϕ_3 goes high, the charge on C_I is

$$Q_I^{\phi_3} = C_I \cdot (V_{CM} - V_{CM} \pm V_{OS}) \tag{30.47}$$

The difference between $Q_I^{\phi_1}$ and $Q_I^{\phi_3}$ is transferred to C_F when ϕ_3 goes high. The output voltage is then determined knowing charge must be conserved

$$C_F \cdot (v_{out} - V_{CM} \pm V_{OS}) = \overbrace{C_F \cdot (v_{in} - V_{CM} \pm V_{OS})}^{Q_F^{\phi_1}} + \overbrace{C_I \cdot (v_{in} - V_{CM} \pm V_{OS})}^{Q_I^{\phi_1}} - \overbrace{C_I \cdot (V_{CM} - V_{CM} \pm V_{OS})}^{Q_I^{\phi_3}} \tag{30.48}$$

IPR2023-00698
Theta EX2019

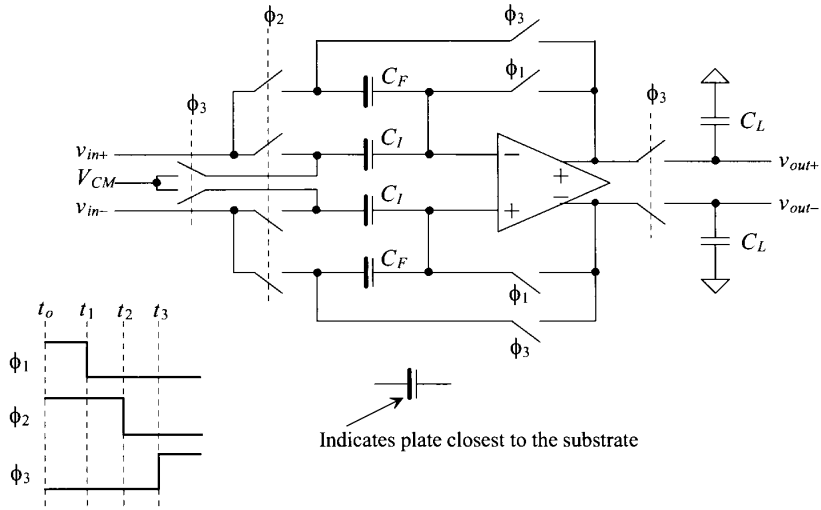


Figure 30.30 Data converter S/H building block.

or when ϕ_3 goes high

$$v_{out} = \left(1 + \frac{C_I}{C_F}\right) \cdot v_{in} - \frac{C_I}{C_F} \cdot V_{CM} \tag{30.49}$$

Notice how the op-amp offset is autozeroed out. The ideal residual offset is V_{OS}/A_{OL}. Practically, the residual offset is limited by the imperfections in the switches (which, once again, forces us to use fully-differential topologies). Also note, in Fig. 30.30, that we have drawn the input capacitance of the next stage as a load, C_L. This was so that the output of the S/H would appear to change only on the rising edge of ϕ_3 (plus the output settling time). Finally, if the S/H is clocked at $f_{clk} = 1/T_{clk}$, the output of the S/H must settle to less than 1/2 LSB, worst-case, in a time of $T_{clk}/2$. The minimum required value of op-amp unity gain frequency was specified in Eq. (30.42).

Equation (30.49) can be used to determine the relationship between v_{in} and v_{out} for fully-differential signals

$$v_{out} = v_{out+} - v_{out-} = \left(1 + \frac{C_I}{C_F}\right) \cdot (v_{in+} - v_{in-}) \tag{30.50}$$

Note how, as we would expect, the common-mode voltage subtracts out of the relationship when we take the difference between v_{out+} and v_{out-}. A block diagram representing the S/H of Fig. 30.30 is shown in Fig. 30.32. The use of block diagrams can be very useful to describe data converter architectures.

Example 30.7

Simulate the operation of the data converter S/H building block shown in Fig. 30.30. Assume $C_I = C_F = 1$ pF and $f_s = 100$ MHz.

IPR2023-00698
Theta EX2019

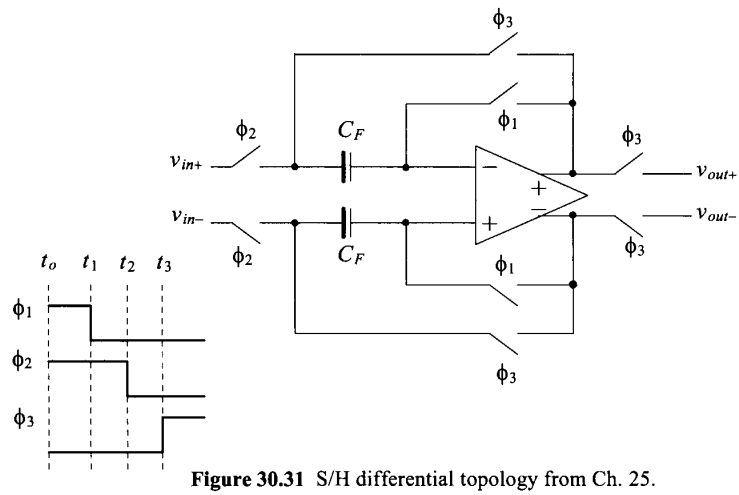


Figure 30.31 S/H differential topology from Ch. 25.

The simulation results are shown in Fig. 30.33. In part (a) the clock signals are shown. Unlike the clock signals shown in Fig. 30.30 where the falling edge of ϕ_2 is delayed from ϕ_1, the simulation sets the signals so they go low at the same time. This was to avoid the outputs of the op-amp changing to very large values for the small amount of time the op-amp operates open-loop with an input signal applied.

In part (b) we show the op-amp outputs. Note how, when ϕ_1 goes high, both outputs are set to the common-mode voltage by forcing the op-amp into a follower configuration (which may lead us to use switches to short the terminals of the op-amp to V_{CM} when ϕ_1 is high if offset isn't important). When ϕ_3 goes high, the circuit behaves as an S/H with a gain of two. Part (c) of the figure shows the outputs connected through ϕ_3 switches, as seen in Fig. 30.30, driving 10 pF load capacitances. ■

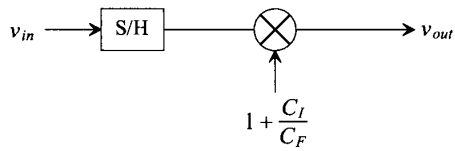


Figure 30.32 Block diagram for the S/H of Fig. 30.30.

A Single-Ended to Differential Output S/H

Note how we have assumed, in the S/H of Fig. 30.30, that the input voltage was fully-differential. In many practical applications the input to the ADC is single-ended. While we can connect v_{in-} to V_{CM} in an attempt to change the single-ended input into a fully-differential sampled output, the practical problem is the variation of the op-amp's

IPR2023-00698
Theta EX2019

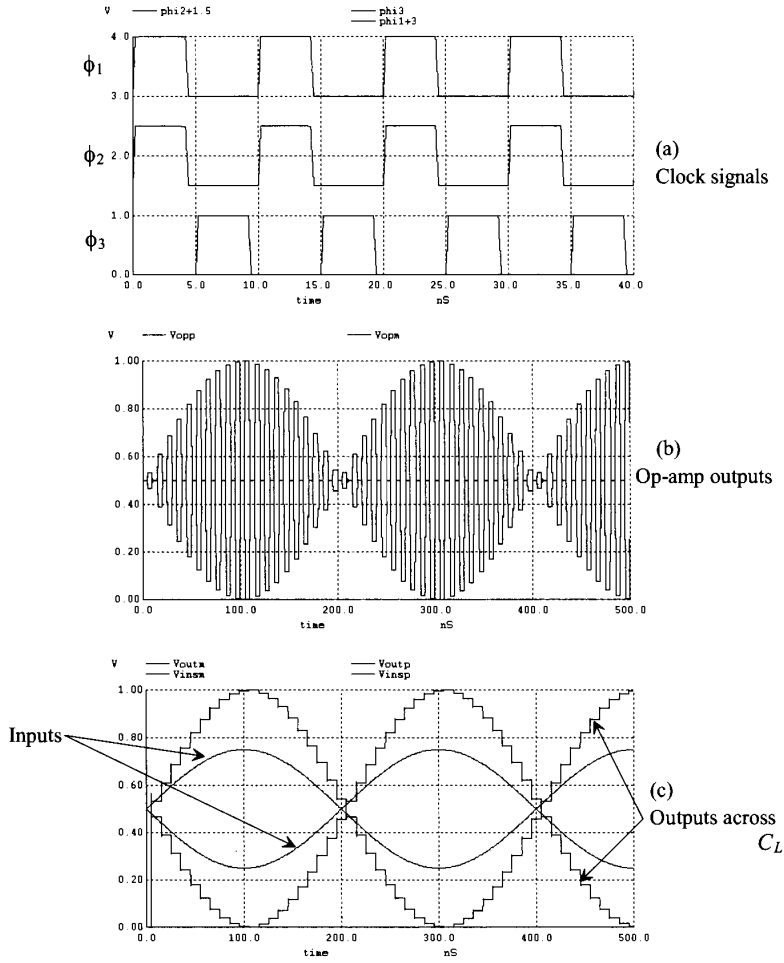


Figure 30.33 SPICE simulations of the operation of Fig. 30.30.

input common-mode voltage. As we've already discussed, precision data converters must use op-amp configurations where the input common-mode voltage is constant. Also, and perhaps more practically, the range of allowable common-mode voltages can be very restricted when designing low-voltage circuits. As we saw when designing mixed-signal op-amps in Ch. 26, the minimum input common-mode voltage can be very close to V_{CM} in nanometer CMOS. A technique to keep the op-amp's input common-mode voltage at V_{CM} when a single-ended input is applied to the S/H is seen in Fig. 30.34. During the sample phase all of the bottom plates of the capacitors are connected to V_{CM} except for one, which is connected to the single-ended input. The op-amp is in the unity-follower configuration

IPR2023-00698
Theta EX2019

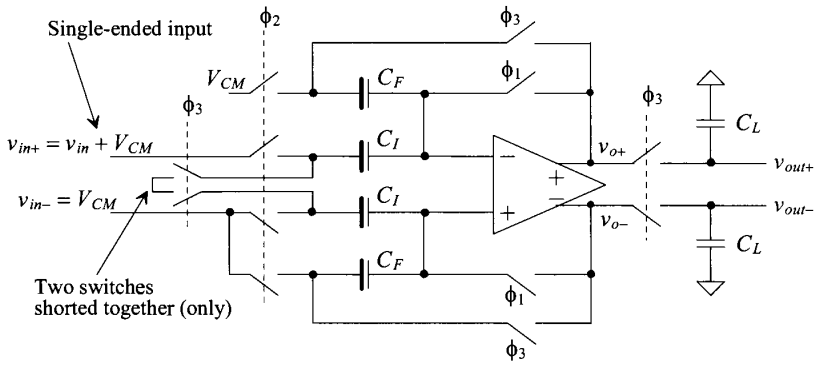


Figure 30.34 Single-ended to differential S/H.

and the bottom-plate sampling techniques we discussed earlier are still used. When the ϕ_3 switches are closed (ϕ_1 and ϕ_2 are open) the bottom plates of the C_I capacitors are shorted together but to nothing else. This causes the charge stored on the bottom-plate of the top C_I capacitor (the capacitor connected to the single-ended input v_{in+}) to redistribute between both C_I capacitors making the input appear to be fully differential. Because of the single-ended input and changed connections to V_{CM}, the gain of the topology is C_I/C_F.

Example 30.8

Simulate the operation of the S/Hs shown in both Figs. 30.31 and 30.34. Assume the S/H is clocked at 100 MHz, v_{in+} is a sinewave that swings from ground to *VDD*, and v_{in-} is connected to V_{CM} (the input signal is single-ended and covers the entire supply range). Show how the op-amp's input common-mode voltage range changes or doesn't change. Assume all capacitors are 1 pF and then show how a 10% mismatch in the capacitors affects the output of the S/H.

Figure 30.35 shows the simulation results for the S/H seen in Fig. 30.34. Notice how the op-amp's input common-mode voltage stays constant at V_{CM}. While the simulation results show the op-amp's outputs, v_{o+} and v_{o-}, and not the difference in the outputs, if we do take the difference (with all capacitors at 1 pF and so the gain of the S/H should be one) we indeed see that the gain is one. Next if we modify the simulation and change one C_I capacitor to 1.1 pF the simulation results show a gain error. As we'll discuss in a moment the big benefit of using the topology seen in Fig. 30.31 is that capacitor matching isn't important.

Next, Fig. 30.36 shows the simulation results for the S/H seen in Fig. 30.31. In part (a) the input common-mode voltage of the op-amp is shown. When the ϕ_1 switches are closed, the voltage returns to V_{CM} (the op-amp is placed in a follower configuration where the input signal charges the two capacitors). The input common mode voltage varies between 750 mV and 250 mV, a concern for most op-amp designs using a single input diff-amp. In part (b) the rail-to-rail input signal is converted into a differential S/H output signal. Note that the gain of the S/H is one. Note also how, unlike the op-amp outputs in Fig. 30.33, the outputs

IPR2023-00698
Theta EX2019

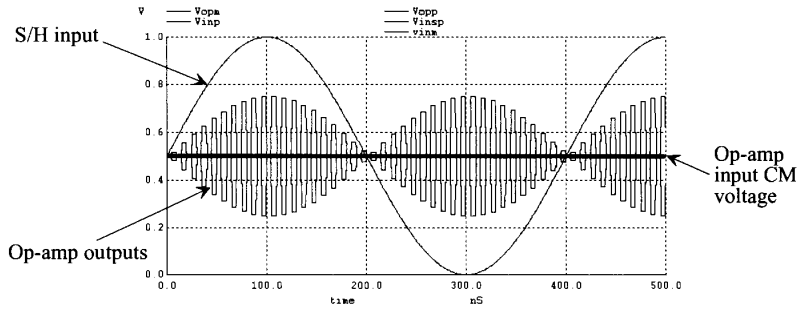
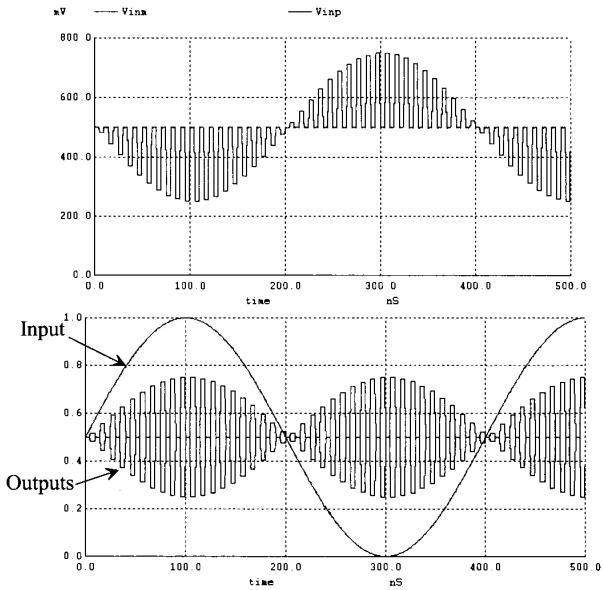


Figure 30.35 Simulating the operation of the S/H in Fig. 30.34.

are limited to $V_{CM} + V_p/2$ where V_p is the peak amplitude of the input sinewave (= 500 mV here). When the input sinewave has an amplitude of 1 V (500 mV above V_{CM}), the positive output is 750 mV and the negative op-amp output is 250 mV. Subtracting the op-amp outputs results in 500 mV (the same voltage as the input signal referenced to V_{CM}). It's important to understand how going from a single-ended signal to a fully-differential signal may result in a reduction in the op-amp output swing.



(a) Showing how the input common-mode voltage can vary in a S/H with a single-ended input.

(b) Input and outputs of a S/H performing single-ended to differential coversion.

Figure 30.36 Simulating the S/H in Fig. 30.31.

IPR2023-00698
Theta EX2019

The simulation results seen in Fig. 30.36 were generated using 0.9 pF and 1.0 pF sampling capacitors (labeled C_F in Fig. 30.31). Because the feedback factor is unity, these capacitors will not affect the gain of the S/H. The point here is that the matching of the capacitors isn't important for a precise gain of one when using this (Fig. 30.31) S/H topology. (The op-amp open-loop gain, however, is still important.) Again, while the topology seen in Fig. 30.34 is sensitive to capacitor matching, it still may be the S/H topology of choice because the op-amp's input CMR remains constant with a single-ended input signal. ■

Before leaving this section let's remember that a common-mode feedback (CMFB) circuit is still required to precisely balance the outputs of the op-amp. In addition to the op-amp designs discussed in Ch. 26, Fig. 30.37 shows another possible design. When the ϕ_1 switches are closed, the outputs are connected to the CMFB amplifier (see Fig. 30.37a). Also when the ϕ_1 switches are closed the op-amp is placed in the unity feedback configuration (not shown). The CMFB circuit is used to ensure that the input voltages are $V_{CM} \pm V_{OS}$. The op-amp in Fig. 30.37b is derived from the designs presented in Ch. 26 but without the output buffers. The benefit of removing the output buffers when driving purely capacitive loads is the ease of attaining a 90-degree phase margin (and so clean settling behavior). The drawbacks of not using an output buffer are the reduced gain and the need to increase the biasing currents and device sizes to drive a given load capacitance. Again, the gain-boosting amplifiers labeled N and P in part (b) can be compensated, if needed, using capacitors at their outputs to ground (or *VDD*). The CMFB amplifier is seen in Fig. 30.37c.

30.3.2 The Cyclic ADC

Cyclic or algorithmic DACs were first discussed back in Sec. 29.1.5. Here we present the concept of a cyclic ADC. A block diagram of a cyclic ADC is seen in Fig. 30.38 assuming an 8-bit ($N = 8$) conversion. The input signal is sampled on the rising edge of every *eighth* (N) clock pulse. On the rising edge of every clock pulse the comparator determines if the S/H input is above or below the common-mode voltage. If it is below V_{CM}, nothing is subtracted from the S/H output. If it is above V_{CM}, then V_{CM} is subtracted from the S/H output. In either case the resulting output is multiplied by two and cycled back to the S/H input. Each time the comparator output goes high the value is stored in a shift register. When the conversion is complete, the digital word stored in the shift register, which corresponds to the analog input voltage, is shifted into a hold register. The next conversion then begins on the following clock pulse starting with sampling the input voltage, v_{in}. Note that it takes N clock cycles for one conversion.

Example 30.9

Determine the output of the ADC in Fig. 30.38 if the input voltage is 1 V.

We begin by sampling the input voltage of 1 V. The output of the comparator is a logic 1 (MSB). Next, V_{CM} (= 0.5 V) is subtracted from the S/H output resulting in an output of 0.5 V. This output is multiplied by 2 resulting in 1 V. This 1 V output (the output of the multiplier) is cycled back to the S/H input.

Next, we sample the fed back voltage of 1 V, the output of the comparator is, again, a logic 1 (MSB – 1). V_{CM} (= 0.5 V) is subtracted from the S/H output

IPR2023-00698
Theta EX2019

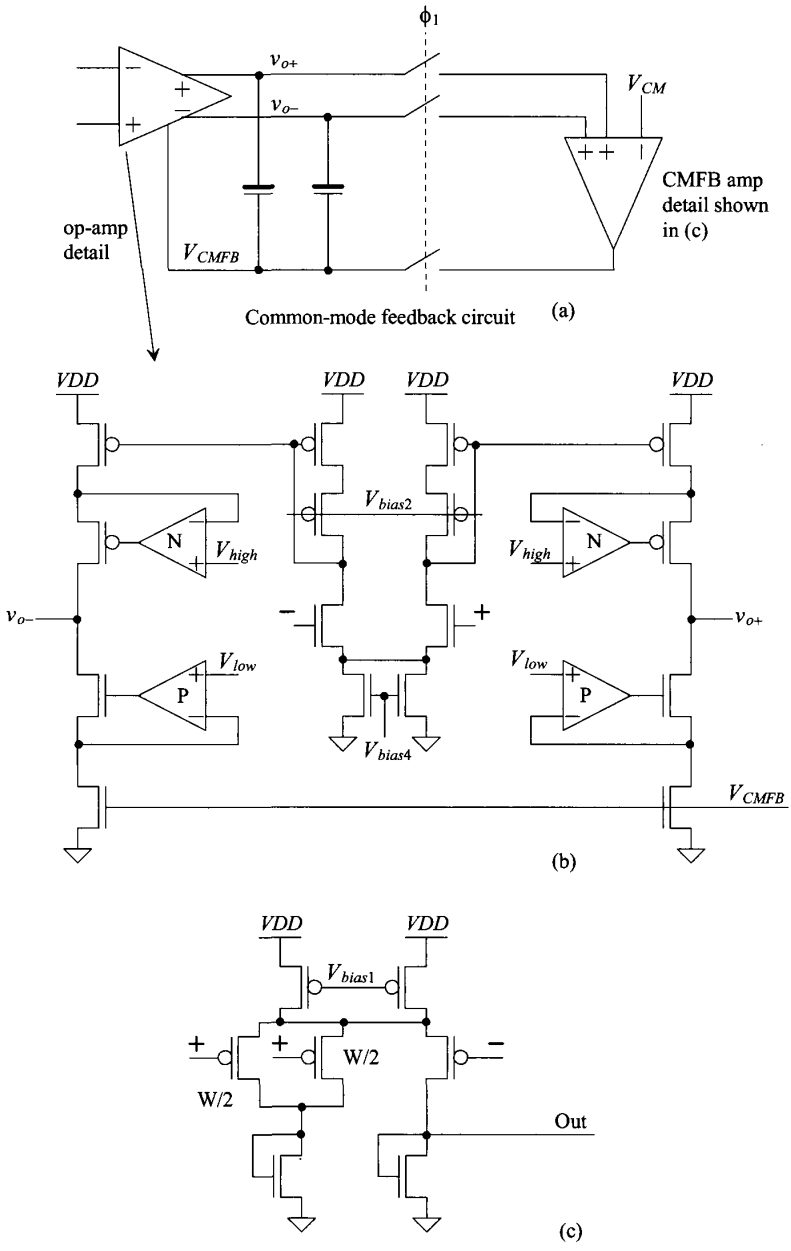


Figure 30.37 Mixed-signal op-amp for use in a S/H with CMFB.

IPR2023-00698
Theta EX2019

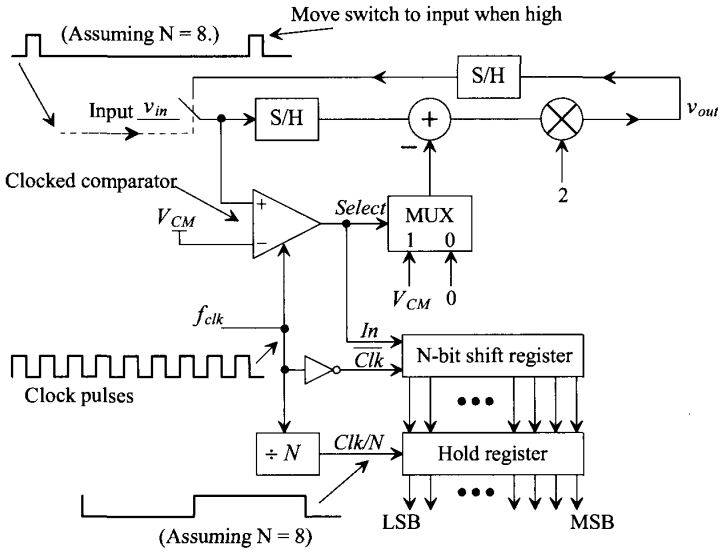


Figure 30.38 Block diagram of a cyclic ADC.

resulting in 0.5 V. This output is multiplied by 2 resulting in 1 V. This 1 V output (the output of the multiplier) is cycled back to the S/H input.

This continues and the final output of the ADC hold register is 1111 1111 (binary offset format). ■

Example 30.10

Repeat Ex. 30.9 if the Cyclic ADC input is 600 mV.

1. Sample the 600 mV input voltage. The comparator output goes high (MSB, b_7, = 1). The output of the multiply by 2, after subtracting V_{CM} (= 500 mV) from the S/H output, is 200 mV.

2. Sample the 200 mV fed back voltage. The comparator output goes low (b_6 = 0). The output of the multiplier is 400 mV.

3. Sample 400 mV. The comparator output goes low (b_5 = 0). The output of the multiplier is 800 mV.

4. Sample 800 mV. The comparator output goes high (b_4 = 1). The output of the multiplier is 600 mV.

5. Sample 600 mV (b_3 = 1) output of the multiplier is 200 mV.

6. Sample 200 mV (b_2 = 0) output of the multiplier is 400 mV.

IPR2023-00698
Theta EX2019

7. Sample 400 mV ($b_1 = 0$) output of the multiplier is 800 mV.

8. Sample 800 mV ($b_0 = 1$) output of the multiplier is 600 mV.

9. Sample the new input voltage and begin conversion again. The output word in the hold register is 1001 1001 (binary offset). ■

Comparator Placement

We showed the inverting input of the comparator in Fig. 30.38 connected to the common-mode voltage. In practice, however, we know that the comparator will have an offset or that the fed back signal may have a common-mode voltage slightly different than the ideal value. If the common-mode voltage of the fed back signals was, for example, 10 mV different than the ideal value, the comparator can make a wrong decision. Further, if the common-mode voltage is varying because of power supply noise or temperature changes then a wrong decision can occur even if some calibration scheme is employed. To avoid a wrong decision, the comparator is most often used in a fully-differential configuration, as seen in Fig. 30.39, with offset storage. The comparator can have significant kickback noise. By adding the ϕ_2 switches in series with the comparator input, we ensure the kickback noise doesn't corrupt the S/H input voltage. Note that since ϕ_3 and ϕ_2 are nonoverlapping, we guarantee that the comparator and S/H are disconnected when the comparator is clocked (and the kickback noise is generated). When the ϕ_1 switches are closed, the offset voltage of the comparator or the op-amp is zeroed out. The performance requirements of the comparator (gain and offset) can be greatly reduced (offset storage is not required) if we use 1.5 bits per clock cycle instead. We discuss this further in the next section.

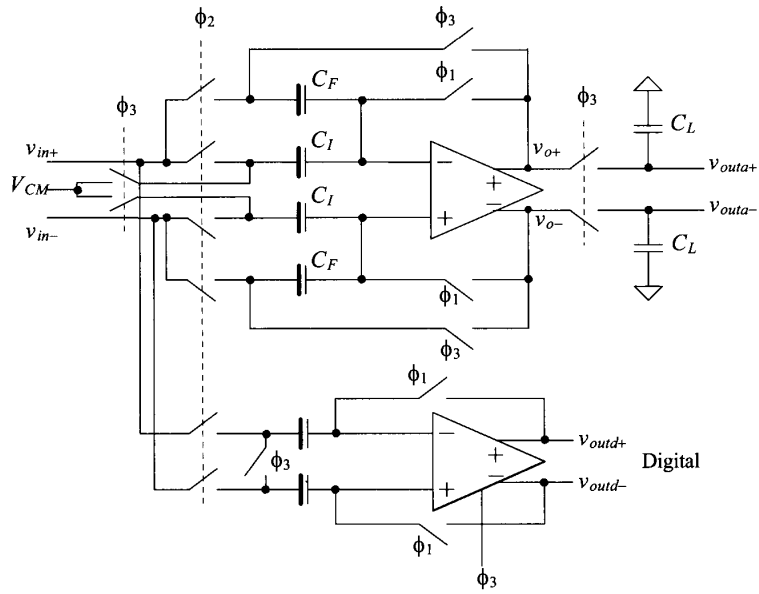


Figure 30.39 Implementation of the comparator with an S/H for use in a cyclic ADC.

IPR2023-00698
Theta EX2019

Example 30.11

Estimate the gain required of a comparator used in a 10-bit cyclic ADC if V_{REF+} = 1 V and $V_{REF-} = 0$.

The LSB of this converter is $1/2^{10} \approx 1\ mV$. This means that the comparator gain must be large enough so that the output can fully transition to either *VDD* or ground with less than 1 mV difference on its inputs (assuming the offset voltage is zeroed out). In other words, the gain must be well above 1/1 mV or 2^{10}. Clearly this presents a real design concern. The gain of the positive feedback comparator may be very large because of the positive feedback used. However, the delay time of the comparator (time delay between the clock going high and the outputs transitioning all the way to *VDD* and ground) may be too long with such a little input voltage difference. Increasing the gain of the comparator, without care, can result in the comparator being unstable when placed in the unity feedback condition (the ϕ_1 switches closed in Fig. 30.39). To increase the comparator gain and avoid instability, a diff-amp (or two) can be added as a pre-amp in front of the positive feedback portion of the comparator. The offset of the diff-amp can then be zeroed out by placing the ϕ_1 switches between the diff-amp's output and its input. ■

Implementing Subtraction in the S/H

Notice in Fig. 30.38 how we can implement the S/H and then multiply by two by simply setting $C_F = C_I$ in Fig. 30.30. Reviewing Fig. 30.38 we see that it would also be useful to implement the subtraction in the S/H. In this figure we see that if the output of the MUX is 0 V, nothing needs to be changed in Fig. 30.30. However, if the MUX output is V_{CM}, then the S/H output must be reduced by V_{CM}. Consider what happens if, when ϕ_3 goes high, instead of connecting the bottom plate of C_I to V_{CM} in Fig. 30.30 we connect it to a voltage V_{CI+} (see Fig. 30.40). Doing this, after reviewing Eqs. (30.46) to (30.49), results in

$$Q_I^{\phi_3} = C_I \cdot (V_{CI+} - V_{CM} \pm V_{OS}) \tag{30.51}$$

or

$$v_{out+} = \left(1 + \frac{C_I}{C_F}\right) \cdot v_{in+} - \frac{C_I}{C_F} \cdot V_{CI+} \tag{30.52}$$

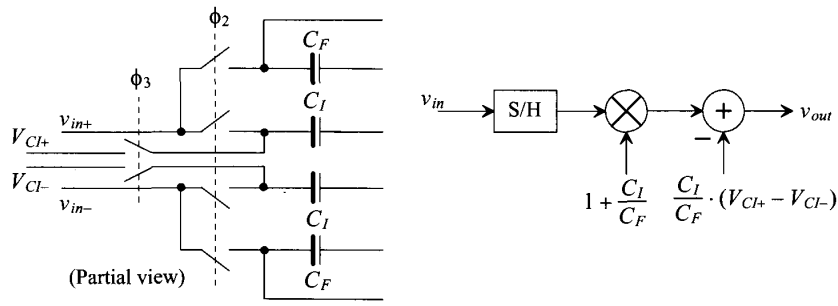


Figure 30.40 Implementing subtraction in the S/H.

IPR2023-00698
Theta EX2019

The differential output voltage is then given by

$$v_{out} = v_{out+} - v_{out-} = \left(1 + \frac{C_I}{C_F}\right) \cdot (v_{in+} - v_{in-}) - \frac{C_I}{C_F} \cdot (V_{CI+} - V_{CI-}) \qquad (30.53)$$

We can easily rearrange the block diagram of Fig. 30.40 so that it more closely resembles the block diagram of the cyclic converter (Fig. 30.38), as seen in Fig. 30.41. If $C_F = C_I$, for our needed gain of two, we end up subtracting V_{CM} when V_{CI+} is $1.5V_{CM}$ and V_{CI-} is $0.5V_{CM}$. For the fully-differential I/O signal case then we are actually *subtracting* $V_{CM}/2$ when $v_{in+} > v_{in-}$ and *adding* $V_{CM}/2$ when $v_{in+} < v_{in-}$.

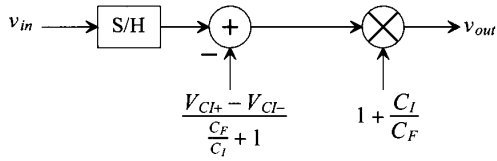


Figure 30.41 Block diagram of Fig. 30.30 with bottom plates of C_I tied to V_{CI}.

Example 30.12

Simulate the operation of the S/H shown in Fig. 30.42 if f_s = 100 MHz, $C_F = C_I$ = 1 pF, $V_{CI+} = 1.5V_{CM}$, and V_{CI-} is $0.5V_{CM}$. Comment on the resulting output.

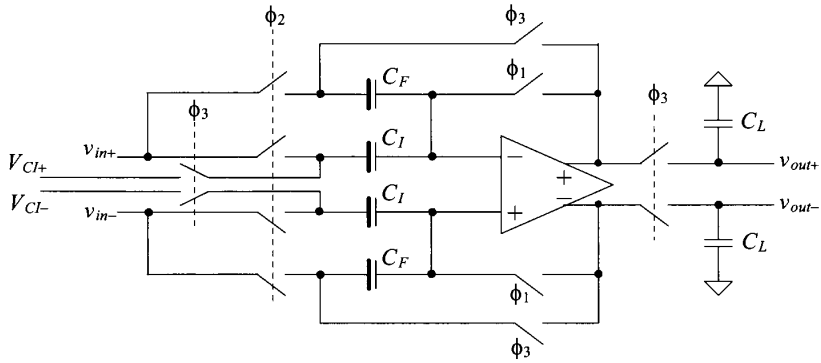


Figure 30.42 S/H used in Ex. 30.12

The simulation results are shown in Fig. 30.43. We only show the situation when we would want to subtract $V_{CM}/2$ from the differential input signal of the cyclic ADC, that is, when $v_{in+} > v_{in-}$. When the inputs are approximately the same voltage, the + input is approximately the same voltage as the – input and v_{in} is 0. We subtract 250 mV ($V_{CM}/2$) from the input and multiply by 2. The result, when the input is 0, is – 500 mV ($-V_{CM}$). When this happens, v_{out+} is 250 mV and v_{out-} is

IPR2023-00698
Theta EX2019

750 mV. Taking the difference in these signals results in – 500 mV. At 100 ns in Fig. 30.43, for example, v_{in} is + V_{CM}. After we subtract $V_{CM}/2$ and multiply by 2, we get V_{CM} again (as indicated in the figure). ■

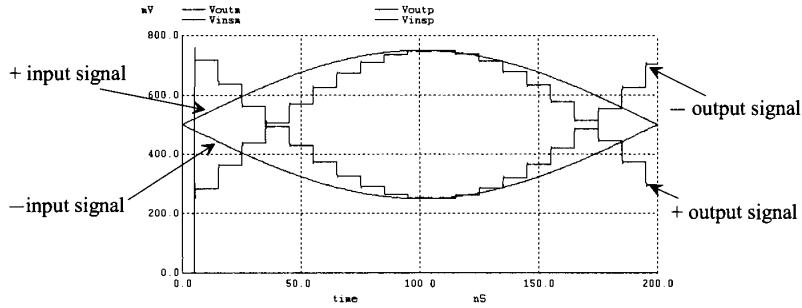


Figure 30.43 Simulation results for Ex. 30.12.

Example 30.13
Repeat Ex. 30.12 if we want to add $V_{CM}/2$ to the input signal.

We only want to add $V_{CM}/2$ to the input signal when $v_{in+} < v_{in-}$. In this situation we set $V_{CI+} = 0.5V_{CM}$ and V_{CI-} to $1.5V_{CM}$. The simulation results are shown in Fig. 30.44. ■

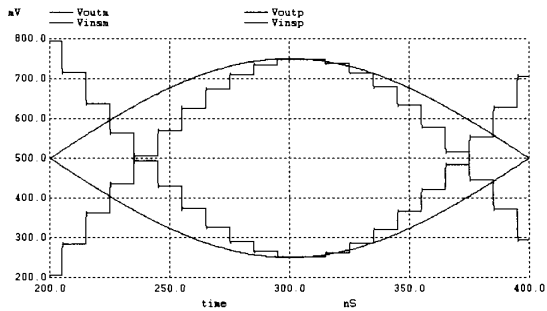


Figure 30.44 Simulation results for Ex. 30.13.

Let's write the analog output of the single-ended cyclic stage as

$$v_{out} = 2 \cdot (v_{in} + \overline{b} \cdot 0 - b \cdot V_{CM}) \tag{30.54}$$

where b is the digital (1 or 0) output of the comparator. Figure 30.45 shows the transfer curve for the cyclic stage. If the comparator output, b, is a 1 ($v_{in} > V_{CM}$), then we subtract V_{CM} from v_{in} before multiplying by two. Note that we have assumed

$$\frac{C_I}{C_F} = 1 \tag{30.55}$$

IPR2023-00698
Theta EX2019

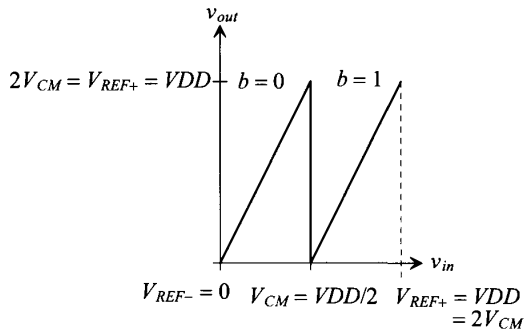


Figure 30.45 Transfer curve for the cyclic ADC (single-ended case).

Understanding Output Swing

After reviewing Fig. 30.45, the reader may wonder how the ADC will perform when the op-amp output must swing all the way up to *VDD* and down to ground. Any practical op-amp output voltage will become nonlinear as it approaches the supply rails. What must be realized is that the single-ended voltage, which ranges from 0 to *VDD*, is changed into a fully-differential voltage using the circuit of Fig. 30.30, which varies from *VDD*/4 to 3*VDD*/4 (as seen in Fig. 30.37). If we use our common-mode voltage as a reference (V_{CM} = *VDD*/2) for single-ended inputs, then we can show some conversions from single-ended to differential.

Single-ended input	Differential outputs
1. $v_{in} = 0.5VDD$	$v_{out+} = v_{out-} = V_{CM}$, $v_{out} = v_{out+} - v_{out-} = 0$
2. $v_{in} = VDD$	$v_{out+} = 0.75VDD, v_{out-} = 0.25VDD$, or $v_{out} = VDD/2$
3. $v_{in} = 0$	$v_{out+} = 0.25VDD, v_{out-} = 0.75VDD$, or $v_{out} = -VDD/2$
4. $v_{in} = 0.6VDD$	$v_{out+} = 0.55VDD, v_{out-} = 0.45VDD$, or $v_{out} = 0.1VDD$

Figure 30.46 shows the transfer curve of Fig. 30.45 redrawn to indicate that the signals, both input and output, are fully-differential. Note, as seen in Fig. 30.39, that the comparator output transitions from a 0 to a 1 when $V_{in+} > V_{in-}$ ($V_{in} > 0$).

We can rewrite Eq. (30.54) for fully-differential signals by looking at Fig. 30.46 and noticing that now, because the inputs and outputs are fully-differential and referenced to V_{CM} instead of 0 so to must be the voltages we add and subtract from the input, Eq. (30.54) can be written as

$$v_{out+} - v_{out-} = 2 \cdot \left(v_{in+} - v_{in-} + \overline{b} \cdot 0 - b \cdot V_{CM} \right) + \overline{b} \cdot V_{CM} + b \cdot V_{CM} \qquad (30.56)$$

or

$$v_{out+} - v_{out-} = 2 \cdot \left(v_{in+} - v_{in-} + \overline{b} \cdot 0 - b \cdot V_{CM} + \overline{b} \cdot \frac{V_{CM}}{2} + b \cdot \frac{V_{CM}}{2} \right) \qquad (30.57)$$

IPR2023-00698
Theta EX2019

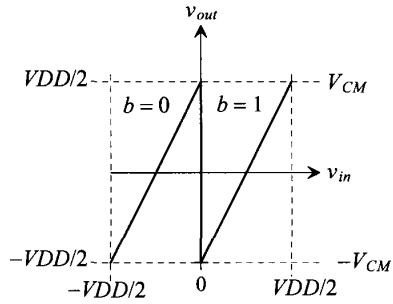


Figure 30.46 Transfer curve for the cyclic ADC when using fully-differential signals.

and finally

$$\overbrace{v_{out+} - v_{out-}}^{v_{out}} = 2 \cdot \left(\overbrace{v_{in+} - v_{in-}}^{v_{in}} + \bar{b} \cdot \frac{V_{CM}}{2} - b \cdot \frac{V_{CM}}{2} \right) \qquad (30.58)$$

Example 30.14

Using the fully-differential S/H stage of Fig. 30.42 simulate the transfer curve shown in Fig. 30.46.

The results of the simulation are shown in Fig. 30.47. To simulate v_{in}, the v_{in+} input is a ramp that varies from 250 mV to 750 mV while, at the same time, the v_{in-} is a ramp that varies from 750 mV to 250 mV. This results in v_{in} changing linearly from $-V_{CM}$ to $+V_{CM}$ (knowing $V_{CM} = VDD/2$). When the two ramps are equal, that is, when they are both V_{CM} ($v_{in} = 0$), V_{CI+} changes from $0.5V_{CM}$ to $1.5V_{CM}$, and V_{CI-} changes from $1.5V_{CM}$ to $0.5V_{CM}$ so that we go from adding $V_{CM}/2$ when $v_{in} < 0$ to subtracting $V_{CM}/2$ when $v_{in} > 0$. ■

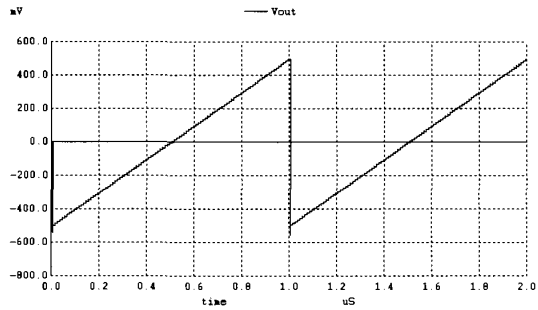


Figure 30.47 Simulating the transfer curves of a cyclic ADC stage.

IPR2023-00698
Theta EX2019

30.3.3 The Pipeline ADC

One drawback of the cyclic ADC discussed in the previous section is the requirement of N clock cycles for each N-bit conversion. From Ch. 29 we know that the flash and pipeline (after an N-bit latency) topologies can perform an analog-to-digital conversion in one clock cycle. Another possibility for an N-bit conversion in one clock cycle is to use a time-interleaved topology. The basic idea is seen in Fig. 30.48. The S/Hs are sequentially clocked so that during each clock cycle the input voltage, v_{in}, is sampled, held, and applied to the input of an N-bit ADC. If the outputs of each ADC are then sequentially available through a MUX, the overall topology behaves as if it were a single ADC with flash-like performance. The practical problem with this topology is the matching between the ADCs and the sensitivity to clock skew and jitter. Differences in the DC characteristics of each ADC, for example, can result in measuring digital output values that change with time when a DC input signal is applied (a ripple on the output similar to what was seen in a noise-shaping data converter output). Mismatches in the ADCs can also result in a reduced (from ideal) signal-to-noise plus distortion ratio (*SNDR*).

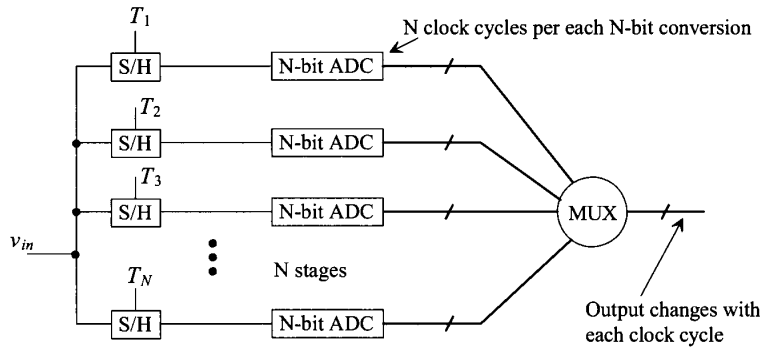


Figure 30.48 Time-interleaved operation of an ADC.

The pipelined ADC (see Sec. 29.2.3) can be thought of as an amplitude-interleaved topology where errors from one stage are correlated with errors from previous stages. The basic block diagram implementation of an N-bit pipelined ADC using the cyclic stage (see Fig. 30.42 for example) is seen in Fig. 30.49. Instead of cycling the analog output of the 1 bit/stage section back to its input, we feed the output into the next stage. The stages are clocked with opposite phases of the master clock signal. The comparator outputs are labeled *digital* in the figure. The digital comparator outputs are delayed through latches so that the final digital output word corresponds to the input signal sampled N clock cycles earlier. The first stage in Fig. 30.49 must be N-bit accurate (again see Sec. 29.2.3). It must amplify its analog output voltage, v_{N-1}, to within 1 LSB of the ideal value (after the subtraction of 0 or V_{CM} from the input signal and the multiplication by two). The second stage output, v_{N-2}, must be an analog voltage within 2 LSBs of its ideal value. The third stage output, v_{N-3}, must be an analog voltage within 4 LSBs of its ideal value and so forth. The important point here is that because the required accuracy of each stage decreases as we move down the line, the settling time, gain

IPR2023-00698
Theta EX2019

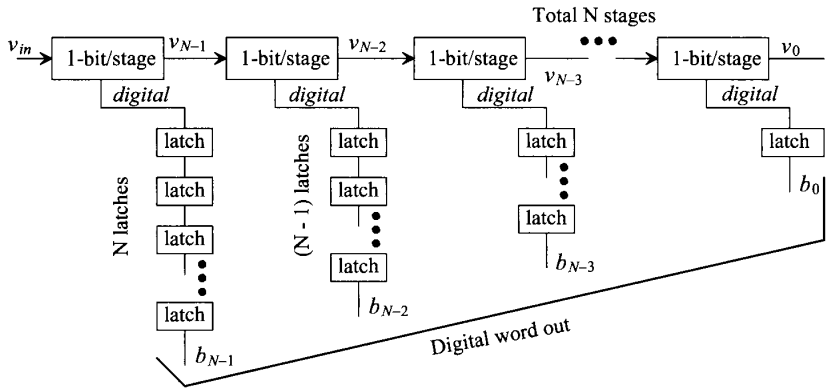


Figure 30.49 Pipelined ADC based on the cyclic stages discussed in the last section.

accuracy, and offsets all become less important. Smaller (and thus lower power) stages can be used for the later stages having, possibly, a dramatic effect on both layout size and power dissipation. While we're showing 1 bit/stage in Fig. 30.49, most commercially available pipeline ADCs use the *digital error correction* present in the 1.5 bit/stage topology discussed later.

We know from our analysis of pipeline ADC errors in Ch. 29 that they can be the result of comparator offsets, gain or linearity errors, and amplifier offsets. In the remainder of this section we discuss how to reduce the effects of these errors. We begin with a discussion of using the 1.5 bit/stage topology to make the comparator offsets (and the reference voltages used with the comparators) a "don't care." While we might think that using six stages with 1.5 bits/stage would result in an ADC with 9-bit resolution, we find that the extra 0.5 bit/stage is used to correct for errors introduced by the comparators. Using six stages will still result in a 6-bit ADC. We then cover the use of capacitor error averaging to set the gains of the amplifiers to precisely two. The cost of this technique is the increase in the number of clock cycles required for each stage's operation and a slightly more complex switching scheme. Finally, we cover some other topologies useful in amplifier offset removal and discuss offsets in general.

Using 1.5 Bits/Stage

As we saw in Ex. 30.11, the gain of the comparator (and the offset) can present a practical limitation to the operation of the ADC at high resolutions. The transfer curve of Fig. 30.45 relates the analog input of the cyclic ADC to its analog output voltage. The important point in this figure, besides the gain and linearity, is where the output of the comparator, *b*, transitions from a 0 to a 1. One-bit corresponds to two levels, i.e., a 0 or a 1. Two bits corresponds to four levels, i.e., 00, 01, 10, and 11. If we were to use three levels then we would get 1.5 bits of resolution. Using a thermometer code or decimal numbers, the three levels would then be

IPR2023-00698
Theta EX2019

Thermometer, *ab*	Decimal
11	3
01	1
00	0

Next consider the transfer curves shown in Fig. 30.50 where three levels are used (1.5 bits). We can rewrite Eq. (30.54) for the 1.5 bit case as

$$v_{out} = 2 \cdot (v_{in} - \bar{a}\bar{b} \cdot 0 - \bar{a}b \cdot V_{CM} - ab \cdot 2V_{CM}) + V_{CM} \qquad (30.59)$$

or

$$v_{out} = 2 \cdot (v_{in} + \bar{a}\bar{b} \cdot \frac{V_{CM}}{2} - \bar{a}b \cdot \frac{V_{CM}}{2} - ab \cdot \frac{3V_{CM}}{2}) \qquad (30.60)$$

or if ab = 00 ($v_{in} < V_{CM}/2$), then $v_{out} = 2 \cdot (v_{in} + V_{CM}/2)$, if ab = 01 ($V_{CM} < v_{in} < 3V_{CM}/2$) then $v_{out} = 2 \cdot (v_{in} - 0.5V_{CM})$, and if ab = 11 then $v_{out} = 2 \cdot (v_{in} - 1.5V_{CM})$. For the fully-differential situation Eq. (30.59) can be rewritten as

$$v_{out+} - v_{out-} = 2 \cdot (v_{in+} - v_{in-} + \bar{a}\bar{b} \cdot V_{CM} - \bar{a}b \cdot 0 - ab \cdot V_{CM}) \qquad (30.61)$$

where all we did was reference, to V_{CM}, the voltages added/subtracted to v_{in} as seen in Eq. (30.58).

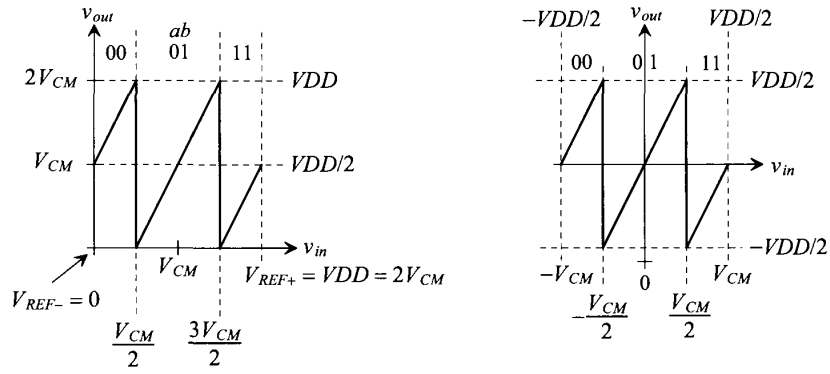


Figure 30.50 Transfer curves for using 1.5-bits per clock cycle.

Figure 30.51a shows how the comparators would be set up to determine *ab* and how we would provide the outputs. The outputs of the comparators are used as address inputs to a MUX. The MUX is used to set the bottom plate voltage of the C_1 capacitor in Fig. 30.42. Figure 30.51b shows how we would implement the comparators if the input and output signals are double-ended.

IPR2023-00698
Theta EX2019

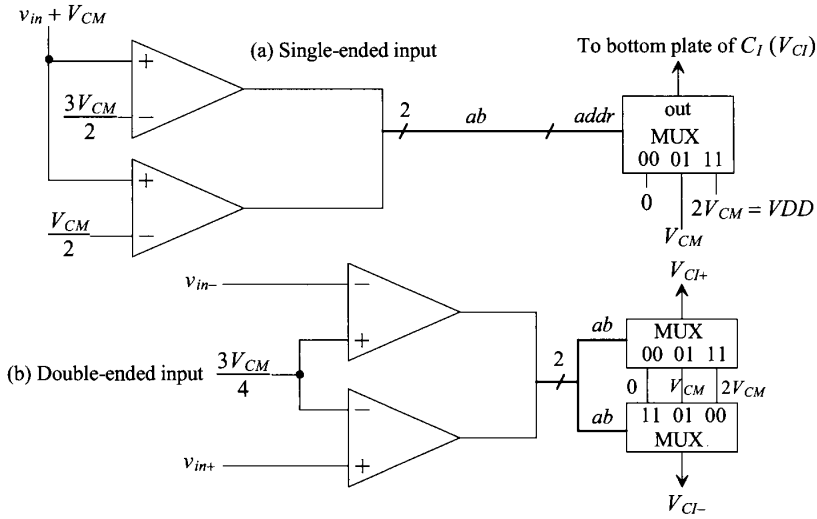
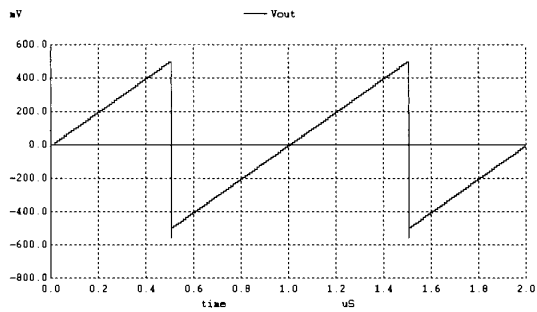


Figure 30.51 Implementing comparators and MUX for 1.5 bits.

Example 30.15
Repeat Ex. 30.14 if 1.5 bits/stage are used.

The simulation results are shown in Fig. 30.52. The same signals were used for the inputs (two ramps) here as were used in Ex. 30.14. The voltage V_{C1+} is 0 V when v_{in} is less than – 250 mV and is 0.5 when v_{in} is between – 250 mV and + 250 mV, while it is 1 V (= $VDD = 2V_{CM}$) when v_{in} is greater than + 250 mV. Notice how we only need three precision voltages, unlike the 1-bit/stage case, that is, VDD, V_{CM}, and 0. ■

Before going any further, let's answer, *"How can using 1.5 bits/stage eliminate the need for precision comparators?"* After reviewing Fig. 30.50, we see that an output of 11 cannot be followed by another output of 11 because we subtract V_{CM} from the input prior to multiplication by two. Even if the comparator has a reasonable offset (say less than 100



$VDD = 1\ V$

Figure 30.52 Simulating the transfer curves for 1.5 bits/stage.

IPR2023-00698
Theta EX2019

mV) or low gain (say less than 50) resulting in a wrong decision, it's impossible for a 11 to be followed by another 11 or for a 00 to be followed by another 00. It is possible, however, to get a continuous string of 01 outputs (a simple example is when V_{CM} is applied to the ADC).

Let's now discuss how the outputs of the 1.5-bit stage are combined into the final ADC output word. Let's assume, again, that $V_{REF+} = VDD = 1$ V, $V_{REF-} = 0$, and $V_{CM} = 500$ mV. We know that the final N-bit output word can be converted back into an output voltage (with the unwanted quantization noise) using a DAC with the following weighting, Eq. (30.27),

$$v_{out} \text{ (if a DAC or } v_{in} \text{ if an ADC)} = b_{N-1} \cdot V_{CM} + b_{N-2} \cdot \frac{V_{CM}}{2} + b_{N-3} \cdot \frac{V_{CM}}{4} + \ldots + b_0 \cdot \overbrace{\frac{V_{CM}}{2^{N-1}}}^{1\ LSB} \tag{30.62}$$

Reviewing Eq. (30.54) note how if, on the first cycle, $b = 1$, we subtract V_{CM} from the input and then multiply the result by two. After a little thought, we should be able to see how we derive Eq. (30.62) from Eq. (30.54) after N clock cycles.

If the first thermometer code output of the 1.5-bit stage is labeled $a_{1.5N-1}b_{1.5N-1}$ and the second output is $a_{1.5N-2}b_{1.5N-2}$, etc., then we can write, with the help of Eq. (30.59), the relation between the ADC input (analog) and the ADC outputs (digital) as

$$\begin{aligned} v_{in} = {} & \overline{a_{1.5N-1}}b_{1.5N-1} \cdot V_{CM} + a_{1.5N-1}b_{1.5N-1} \cdot 2V_{CM} - \frac{V_{CM}}{2} \\ & + \overline{a_{1.5N-2}}b_{1.5N-2} \cdot \frac{V_{CM}}{2} + a_{1.5N-2}b_{1.5N-2} \cdot \frac{2V_{CM}}{2} - \frac{V_{CM}}{4} \\ & + \overline{a_{1.5N-3}}b_{1.5N-3} \cdot \frac{V_{CM}}{4} + a_{1.5N-3}b_{1.5N-3} \cdot \frac{2V_{CM}}{4} - \frac{V_{CM}}{8} + \ldots \end{aligned} \tag{30.63}$$

noting that we can group

$$-\frac{V_{CM}}{2} - \frac{V_{CM}}{4} - \frac{V_{CM}}{8} - \ldots = -V_{CM} \cdot \frac{2^N - 1}{2^N} = -V_{CM} + 0.5\ LSB + \overbrace{\frac{V_{REF-}}{2^N}}^{=0 \text{ here}} \tag{30.64}$$

which is nothing more than a level shift. Clearly we can combine outputs in the following manner to arrive at an equation similar to Eq. (30.62)

$$\begin{aligned} v_{in} = {} & (\overline{a_{1.5N-1}}b_{1.5N-1} + a_{1.5N-2}b_{1.5N-2}) \cdot V_{CM} + (\overline{a_{1.5N-2}}b_{1.5N-2} + a_{1.5N-3}b_{1.5N-3}) \cdot \frac{V_{CM}}{2} \\ & + (\overline{a_{1.5N-3}}b_{1.5N-3} + a_{1.5N-4}b_{1.5N-4}) \cdot \frac{V_{CM}}{4} + \ldots \\ & a_{1.5N-1}b_{1.5N-1} \cdot 2V_{CM} - V_{CM} + 0.5\ LSB \end{aligned} \tag{30.65}$$

Next notice that, when using a thermometer code, the only time $a_{1.5N-X}b_{1.5N-X}$ (the logical AND of ab) can be high is when both are high. The term $a_{1.5N-X}$ cannot be high while $b_{1.5N-X}$ is low (there is no such output code as 10, see Fig. 30.50). This means that we can replace $a_{1.5N-X}b_{1.5N-X}$ with simply $a_{1.5N-X}$. We can rewrite Eq. (30.65) as

IPR2023-00698
Theta EX2019

$$v_{in} = a_{1.5N-1} \cdot 2V_{CM} + (\overline{a_{1.5N-1}} b_{1.5N-1} + a_{1.5N-2}) \cdot V_{CM} + (\overline{a_{1.5N-2}} b_{1.5N-2} + a_{1.5N-3}) \cdot \frac{V_{CM}}{2}$$

$$+ (\overline{a_{1.5N-3}} b_{1.5N-3} + a_{1.5N-4}) \cdot \frac{V_{CM}}{4} + \ldots + \overline{a_{1.50}} b_{1.50} \cdot \frac{V_{CM}}{2^{N-1}} - (V_{CM} - 0.5\ LSB) \qquad (30.66)$$

The + symbol in Eq. (30.66) indicates addition rather than a logical OR. If $\overline{a_{1.5N-1}} b_{1.5N-1} = 1$ and $a_{1.5N-2} = 1$, then the second term in this equation is $2V_{CM}$ and the first term must be 0. When this occurs, the addition of the two terms is 0 and a carry is generated. We can now use this information to write the relationship between Eq. (30.62) and Eq. (30.63) knowing $\oplus$ indicates an exclusive OR

$$b_0 = \overline{a_{1.50}} b_{1.50} \text{ with carry} = c_0 = a_{1.50} \qquad (30.67)$$

$$b_1 = \overline{a_{1.51}} b_{1.51} \oplus c_0 \text{ with } c_1 = \overline{a_{1.51}} b_{1.51} c_0 \qquad (30.68)$$

$$b_2 = \overline{a_{1.52}} b_{1.52} \oplus a_{1.51} \oplus c_1 \text{ with } c_2 = \overline{a_{1.52}} b_{1.52} a_{1.51} + c_1(\overline{a_{1.52}} b_{1.52} + a_{1.51}) \qquad (30.69)$$

noting each bit and carry are the outputs of a full adder. To simplify the carry equation, and the subsequent equations, we can substitute c_1 to get

$$c_2 = \overline{a_{1.52}} b_{1.52} a_{1.51} + c_1 \overline{a_{1.52}} b_{1.52} \qquad (30.70)$$

We can write the general form of b_2 through b_{N-1}, using full adders, as

$b_{N-1} = \overline{a_{1.5N-1}} b_{1.5N-1} \oplus a_{1.5N-2} \oplus c_{N-2}$ and

$$c_{N-1} = \overline{a_{1.5N-1}} b_{1.5N-1} a_{1.5N-2} + c_{N-2}(\overline{a_{1.5N-1}} b_{1.5N-1} + a_{1.5N-2}) \qquad (30.71)$$

The bit b_N has a weighting of $2V_{CM}$ and thus the final output word size is $N + 1$

$$b_N = a_{1.5N-1} \oplus c_{N-1} \qquad (30.72)$$

Before we sketch the implementation of this digital circuit, let's make a few comments. To begin, notice that the word size is one larger than N (the resolution). In the 1 bit/stage circuit our maximum output is (assuming $N = 8$)

$$1111\ 1111 = VDD - 1\ \text{LSB} = 2V_{CM} - 1\ \text{LSB}\ \text{(1 bit/stage)}$$

Now our maximum output is

$$1\ 0000\ 0000 = VDD = 2V_{CM}\ \text{(1.5 bit/stage)}$$

The resolution is still essentially eight bits; we just have a slightly larger (1 LSB) output range. To make the words exactly match, we can throw out the MSB and assume 1 0000 0000 is an out-of-range condition. Note that an input of VDD – 1 LSB using 1.5 bits/stage gives an output code of 0 1111 1111.

One more comment: if we go through the logic in Eq. (30.66), we don't get the correct outputs unless we subtract V_{CM} – 0.5 LSB. The binary offset representation can be written as

$$V_{CM} - 0.5\ \text{LSB} = 0011111111... \qquad (30.73)$$

For example, applying V_{CM} (single-ended, see Fig. 30.51) to the ADC input results in a continuous output (ab) of 01. The output prior to subtraction, from Eqs. (30.67) to (30.72), is then 01111111... (b_N = 0, b_{N-1} = 1, b_{N-2} = 1, etc.). After subtracting 0011111111..., we get 01000000 or V_{CM} knowing the weighting of the second bit is V_{CM}.

IPR2023-00698
Theta EX2019

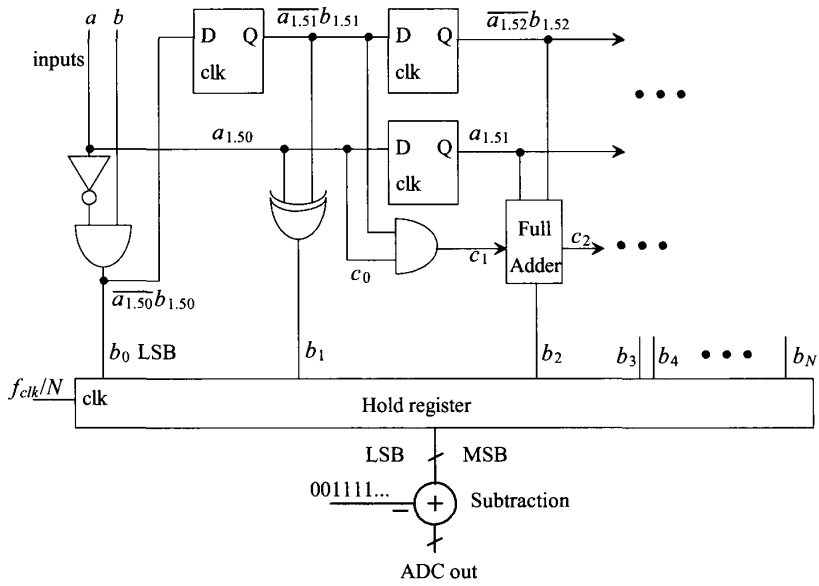


Figure 30.53 Combining the outputs of the cyclic ADC when 1.5 bits/stage is used.

Figure 30.53 shows one possible implementation of Eqs. (30.67)-(30.72) for a cyclic ADC. The state shown, i.e., the *ab* values, is valid at the end of the conversion (after N clock cycles). When starting the algorithm (on the first rising edge of clock), all latches are reset to zeroes and the first comparator outputs, *ab*, are applied. On the second rising edge of clock the output b_2 corresponds to the final b_N. After N clock cycles, the hold register is clocked (and the latches are reset). The final ADC output is the contents of the hold register after subtracting 00111111.... Changing the numbers to two's complement may be useful when implementing this stage (assuming the MSB has a weighting of V_{CM}, i.e., throw out b_N). Note the subtraction can precede the hold register.

Example 30.16

Repeat Ex. 30.10 if 1.5 bits/stage are used. Assume the converter is ideal and the comparators switch precisely at $V_{CM}/2$ (= 250 mV here) and $3V_{CM}/2$ (= 750 mV here). Assume all latches initially contain zeroes.

v_{in}	$a_{1.5X}b_{1.5X}$	v_{out}	Digital out
600 mV ($N-1=7$)	01	700 mV	$b_7 = \overline{a_{1.57}}b_{1.57} \oplus a_{1.56} \oplus c_6 = 0$
$c_7 = \overline{a_{1.57}}b_{1.57}a_{1.56} + c_6(\overline{a_{1.57}}b_{1.57} + a_{1.56}) = 1$			
700 mV ($N-2=6$)	01	900 mV	$b_6 = \overline{a_{1.56}}b_{1.56} \oplus a_{1.55} \oplus c_5 = 0$
$c_6 = \overline{a_{1.56}}b_{1.56}a_{1.55} + c_5(\overline{a_{1.56}}b_{1.56} + a_{1.55}) = 1$			

IPR2023-00698
Theta EX2019

900 mV ($N-3=5$) 11 300 mV $b_5 = \overline{a_{1.55}} b_{1.55} \oplus a_{1.54} \oplus c_4 = 0$

$c_5 = \overline{a_{1.55}} b_{1.55} a_{1.54} + c_4(\overline{a_{1.55}} b_{1.55} + a_{1.54}) = 0$

300 mV ($N-4=4$) 01 100 mV $b_4 = \overline{a_{1.54}} b_{1.54} \oplus a_{1.53} \oplus c_3 = 1$

$c_4 = \overline{a_{1.54}} b_{1.54} a_{1.53} + c_3(\overline{a_{1.54}} b_{1.54} + a_{1.53}) = 0$

100 mV ($N-4=3$) 00 700 mV $b_3 = \overline{a_{1.53}} b_{1.53} \oplus a_{1.52} \oplus c_2 = 1$

$c_3 = \overline{a_{1.53}} b_{1.53} a_{1.52} + c_2(\overline{a_{1.53}} b_{1.53} + a_{1.52}) = 0$

700 mV ($N-6=2$) 01 900 mV $b_2 = \overline{a_{1.52}} b_{1.52} \oplus a_{1.51} \oplus c_1 = 0$

$c_2 = \overline{a_{1.52}} b_{1.52} a_{1.51} + c_1(\overline{a_{1.52}} b_{1.52} + a_{1.51}) = 1$

900 mV ($N-7=1$) 11 300 mV $b_1 = \overline{a_{1.51}} b_{1.51} \oplus c_0 = 0$

$c_1 = \overline{a_{1.51}} b_{1.51} a_{1.50} + c_0(\overline{a_{1.51}} b_{1.51} + a_{1.50}) = 0$

300 mV ($N-8=0$) 01 100 mV $b_0 = \overline{a_{1.50}} b_{1.50} = 1$

$c_0 = a_{1.50} = 0$

noting that $b_8 = a_{1.57} \oplus c_7 = 1$. We can reorder the bits so the MSB is on the left, the LSB is on the right, yielding 1 0001 1001 and subtract 0 0111 1111 yielding

$$\begin{array}{r} 1\ 0001\ 1001\ (281) \\ -\ \underline{0\ 0111\ 1111}\ (127) \\ 0\ 1001\ 1010\ (154) \end{array}$$

This is the result given in Ex. 30.10 (1001 1001, or decimal 153) plus 1 LSB. The 1 LSB discrepancy can be traced to Eq. (30.66) where we used 0.5 LSBs. Because our resolution is at best 1 LSB, sometimes the result will experience a round-off error. To understand this in the subtraction above, the more correct decimal representation of V_{CM} – 0.5 LSBs is 127.5 and the more correct decimal output is 153.5. ■

Example 30.17

Repeat Ex. 30.16 if the comparators switch at 305 mV (a 55 mV offset) and 675 mV (a –75 mV offset).

v_{in}	$a_{1.5X} b_{1.5X}$	v_{out}	Digital out
600 mV ($N-1=7$)	01	700 mV	$b_7 = \overline{a_{1.57}} b_{1.57} \oplus a_{1.56} \oplus c_6 = 0$
$c_7 = \overline{a_{1.57}} b_{1.57} a_{1.56} + c_6(\overline{a_{1.57}} b_{1.57} + a_{1.56}) = 1$			
700 mV ($N-2=6$)	11	–100 mV	$b_6 = \overline{a_{1.56}} b_{1.56} \oplus a_{1.55} \oplus c_5 = 0$
$c_6 = \overline{a_{1.56}} b_{1.56} a_{1.55} + c_5(\overline{a_{1.56}} b_{1.56} + a_{1.55}) = 0$			
–100 mV ($N-3=5$)	00	300 mV	$b_5 = \overline{a_{1.55}} b_{1.55} \oplus a_{1.54} \oplus c_4 = 0$
$c_5 = \overline{a_{1.55}} b_{1.55} a_{1.54} + c_4(\overline{a_{1.55}} b_{1.55} + a_{1.54}) = 0$			
300 mV ($N-4=4$)	00	1.1 V	$b_4 = \overline{a_{1.54}} b_{1.54} \oplus a_{1.53} \oplus c_3 = 1$
$c_4 = \overline{a_{1.54}} b_{1.54} a_{1.53} + c_3(\overline{a_{1.54}} b_{1.54} + a_{1.53}) = 0$			

IPR2023-00698
Theta EX2019

1100 mV ($N-4=3$) 11 700 mV $b_3 = \overline{a_{1.53}} b_{1.53} \oplus a_{1.52} \oplus c_2 = 1$

$c_3 = \overline{a_{1.53}} b_{1.53} a_{1.52} + c_2(\overline{a_{1.53}} b_{1.53} + a_{1.52}) = 0$

700 mV ($N-6=2$) 11 −100 mV $b_2 = \overline{a_{1.52}} b_{1.52} \oplus a_{1.51} \oplus c_1 = 0$

$c_2 = \overline{a_{1.52}} b_{1.52} a_{1.51} + c_1(\overline{a_{1.52}} b_{1.52} + a_{1.51}) = 0$

−100 mV ($N-7=1$) 00 300 mV $b_1 = \overline{a_{1.51}} b_{1.51} \oplus c_0 = 0$

$c_1 = \overline{a_{1.51}} b_{1.51} a_{1.50} + c_0(\overline{a_{1.51}} b_{1.51} + a_{1.50}) = 0$

300 mV ($N-8=0$) 00 1.1V $b_0 = \overline{a_{1.50}} b_{1.50} = 0$

$c_0 = a_{1.50} = 0$

and $b_8 = a_{1.57} \oplus c_7 = 1$. Again, we can reorder the bits so the MSB is on the left, the LSB is on the right, yielding 1 0001 1000 and subtract 0 0111 1111 yielding

```
  1 0001 1000 (280)
- 0 0111 1111 (127)
  0 1001 1001 (153)
```

This is the exact result given in Ex. 30.10 (1001 1001, or decimal 153). In this case the 0.5 LSB round-off worked in our favor.

While the comparator performance can be extremely poor, the circuit of Fig. 30.42 must subtract and amplify to an accuracy set by the least significant bit of the converter. When we calculated values in Ex. 30.16 and here in this example we assumed subtractions of exactly 0, V_{CM}, and $2V_{CM}$ followed by a multiplication of exactly two.

Finally, notice that the negative output of −100 mV (single-ended) and the 1.1 V output that is greater than *VDD* (= 1 V here) are easily accommodated when using fully-differential op-amps. ■

Capacitor Error Averaging

While using 1.5 bits/stage has made the comparator offset unimportant, we still have to amplify the signal by a precise factor of 2. Toward the goal of precision gain consider the redrawn version of our basic S/H, Fig. 30.42, shown in Fig. 30.54. In this figure we've shown the clock phases (but not the slightly delayed clock signals also used) and the capacitors with mismatch. Ideally, $\Delta C_{+,-} = 0$ and the gain of the S/H is precisely 2 (assuming sufficiently high op-amp open-loop gain, see Eq. [30.38]). From Eq. (30.52) we can write

$$v_{out+} = \left(2 + \frac{\Delta C_+}{C_+}\right) \cdot v_{in+} - \left(1 + \frac{\Delta C_+}{C_+}\right) V_{CI+} \qquad (30.74)$$

where $\Delta C/C$ is the capacitor mismatch (say, ±1% or ±0.01, noting ΔC may be positive or negative). For the negative output

$$v_{out-} = \left(2 + \frac{\Delta C_-}{C_-}\right) \cdot v_{in-} - \left(1 + \frac{\Delta C_-}{C_-}\right) V_{CI-} \qquad (30.75)$$

where $v_{out} = v_{out+} - v_{out-}$. Note that when using this topology in a pipeline converter one stage can be in the hold state while the next stage can be in the sampling state effectively

IPR2023-00698
Theta EX2019

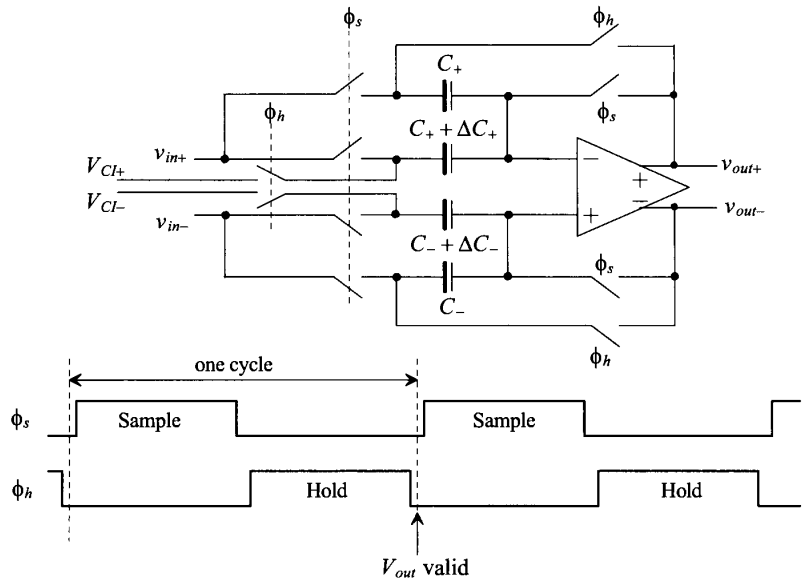


Figure 30.54 S/H of Fig. 30.42 with mismatched capacitors.

sharing the time. The result of this sharing is the need for only one clock cycle for each stage in the conversion.

Next consider what happens if, instead of sampling the input voltage again, we simply switch the positions of the C and $C+\Delta C$ capacitors, that is, we connect C to V_{CI} and $C+\Delta C$ to V_{out}. Figure 30.55 shows this switch and the modified S/H using an extra half-clock cycle. The sample and amplify phases of the clock are exactly the same as before, and so Eqs. (30.74) and (30.75) are still valid. We denote the outputs at the end (falling edge) of the amplify phase as v_{outa+} and v_{outa-}

$$v_{outa+} = 2 \cdot v_{in+} - V_{CI+} + \frac{\Delta C_+}{C_+} \cdot (v_{in+} - V_{CI+}) \qquad (30.76)$$

and

$$v_{outa-} = 2 \cdot v_{in-} - V_{CI-} + \frac{\Delta C_-}{C_-} \cdot (v_{in-} - V_{CI-}) \qquad (30.77)$$

where the ideal situation is $\Delta C = 0$.

At the end of the amplify phase the charge on the capacitors is (assuming the op-amp input voltages are 0 and only looking at the + path to simplify the equations)

$$Q_{a+} = (C_+ + \Delta C_+) \cdot (V_{CI+}) + C_+ \cdot (v_{outa+}) \qquad (30.78)$$

and at the end of the hold phase

$$Q_{h+} = (C_+ + \Delta C_+) \cdot (v_{outh+}) + C_+ \cdot (V_{CI+}) \qquad (30.79)$$

IPR2023-00698
Theta EX2019

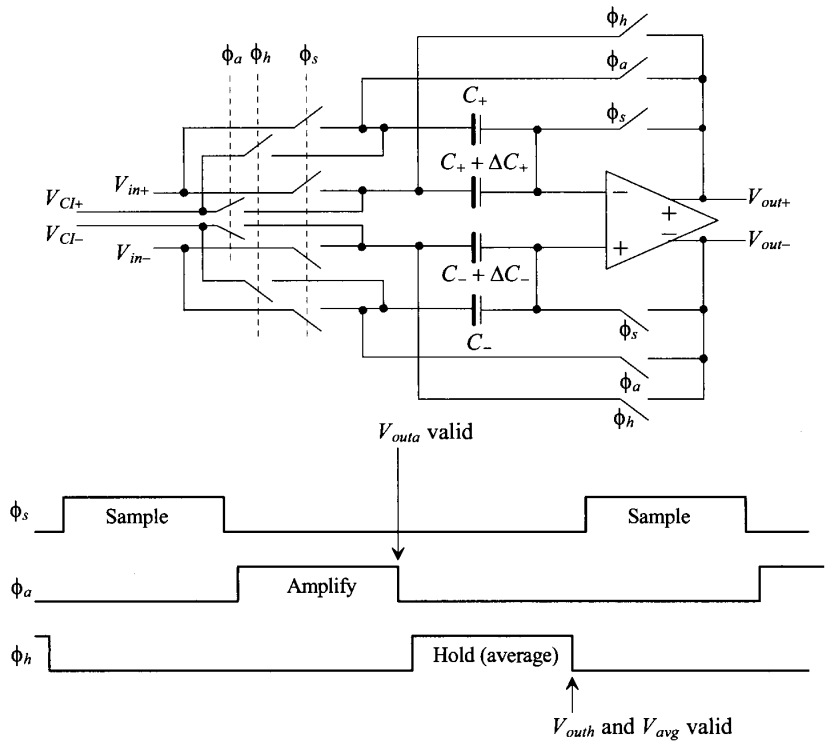


Figure 30.55 S/H using capacitor error averaging.

Because charge must be conserved, $Q_{a+} = Q_{h+}$, and therefore

$$v_{outh+} = V_{CI+} + \frac{C_+}{C_+ + \Delta C_+} \cdot v_{outa+} - \frac{C_+}{C_+ + \Delta C_+} \cdot V_{CI+} \tag{30.80}$$

It will be useful to use

$$\frac{C_+}{C_+ + \Delta C_+} = \frac{1}{1 + \Delta C_+ / C_+} \approx 1 - \frac{\Delta C_+}{C_+} \tag{30.81}$$

Rewriting Eq. (30.80) gives

$$v_{outh+} = \left(1 - \frac{\Delta C_+}{C_+}\right) \cdot v_{outa+} + \frac{\Delta C_+}{C_+} \cdot V_{CI+} \tag{30.82}$$

Substituting Eq. (30.76) for v_{outa+} results in

$$v_{outh+} = 2v_{in+} - V_{CI+} - \frac{\Delta C_+}{C_+} \cdot (v_{in+} - V_{CI+}) \tag{30.83}$$

where the terms containing $(\Delta C_+/C_+)^2$ are assumed negligible. If the matching is 1%, then $(\Delta C_+/C_+)^2 = 10^{-4}$.

IPR2023-00698
Theta EX2019

Clearly averaging v_{outa+} (Eq. [30.76]) and v_{outh+} (Eq. [30.83]) results in the precise desired gain. The question now becomes: "How do we perform the averaging without introducing more error?" Ideally we want

$$\begin{aligned} v_{avg+} &= \frac{v_{outa+} + v_{outh+}}{2} \\ &= \frac{1}{2} \cdot \left(2v_{in+} - V_{CI+} + \frac{\Delta C_+}{C_+} \cdot (v_{in+} - V_{CI+}) + 2v_{in+} - V_{CI+} - \frac{\Delta C_+}{C_+} \cdot (v_{in+} - V_{CI+}) \right) \\ &= 2v_{in+} - V_{CI+} \end{aligned} \tag{30.84}$$

or more generally

$$v_{avg} = v_{avg+} - v_{avg-} = 2(v_{in+} - v_{in-}) - (V_{CI+} - V_{CI-}) \tag{30.85}$$

This equation should be compared to Eq. (30.61).

Consider the averaging amplifier shown in Fig. 30.56. We can write the charge on the four capacitors when the ϕ_a switches are closed (actually at the falling edge of the ϕ_a clock, assuming complete settling) as

$$Q_+^{\phi_a} = (v_{outa+} - V_{CM} \pm V_{OS}) \cdot C_F + (v_{outa-} - V_{CM} \pm V_{OS}) \cdot C_I \tag{30.86}$$

and

$$Q_-^{\phi_a} = (v_{outa-} - V_{CM} \pm V_{OS}) \cdot C_F + (v_{outa+} - V_{CM} \pm V_{OS}) \cdot C_I \tag{30.87}$$

Note that the common-mode voltage and op-amp offset subtract out when we take the difference between the balanced signals and so we do not include V_{CM} or V_{OS} in the remaining analysis. (The offset from the common-mode feedback circuit results in an unharmful variation in V_{CM} and is discussed later in this section.) On the falling edge of ϕ_h, and following the same procedure as used in Eq. (30.48), we can write

$$C_F \cdot v_{avg+} = C_F \cdot v_{outa+} + C_I \cdot v_{outa-} - C_I \cdot v_{outh-} \tag{30.88}$$

and

$$C_F \cdot v_{avg-} = C_F \cdot v_{outa-} + C_I \cdot v_{outa+} - C_I \cdot v_{outh+} \tag{30.89}$$

We can then write

$$v_{avg+} = v_{outa+} + (v_{outa-} - v_{outh-}) \cdot \frac{C_I}{C_F} \tag{30.90}$$

$$v_{avg-} = v_{outa-} + (v_{outa+} - v_{outh+}) \cdot \frac{C_I}{C_F} \tag{30.91}$$

noting that if $v_{outa-} = v_{outh-}$ and $v_{outa+} = v_{outh+}$, the matching of the capacitors in Fig. 30.55 is perfect and $v_{avg+} - v_{avg-} = v_{outa+} - v_{outa-}$ (the output perfectly follows Eq. [30.61]). Taking the difference of these two equations results in

$$v_{avg+} - v_{avg-} = v_{outa+} - v_{outa-} - \overbrace{\frac{(v_{outa+} - v_{outh+}) - (v_{outa-} - v_{outh-})}{2}}^{\text{Error adjustment term}} \tag{30.92}$$

If we substitute Eqs. (30.76), (30.77), and (30.83) into this equation, we get

IPR2023-00698
Theta EX2019

$$v_{avg+} - v_{avg-} = 2(v_{in+} - v_{in-}) - (V_{CI+} - V_{CI-}) + (v_{in+} - V_{CI+}) \cdot \frac{\Delta C_+}{C_+} - (v_{in-} - V_{CI-}) \cdot \frac{\Delta C_-}{C_-} - \frac{2(v_{in+} - V_{CI+})}{C_F/C_I} \cdot \frac{\Delta C_+}{C_+} + \frac{2(v_{in-} - V_{CI-})}{C_F/C_I} \cdot \frac{\Delta C_-}{C_-} \quad (30.93)$$

which reduces to Eq. (30.85) or (30.61) assuming C_I/C_F is 1/2 (C_F is twice as large as C_I). Note how the selection (and matching) of the capacitor ratios, C_F/C_I, is not that important. The capacitors do not have to match to the final ADC resolution. A matching of 1% will result in an error term that is one-hundreth of the error that would be present if the error averaging were not used. Also note, as indicated at the beginning of the section, that only the first stages (say the first five in a 14-bit ADC) need use capacitor error averaging because of the reduced accuracy requirements placed on the later stages as we move through the converter.

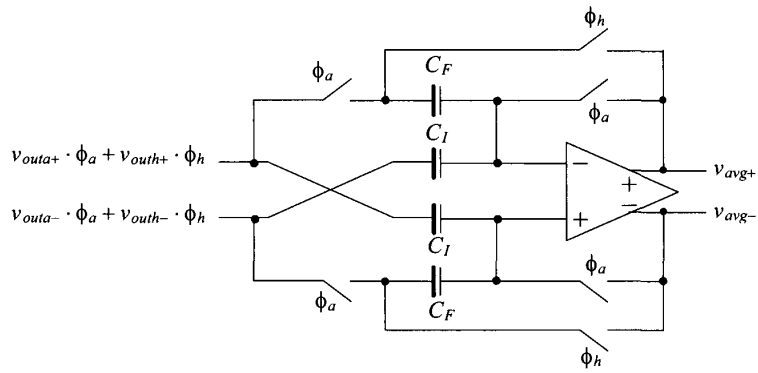


Figure 30.56 Averaging amplifier for use on the output of Fig. 30.55.

The penalty for this precision technique is the increase in conversion time used in each stage (and increased noise because two op-amps are used in each 1.5-bit section). As seen in Fig. 30.55, an extra half-clock cycle is required. For 20 Msamples/s a 30 MHz clock would be required. Again while one stage is in the hold (average) state, the next stage in the pipeline can be in the sample state so that the conversion time is shared (but still requiring 1.5 clock cycles). Using the capacitor error averaging technique for precision gain (and subtraction) and using 1.5-bit/stage sections, a 14-bit ADC, using fully-differential input signals, has been attained at 20 Msamples/s without trimming or calibration [5].

Example 30.18

Simulate the operation of the circuit shown in Fig. 30.57. Comment on the ideal outputs and the simulation results.

The capacitor values were chosen arbitrarily. The input voltage, v_{in}, is 0.75 – 0.25 or 0.5 V. The ideal output voltage, v_{out}, is then 1 V. v_{out+} should ideally be 1 V, and v_{out-} should ideally be 0 V. The simulation results are shown in Fig. 30.58. The

IPR2023-00698
Theta EX2019

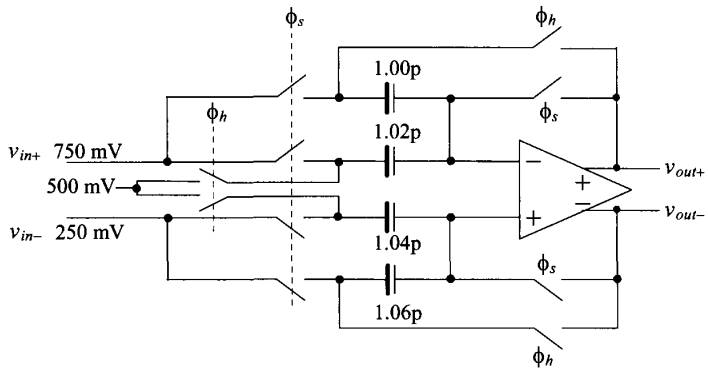


Figure 30.57 Matching errors in capacitors (see Ex. 30.18).

error plotted in this figure is the result of taking the difference between the ideal output and the actual output. The error for the capacitor values shown in Fig. 30.57 is approximately 5 mV. Clearly, at the risk of stating the obvious, capacitor mismatch is a fundamental limitation to ADC accuracy. ■

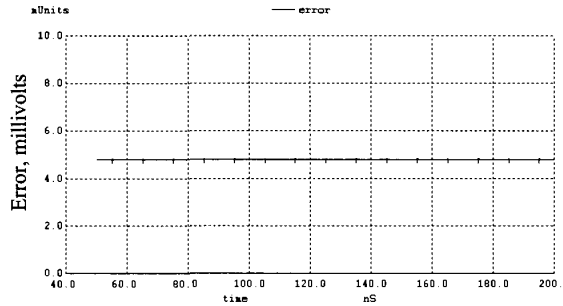


Figure 30.58 Simulation results from Ex. 30.18 showing amplification error.

Example 30.19

Simulate the operation of the circuit shown in Fig. 30.59 (a cascade of Figs. 30.55 and 30.56). Comment on the ideal outputs and the simulation results.

This circuit shows the same mismatched capacitors in the multiply-by-two stage as we saw in Ex. 30.18. The averaging stage also shows mismatched capacitors with arbitrary values. Again, as in Ex. 30.18, the ideal output voltage is 1.0 V (v_{out+} = 1 V and v_{out-} = 0 V). The simulation results are shown in Fig. 30.60. The error has dropped from 5 mV to – 67 μV. It may be instructive to resimulate the capacitor error averaging topology of Fig. 30.59 using different values of capacitors in order to get a feeling for just how forgiving the topology is to mismatches.

IPR2023-00698
Theta EX2019

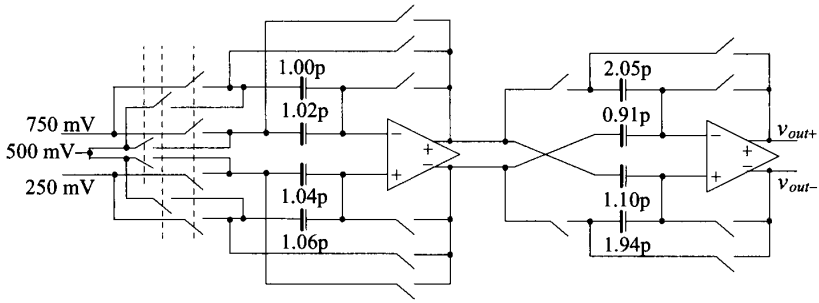


Figure 30.59 Implementation of error averaging (see Ex. 30.19).

Note that in the simulation netlist, where we are using near ideal op-amps with open loop gains of 100 million, we added switches in series with the C_1 capacitors in the averaging circuit to avoid the situation of the op-amp operating open-loop with its outputs going to millions of volts when both ϕ_a and ϕ_h are low. (The op-amp operates open loop when ϕ_s is high, which is usually not a problem in a practical circuit). ■

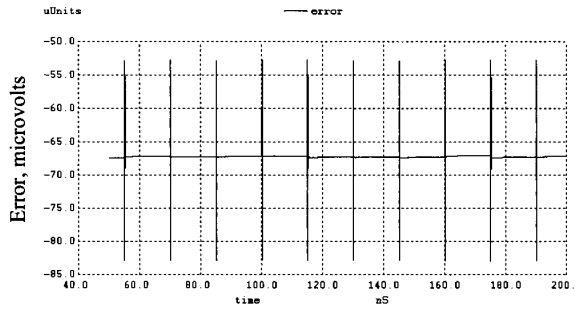


Figure 30.60 Simulation results from Ex. 30.19 showing amplification error.

Example 30.20

Repeat Ex. 30.19 if the op-amps used have open-loop gains of 1,000.

The simulation results are shown in Fig. 30.61. The error has increased by a factor greater than 10. As indicated earlier in Eq. (30.39), the minimum op-amp open-loop gain is set by the resolution of the data converter. If we were designing a 14-bit ADC, the op-amp used would require open-loop gains greater than 2^{16} or 64k (96 dB). ■

IPR2023-00698
Theta EX2019

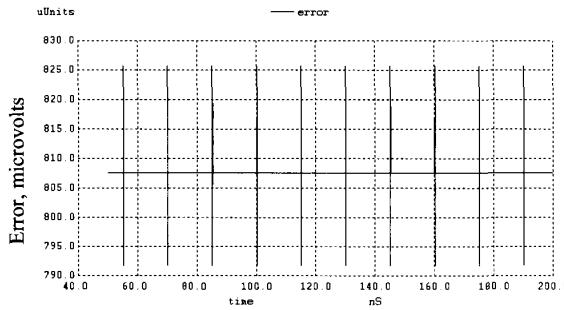


Figure 30.61 Regeneration of Fig. 30.60 using op-amps with open-loop gains of 1,000.

Comparator Placement

The implementation of a pipeline ADC showing how the clock signals are used in each error averaging stage is shown in Fig. 30.62. The important thing to note is the use of only three phases of an input clock signal. The connection of the clock phases to the switches changes in each stage allowing a stage to sample an input signal while the previous stage is in the hold mode. We'll discuss the generation of these clock signals in the next section; here we discuss comparator placement and performance.

As seen in Fig. 30.55 the voltages V_{CI+} and V_{CI-} must be valid and stable when the amplify-and-hold clock phases are high. Reviewing Fig. 30.51 we can implement the two comparators using a clock signal followed by a latch, Fig. 30.63. The comparators can be clocked on the rising edge of the amplifying clock (which is a different clock phase in each of the three possible clocking schemes). It's important to place a separate clocked latch on the outputs of each comparator (*clocked with a slightly delayed clock signal which isn't shown in the figure*) to ensure that comparator metastability isn't a factor when generating the MUX addressing (select inputs). We can get away with this type of clocking scheme because we are using three levels/stage (1.5 bits/stage). The comparators don't have to be N-bit accurate, as discussed earlier, but can withstand offsets/errors approaching $V_{CM}/4$. The signals V_{CI+} and V_{CI-} do have to be N-bit accurate, though, and must settle and stay settled to this accuracy by the end of the amplify phase.

The two bits coming out of the comparators (actually the latches connected to the comparator outputs) can be thought of as the first delay element shown in Fig. 30.49 (except now it is a 2-bit delay element). Because each of these first delays are clocked on the rising edge of one phase of the clock signal, we have a problem with synchronizing the bits together prior to application to the digital correction logic (similar to Fig. 30.53). Reviewing the clock signals in Fig. 30.62, we see that if we clock all other delay elements in the pipeline of Fig. 30.49 on the rising edge of "amplify, one" we can synchronize all of the digital data together.

Clock Generation

Generating the nonoverlapping clock signals used in, for example, Fig. 30.55 can be accomplished by using the basic circuit shown in Fig. 30.64. To understand the operation of this circuit, note that when one of the input signals goes low the corresponding output

IPR2023-00698
Theta EX2019

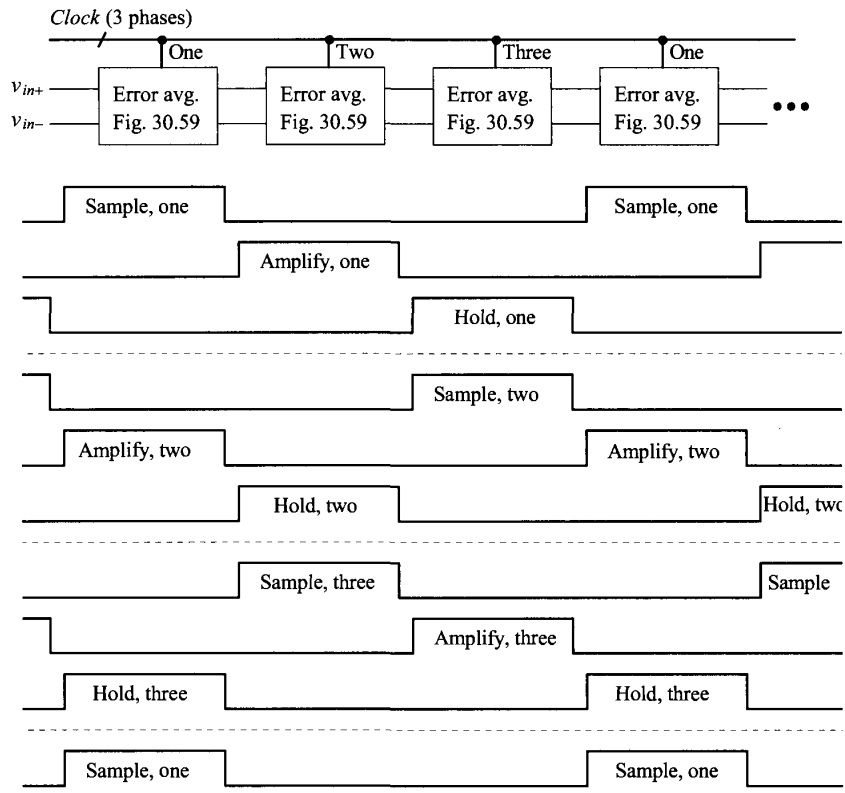


Figure 30.62 Implementation of a pipeline ADC using error averaging (Fig. 30.59).

phase goes high. The feedback ensures that the output doesn't go high until the previous phase goes back low (both inputs of the NOR gate must be low before its output goes high). The amount of nonoverlap time, again, is set by the delay in series with the output of the NOR gates.

The input signals (overlapping clock signals) are generated with the circuit shown in Fig. 30.65. The outputs of this circuit change states on the rising edge of the input clock signal, *Clkin*. After a start-up transient time of a couple of clock cycles, the feedback through the NAND gate to the top D-Flip-Flop ensures that only one bit of In1, In2, or In3 is low any given time. Again, the output of this circuit is then used to drive the nonoverlapping clock circuit seen in Fig. 30.64. While we show 3 signals here it's straightforward to expand the number of clock signals. Note that this circuit is essentially a 3-bit shift register. Also note that changing the NAND gate to a NOR gate or simply inverting the In1 - In3 outputs results in a signal that can be used to drive the nonoverlapping clock generator seen back in Fig. 14.19.

IPR2023-00698
Theta EX2019

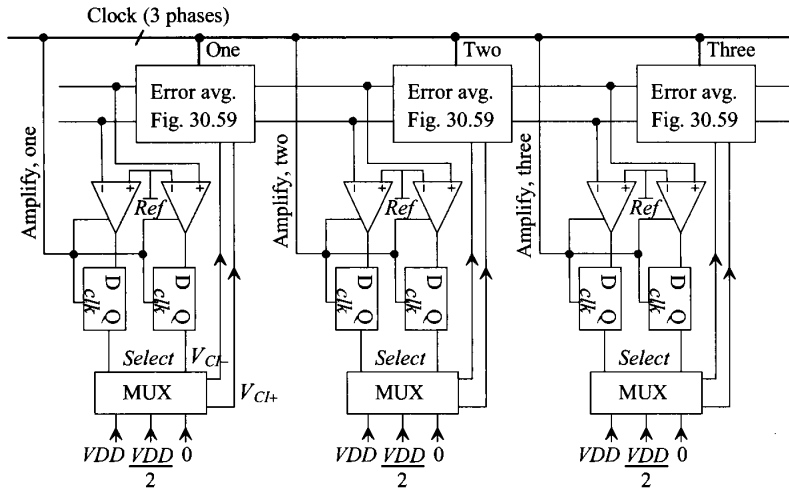


Figure 30.63 Placement of comparators in a pipeline ADC.

Offsets and Alternative Design Topologies

When discussing offsets, we've concentrated on the op-amp's offset. The offset associated with the common-mode voltage hasn't been discussed in detail. What happens if the output signals are moving around a voltage of $V_{CM} + V_{OS}$? Taking the difference in the output signals will eliminate both the common-mode voltage and the offset. A problem

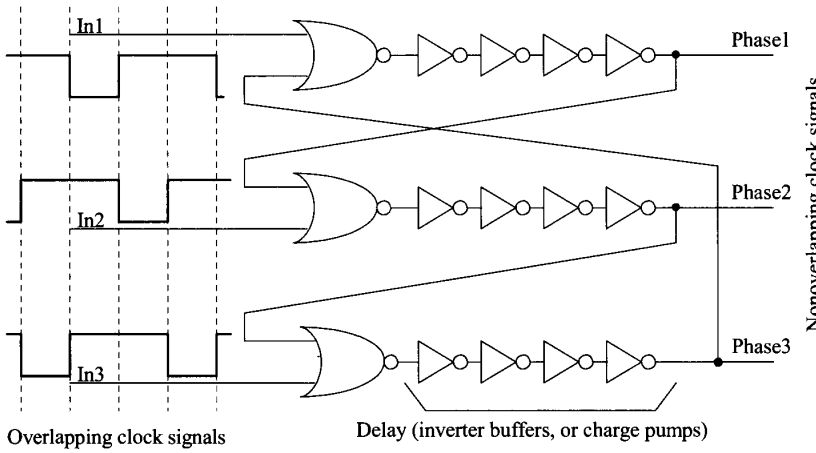


Figure 30.64 Circuit used for generating three phases of a nonoverlapping clock.

IPR2023-00698
Theta EX2019

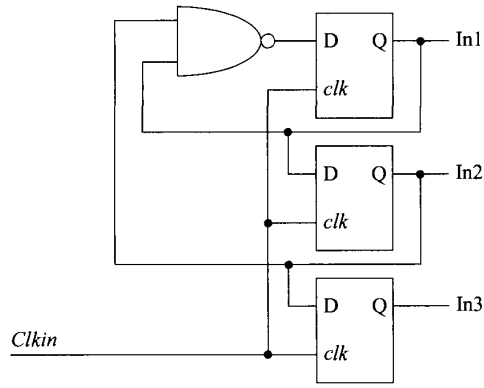


Figure 30.65 Generating the overlapping input clock signals for Fig. 30.64.

could occur if the CMFB circuit doesn't affect each output the same. (The result is a difference-mode signal.)

Example 30.21

Simulate the operation of the circuit shown in Fig. 30.66. Show the input and output signals as the difference between the two differential input signals. Assume the common-mode voltage coming out of the op-amp is precisely 500 mV.

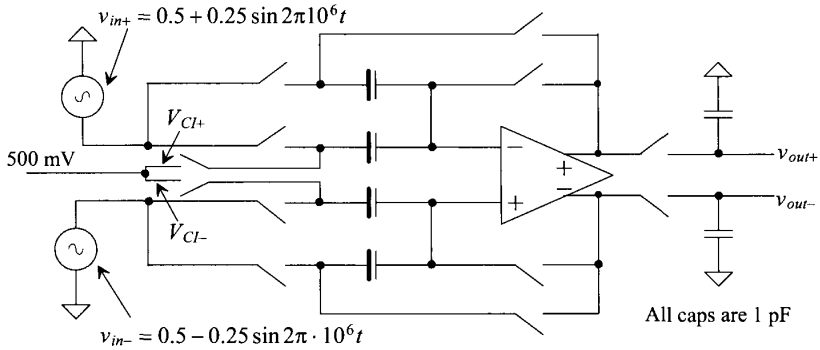


Figure 30.66 ADC building block discussed in Ex. 30.21.

The simulation results are shown in Fig. 30.67. Note how, as we would expect, the gain of the circuit is two. The signals shown are ideal. ■

Example 30.22

Repeat Ex. 30.21 if the common-mode voltage coming out of the op-amp sees an offset of 50 mV, i.e., it is 550 mV.

IPR2023-00698
Theta EX2019

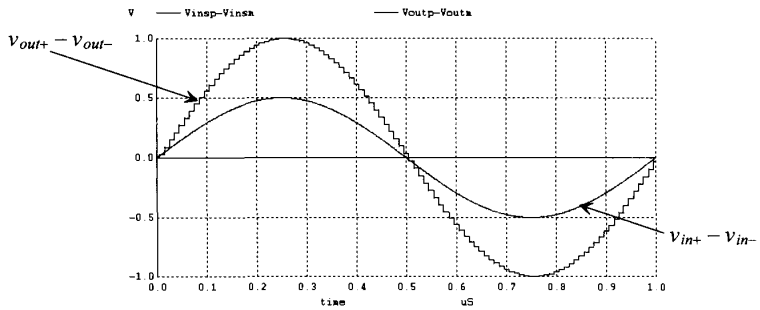


Figure 30.67 Input and output for the circuit of Fig. 30.66.

The differential input and output signals with this output common-mode offset look exactly like what is seen in Fig. 30.67. The single-ended output signals, Fig. 30.68, show the offset (the signals swing around 550 mV in Fig. 30.68). Clearly a common-mode offset in the op-amp output signals isn't a concern (unless it's so large that it limits the op-amp output swing range). Likewise an offset in the common-mode voltage of the input signals isn't a concern (this comment is easy to verify with simulations). What is a concern, though, is the value of the voltages used for V_{CI+} and V_{CI-}. ■

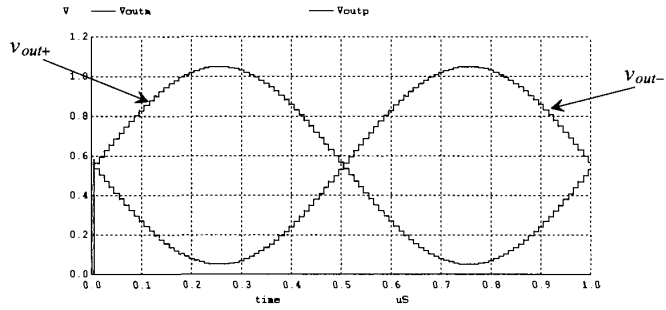


Figure 30.68 Output signals for the circuit of Fig. 30.66 if common-mode voltage is 0.55 V.

In Fig. 30.66 V_{CI+} and V_{CI-} are shorted together and connected, through a switch, to 0.5 V (V_{CM}). It can be shown that when this occurs, the absolute value of the voltage is irrelevant because it is common to both input signal paths, see Eq. (30.53). This means that if we used ground instead of V_{CM} in Fig. 30.66, we would get the same outputs. A problem does occur if the difference in the voltages (when adding or subtracting a voltage from the input signal) isn't exactly what is desired, see Eq. (30.61).

While an offset in the common-mode voltage isn't important, the op-amp's offset is a concern. The op-amp offset voltage is zeroed out when using the topology shown in Fig. 30.66. Equation (30.50) was derived assuming the op-amp had a nonzero offset voltage and shows the offset will not (ideally) affect the building block's output signals. To show the removal of the offset using simulations, consider the following example.

IPR2023-00698
Theta EX2019

Example 30.23

Simulate the operation of the circuit shown in Fig. 30.69. This schematic is Fig. 30.66 redrawn with a 50 mV op-amp offset.

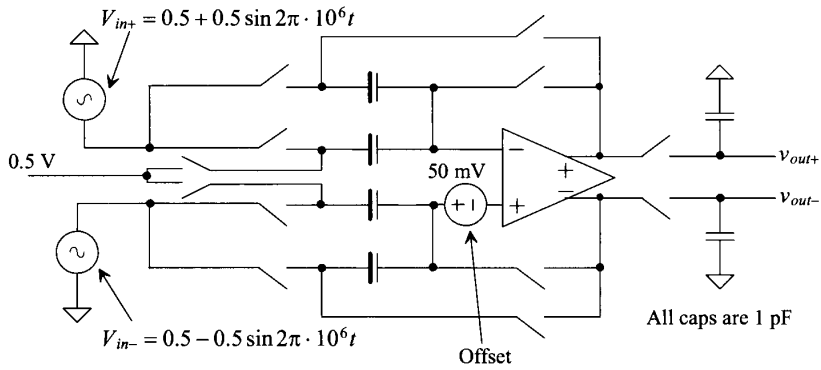


Figure 30.69 ADC building block discussed in Ex. 30.23.

The simulation results are shown in Fig. 30.70. Figure 30.70 should be compared to Fig. 30.67. Note how the large offset has no effect on the building block's output signals. ■

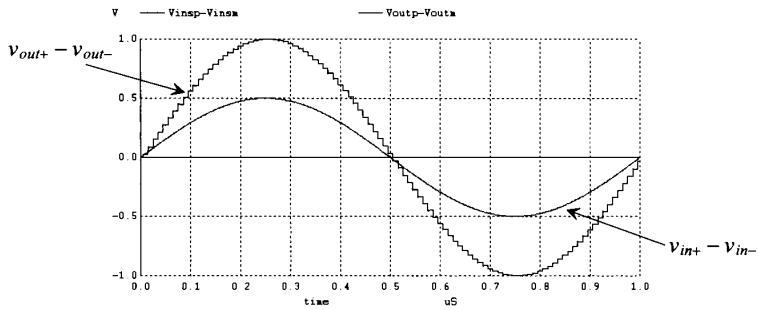


Figure 30.70 Input and output for the circuit of Fig. 30.69 with op-amp offset of 50 mV.

If the op-amp offset is zeroed out and doesn't affect the circuit's outputs when using the basic topology shown in Fig. 30.30, why would we want to consider some other topology? The answer to this question comes from the realization that the op-amp's outputs must settle to the final accuracy of the ADC, referring to Fig. 30.55, by the falling edge of all three phases of the clock. It would be nice to make the settling during one of these three (or two) phases irrelevant. Also, and perhaps more importantly, it would be nice to use other types of CMFB (discussed in the next section).

IPR2023-00698
Theta EX2019

Consider the building-block topology shown in Fig. 30.71. During the sampling phase, the inputs of the op-amp are shorted to the common-mode voltage. During this time the op-amp operates open loop, and settling time is meaningless. The status of the op-amp outputs during this time is discussed in the next section. The charge stored on the capacitors during the storage phase is

$$Q_{I,F}^{\phi_s} = C_{I,F} \cdot (v_{in} - V_{CM}) \tag{30.94}$$

while during the hold phase

$$Q_I^{\phi_h} = C_I \cdot (V_{CI} - V_{CM} \pm V_{OS}) \text{ and } Q_F^{\phi_h} = C_F \cdot (v_{out} - V_{CM} \pm V_{OS}) \tag{30.95}$$

Once again knowing charge must be conserved, the output voltage can be written as

$$v_{out} = v_{out+} - v_{out-} = \left(1 + \frac{C_I}{C_F}\right) \cdot (v_{in+} - v_{in-}) - \frac{C_I}{C_F} \cdot (V_{CI+} - V_{CI-}) \pm \overbrace{\left(1 + \frac{C_I}{C_F}\right) \cdot V_{OS}}^{\text{unwanted term}} \tag{30.96}$$

This equation should be compared to Eq. (30.53) where the offset was zeroed out during the sample phase of the clock (the phases ϕ_1 and ϕ_2). The topology would be useful in an ADC where we can guarantee that the offset contribution is well below the data converter's LSB or in a later stage in the ADC where the accuracy requirements are relaxed.

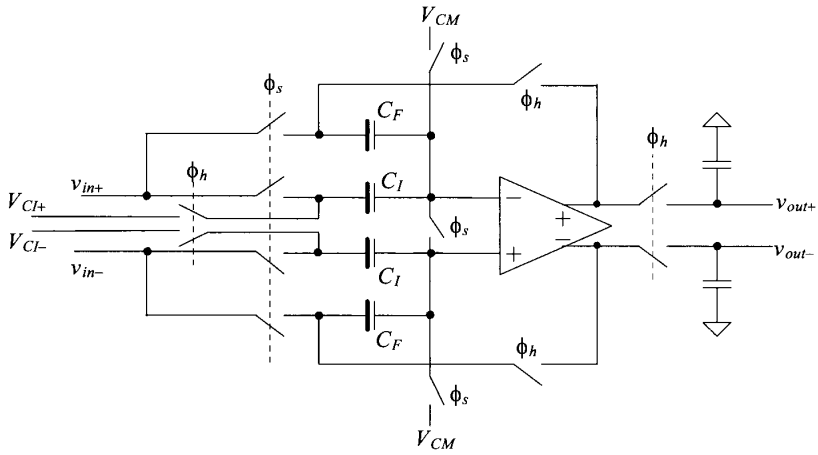


Figure 30.71 Alternative ADC building block.

Example 30.24

Repeat Ex. 30.23 using the topology shown in Fig. 30.71.

The simulation results are shown in Fig. 30.72. The ratio of C_I to C_F is one and so, ideally, the output is just twice as large as the input. The output should swing

IPR2023-00698
Theta EX2019

from 1 V down to −1 V. However, because of the 50 mV offset, which is also multiplied by a factor of 2, the output is shifted downwards so that it swings from 0.9 V to −1.1 V. ■

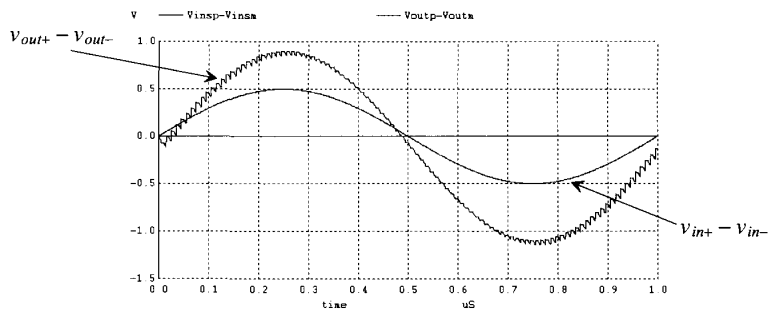


Figure 30.72 Input and output for the circuit of Fig. 30.71 with op-amp offset of 50 mV.

Dynamic CMFB

The CMFB circuits discussed earlier employ an amplifier to sense the average of the outputs and feedback a correction to center the signals around V_{CM}. In Ch. 26 we discussed a CMFB technique used in an op-amp that was dynamic and doesn't employ an amplifier. A simplification of this dynamic CMFB can be realized by noting that during the sampling phase in Fig. 30.71 the op-amp inputs are forced to V_{CM}. The scheme we are about to present won't force the inputs to $V_{CM} \pm V_{OS}$ during the sample phase as in the other topologies based on Fig. 30.30.

The basic dynamic CMFB circuit is shown in Fig. 30.73. During the sample phase of the clock the inputs and outputs of the op-amp are shorted to the common-mode voltage. Also during this time the common-mode feedback voltage, V_{CMFB}, is set to a bias voltage (V_{bias4} if the op-amp of Fig. 30.37 is used). During the hold phase, the CMFB capacitors on the output of the circuit are disconnected from both V_{CM} and V_{bias4} and are used to sense the average value of the outputs. If the outputs move in a balanced fashion, then V_{CMFB} remains equal to V_{bias4}. If the average of the outputs moves upwards above V_{CM}, then V_{CMFB} increases, pulling the output common-mode voltage downwards. Again, because the CMFB loop utilizes negative feedback, an increase in V_{CMFB} must result in a decrease in $(v_{o+} + v_{o-})/2$.

Looking at Fig. 30.73 we see that because of the op-amp's offset voltage the outputs of the op-amp will source/sink a current into V_{CM} during the sample phase of the circuit's operation. By adding an extra pair of switches to disconnect the CMFB capacitors from the op-amp outputs we can avoid this situation. Adding the switches causes the op-amp outputs to approach the power supply rails during the sample phase (because of the offset voltage). This output railing isn't a problem if an OTA (single-stage) topology like the one in Fig. 30.37 is used. The outputs have no capacity to hold charge and so when the CMFB capacitors are reconnected to the op-amp outputs, the outputs immediately go to V_{CM} (neglecting the connection of the feedback capacitors

IPR2023-00698
Theta EX2019

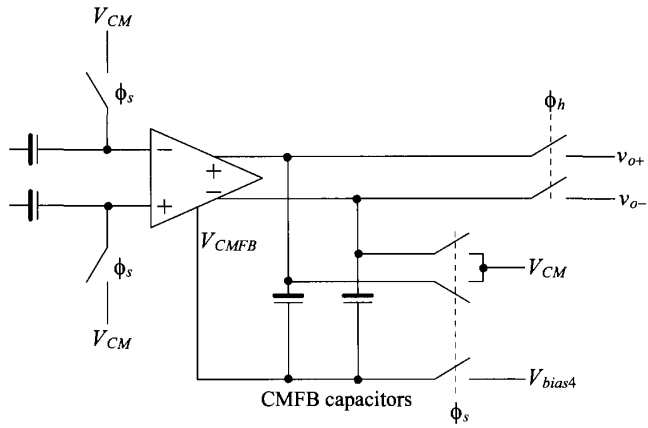


Figure 30.73 Dynamic CMFB.

which are not shown in Fig. 30.73). If a two-stage op-amp is used then care must be taken to ensure that the compensation capacitor (which, of course, does have the capacity to hold charge) doesn't charge to the rail resulting in a large recovery time when the op-amp circuit is put in the hold mode.

One thing that may help with the offset problem (when using the topology of Fig. 30.71) is to AC couple the output of the first op-amp stage to the input of the second op-amp stage, as seen in Fig. 30.74 (add a capacitor between the two stages to remove the DC component). The first stage's offset is stored on the capacitors during the sample phase, while the second stage's inputs are shorted to V_{bias1}. When the op-amp is used in the hold phase of operation, the second stage's offset is referred back to the input of the op-amp after dividing by the first stage's gain. For example, if the first stage of the op-amp has a gain of 10 and the second stage's offset is 10 mV, then the input-referred offset of the op-amp in the hold mode is 1 mV. Note that the gain of the first stage (diff-amp) can't be too large. If the diff-amp's gain were 100 and it had an input-referred offset of 10 mV then the diff-amp's outputs would move towards the supply rails causing devices (one of the PMOS loads and one of the NMOS in the diff-pair) to triode or shut-off. This limits the use of high-impedance loads in the diff-amp. As long as the MOSFETs remain in the saturation region, the offset storage works as desired.

Depending on the load of the diff-amp in Fig. 30.74, the op-amp would be compensated using either a load capacitance (using gate-drain connected loads as seen in the figure so that the topology is an OTA) or using indirect-feedback compensation if a current source load is used in the diff-amp, see Ch. 24. Again, care must be exercised when setting-up the biasing if using a current source load to avoid the PMOS load fighting the diff-amp's tail current for control of the current that flows in the diff-amp.

Layout of Pipelined ADCs

Before leaving this chapter, let's comment on the layout of pipeline ADCs. Using a fixed-height layout structure, with the automatically routed power and ground busses (when the

IPR2023-00698
Theta EX2019

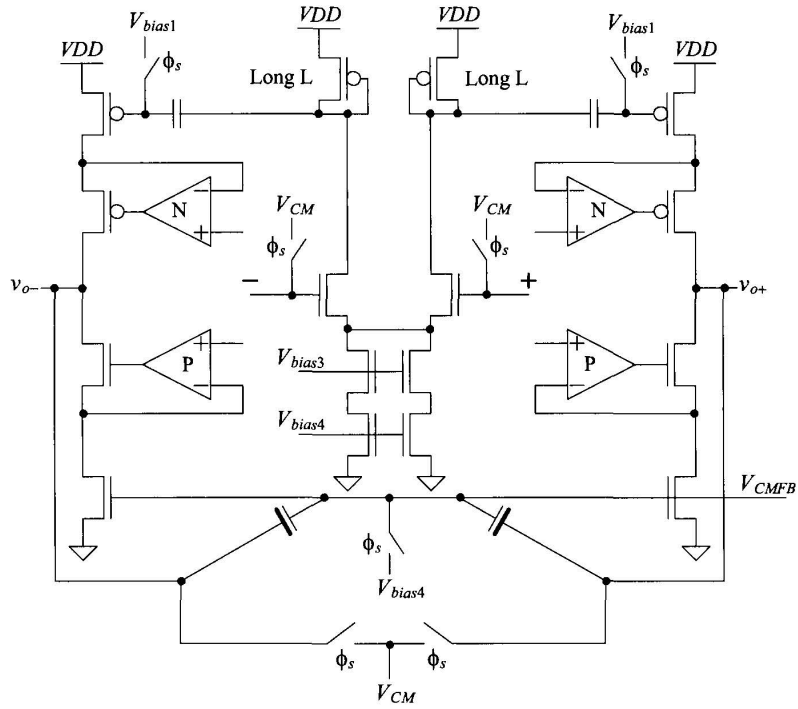


Figure 30.74 Implementation of dynamic CMFB.

cells are laid end-to-end), is a good place to start when laying out the op-amp. As seen in Fig. 30.75, a height (with example values shown) is selected with the width variable. It's important, as discussed in Ch. 28, to keep the analog signals separate, both physically and by distance, from the digital signals.

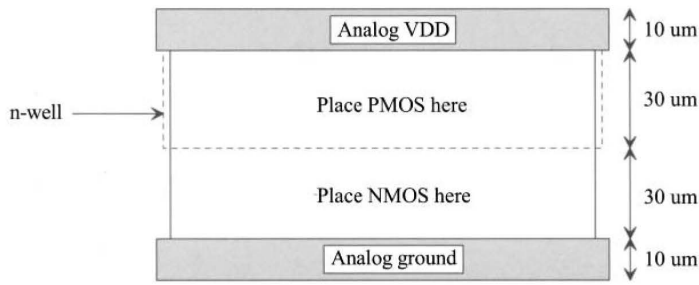


Figure 30.75 Fixed-height layout structure.

IPR2023-00698
Theta EX2019

A possible block diagram of the placement of the fixed-height cells together with the capacitors and switches used to implement a stage of the ADC is seen in Fig. 30.76. Looking at the input signal, we notice that they are laid out next to each other and routed as close as possible to the input of the first stage of the ADC. All of the differential analog signals in the ADC should follow this practice to help make any coupled noise truly common-mode. Notice how we have placed ground pads adjacent to the input signals. These pads are not used for ground connections on-chip. The pads are used to help reduce noise coupling into the input signals. Ideally, these ground pins provide a termination for the noise keeping the input signals "clean." This is especially important when bonding wires connect the integrated circuit to a padframe in the final packaged part. The bonding wires used for digital signals tend to radiate more than enough signal to corrupt the input signal and ruin the ADC's SNR when placed close to an analog low-level signal.

Next notice that we must have a clock signal in our analog domain. This signal, as we have seen, is used for clocking the switches in the S/H stage. Although in the figure we show the placement of the clock adjacent to the input signals, it may be better to move the pad and, if possible, the routing of the clock signals away from the inputs. The main goal when routing the clock signals is to keep the layout regular. Routing clock signals all over the layout is asking for problems.

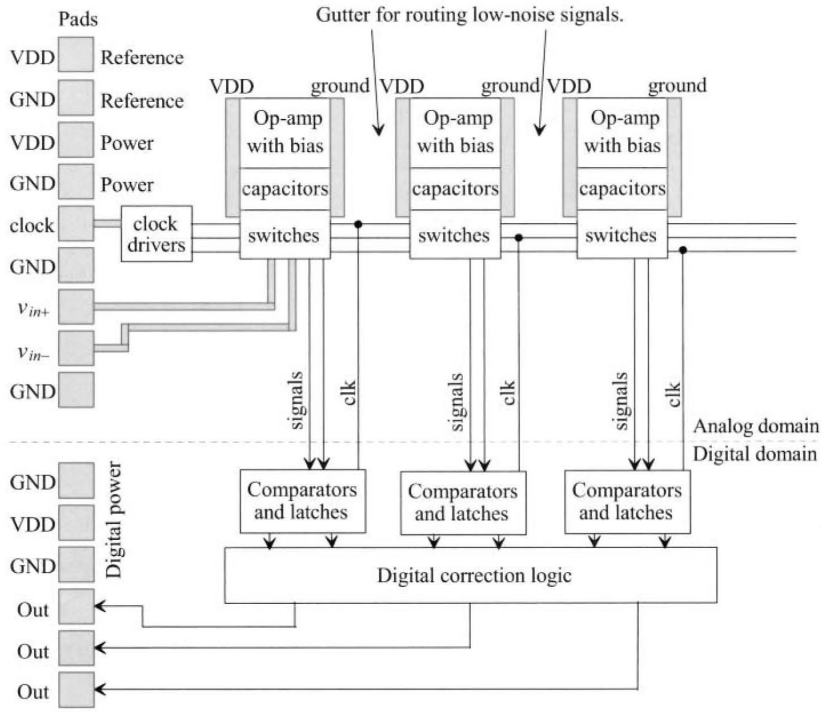


Figure 30.76 Block diagram layout of a pipeline stage.

IPR2023-00698
Theta EX2019

Two sets of power and ground pads are used (more if possible) in the analog domain for the implementation of the ADC. One set of pads is used for supplying power to the op-amps while the other set is used for supplying the reference voltages to each stage. The power and ground supplies are common to both digital and analog sections off-chip. Off-chip the supplies are connected together and decoupled (bypassed) using large capacitors (actually a wide range of capacitor values are connected in parallel between the *VDD* and ground to avoid the increase in a single large capacitor's effective series resistance with frequency). On-chip decoupling capacitors can be used as well. The analog and digital power and ground connections are not shared, and so care must be taken not to decouple the analog *VDD* to digital ground or digital *VDD* to analog ground. Figure 30.77 shows one example of how the decoupling capacitors can be connected.

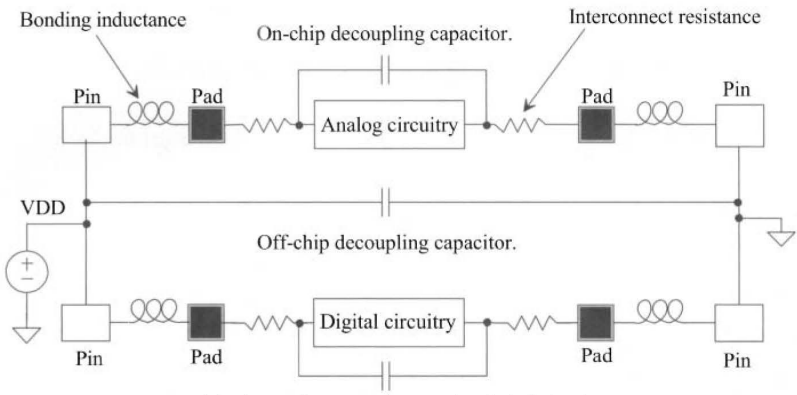


Figure 30.77 Showing how decoupling capacitors are used in a mixed-signal chip.

In general we don't want any DC current flowing on our reference voltages (those voltages used for V_{CI} in our op-amp) because of the possible voltage drop along the supplying line. There may also be a voltage drop along the metal lines supplying power to the op-amps. However, small (DC) changes in these voltages are usually not a significant factor in the precision of the ADC. These changes could be a factor if the output signal approaches the power-supply voltages where the op-amp runs out of head room. AC changes in the power-supply voltages can be a significant factor limiting the ADC's performance.

Ground planes and wide conductors should be used where possible. Using power and ground planes not only provides good distribution of power and ground but also increases the capacitance between the supplies. Areas, labeled "gutter" in Fig. 30.76 should be provided for low-level analog signals. These areas are free to allow the quiet routing of the switch inputs and outputs to the op-amp outputs.

It's also a good idea to use guard rings around the sensitive analog circuits (e.g., the switched capacitors) to avoid coupled substrate noise. Figure 30.78 shows the basic idea. In this figure the capacitors are laid out over an n-well. The n-well is tied to analog

IPR2023-00698
Theta EX2019

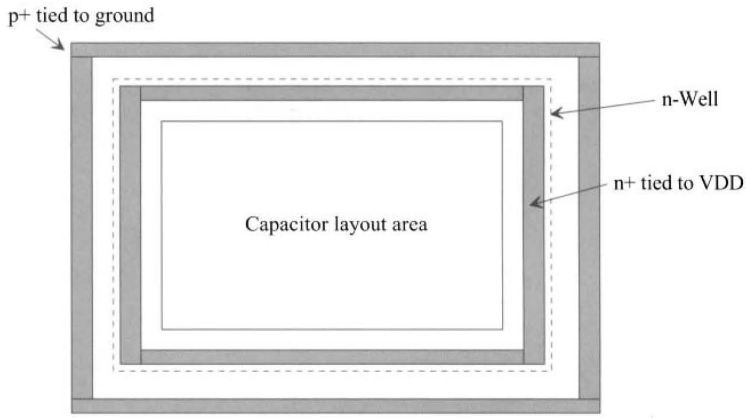


Figure 30.78 Using guard rings for protection in sensitive analog blocks.

VDD through an n+ implant in the n-well and metal. Surrounding the n-well is a ring of p+. This ring is tied to analog ground. The idea is that the p+ will provide a sink for any current injection from the surrounding circuitry. Because ground is the lowest potential in the circuit, the noise will terminate on this p+ and not penetrate the area under the capacitors (and then hopefully not couple into the capacitors). While this works well by itself, it may not be enough. Noise currents may still move deep in the substrate and work their way up under the capacitors. Because the n-well under the capacitor is held at the most postive potential in the circuit, any noise that does get into the n-well will hopefully be swept out through *VDD* and not couple into the capacitors.

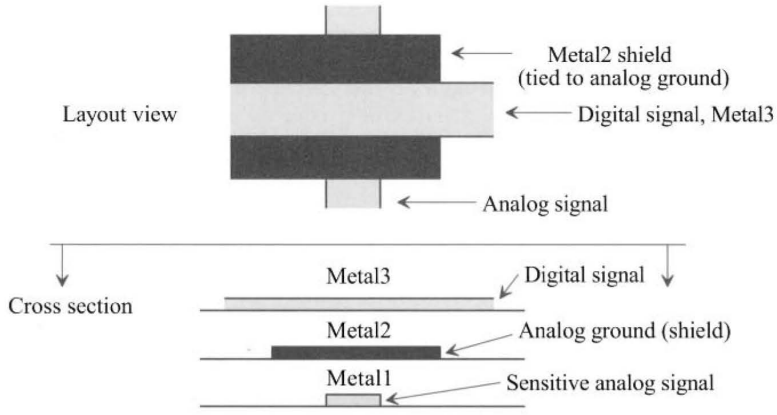


Figure 30.79 Shielding a sensitive analog signal from a digital signal.

IPR2023-00698
Theta EX2019

Finally, if a sensitive analog signal does need to cross a digital signal (an example being the potential need to feed the switch inputs and outputs across the three phases of the switch clock signals in Fig. 30.76) a shield should be used, Fig. 30.79. In this figure the sensitive analog signal is assumed to exist on Metal1, while Metal2 is used for an analog ground shield from the digital signal on Metal3. This shield is used for isolation providing a terminating plane for the electric fields resulting from the voltages on both the digital and analog signals.

ADDITIONAL READING

[1] R. J. Baker, *CMOS: Mixed-Signal Circuit Design, Second Edition*, Wiley-IEEE Press, 2009.

[2] Engineering Staff Analog Devices Inc., *Data Conversion Handbook (Analog Devices)*, Newnes, 2005. ISBN 978-0750678414

[3] R. J. van de Plassche, *CMOS Integrated Analog-to-Digital and Digital-to-Analog Converters*, Second Edition, Springer, 2003.

[4] S. Franco, *Design with Operational Amplifiers and Analog Integrated Circuits*, Third Edition, McGraw-Hill, 2003.

[5] H. S. Chen, B. S. Song, and K. Bacrania, "A 14-b 20-Msample/s CMOS Pipelined ADC," *IEEE Journal of Solid State Circuits*, Vol. 36, No. 6, pp. 997-1001, June 2001.

[6] M. Gustavsson, J. J. Wikner, and N. Tan, *CMOS Data Converters for Communications*, Springer, 2000. ISBN 978-0792377801

[7] C. C. Enz and G. C. Temes, "Circuit Techniques for Reducing the Effects of Op-Amp Inperfections: Autozeroing, Correlated Double Sampling, and Chopper Stabilization," *Proceedings of the IEEE*, Vol. 84, No. 11, pp. 1584-1614, November 1996.

[8] C. G. Yu and R. L. Geiger, "An Automatic Offset Compensation Scheme with Ping-Pong Control for CMOS Operational Amplifiers," *IEEE Journal of Solid-State Circuits*, Vol. 29, No. 5 May 1994, pp. 601-610.

[9] B. Ginetti, P. G. A. Jespers, and A. Vandemeulebroecke, "A CMOS 13-b Cyclic RSD A/D Converter," *IEEE Journal of Solid-State Circuits*, Vol. 27, No. 7, July 1992.

[10] B. S. Song, M. F. Tompsett, and K. R. Lakshmikumar, "A 12-bit, 1-MSample/s Capacitor Error-Averaging Pipelined A/D Converter," *IEEE Journal of Solid State Circuits*, Vol. 23, No. 6, pp. 1324-1333, December 1988.

[11] W. C. Black and D. A. Hodges, "Time Interleaved Converter Arrays," *IEEE Journal of Solid-State Circuits*, Vol. SC-15, No. 6, December 1980.

[12] J. A. Schoeff, "An Inherently Monotonic 12-bit DAC," *IEEE Journal of Solid-State Circuits*, Vol. SC-14, No. 6 December 1979, pp. 904-911.

[13] R. E. Suarez, P. R. Gray, and D. A. Hodges, "All-MOS Charge Redistribution Analog-to-Digital Conversion Techniques - Part II," *IEEE Journal of Solid-State Circuits*, Vol. 10, No. 6, pp. 379-385, December 1975.

IPR2023-00698
Theta EX2019

Problems

30.1 Assuming the DAC shown in Fig. 30.1 is 8 bits and $V_{REF+} = 1$ V and $V_{REF-} = 0$, what are the voltages on each of the R-$2R$ taps?

30.2 Give an example of how the traditional current-mode DAC will have limited output swing.

30.3 Repeat problem 30.1 for the DAC shown in Fig. 30.2.

30.4 For the wide-swing current mode DAC shown in Fig. 30.3, what are the voltages at the taps along the R-$2R$ string assuming 8 bits, $V_{REF+} = 1$ V, $V_{REF-} = 0$, and a digital input code of 0000 0000?

30.5 Can the op-amp shown in Fig. 30.37 be used in fully-differential implementations of the DACs shown in Figs. 30.1 - 30.3? Why or why not?

30.6 Show the detailed derivation of Eqs. (30.12)-(30.14).

30.7 Why would we want to use both current segments and binary-weighted currents to implement a current-mode DAC? (Why use segmentation?)

30.8 Why do we subtract ΔA in Eq. (30.36)? Why not add the gain variation?

30.9 Does the matching of the capacitors matter in the S/H of Fig. 30.31? Why or why not?

30.10 Derive the transfer function of the S/H in Fig. 30.30 if V_{CM} on the left side of the schematic is replaced with ground so that the bottom plates of the C_1 capacitors are grounded when ϕ_3 goes high.

30.11 Determine the transfer function of the S/H in Fig. 30.34 if the top left ϕ_2-controlled switch is connected to the input instead of V_{CM}. Include the effects of offsets and simulate the operation of the circuit to verify your calculations.

30.12 Repeat Ex. 30.10 if the cyclic ADC's input is 0.41 V.

30.13 Is kick-back noise from the comparator a concern for the circuit of Fig. 30.39?

30.14 Derive the transfer function for the circuit shown in Fig. 30.80.

30.15 Repeat Ex. 30.16 if the input voltage is 0.41 V.

30.16 Repeat Ex. 30.17 if the input voltage is 0.41 V.

30.17 Resketch the clock waveforms for Fig. 30.54 if bottom plate sampling is used.

30.18 Show the derivation leading up to Eq. (30.83). Show, using practical values for mismatch, how the squared mismatch terms are negligible.

30.19 What happens to the error adjustment term in Eq. (30.92) if the capacitors in the S/H are perfectly matched?

30.20 Repeat Ex. 30.18 if all capacitors are 1 pF (the ideal situation) and verify that the error out of the stage is zero.

30.21 Sketch a circuit to provide the inputs for the four-phase, nonoverlapping clock generator shown in Fig. 30.81.

IPR2023-00698
Theta EX2019

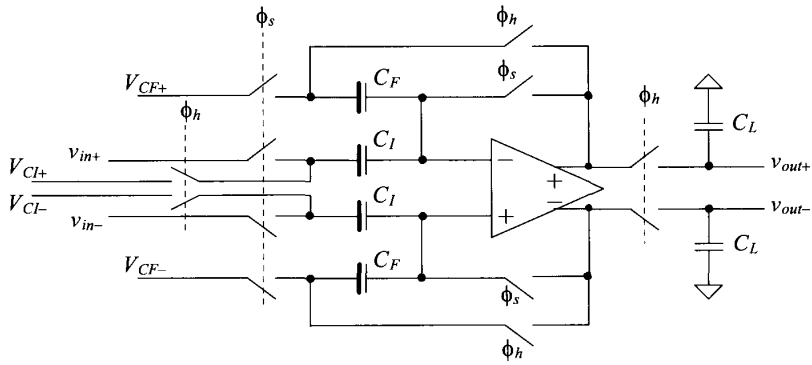


Figure 30.80 Circuit used in problem 30.14.

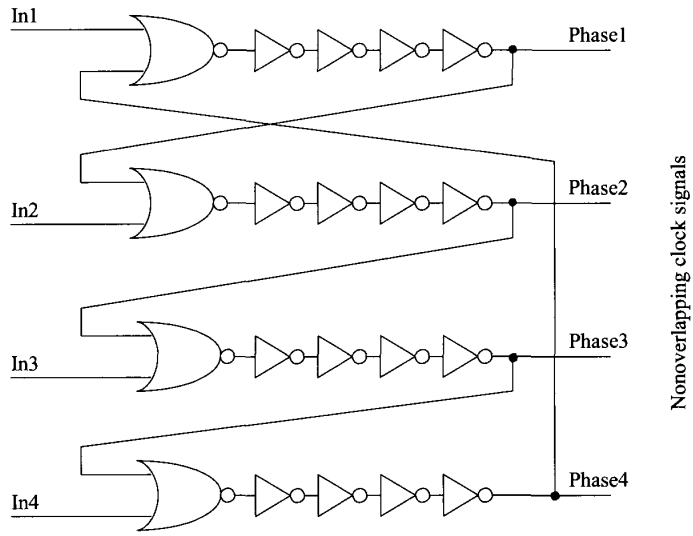


Figure 30.81 Four-phase, nonoverlapping clock generator.

30.22 What is the main advantage of using dynamic CMFB over other CMFB circuits? What is the main disadvantage?

30.23 Can MOSFETs be used to implement the on-chip decoupling capacitors in Fig. 30.77?

IPR2023-00698
Theta EX2019

30.24 Sketch the cross-sectional view of the layout in Fig. 30.78.

30.25 Figure 30.82 shows the implementation of a pipeline DAC. How would we implement this DAC using a topology similar to Fig. 30.42? Sketch the DAC's implementation and the timing signals (clock phases) used.

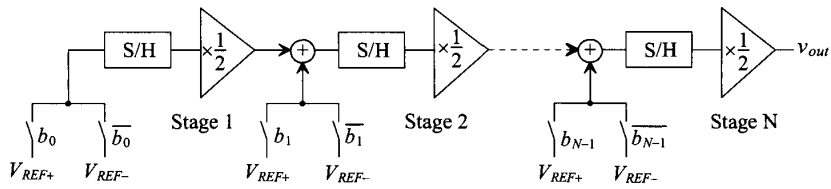


Figure 30.82 A pipeline digital-to-analog converter.

IPR2023-00698
Theta EX2019

Chapter
31

Feedback Amplifiers

Back in Sec. 30.2 we compared op-amps in series-shunt and shunt-shunt feedback topologies without going into the details of what these terms mean. The goals of this chapter are to further develop feedback theory and its application to integrated circuit design by discussing the design of feedback amplifiers.

Feedback, in general, is a very powerful concept that has numerous applications. Examples of feedback systems abound in everyday life. For example, the thermostat on an air conditioner unit uses feedback to maintain a comfortable temperature within a room. Similarly, our bodies use feedback to increase antibodies to fight off an infection. By definition, feedback is the process of combining the output of a system with its input. In the air conditioner example, the temperature of the room is considered to be the output of the system, and the temperature set on the thermostat, the system input. The thermostat subtracts the value of the room temperature from the input value. If the temperature in the room is higher than the temperature set on the thermostat, the air conditioner turns on until the temperature in the room is less than or equal to the desired set value.

Thus, it can be said that feedback stabilizes a system. However, not all types of feedback have this property. There are two types of feedback: positive and negative. Negative feedback stabilizes, while positive feedback has the opposite effect. A good example of positive feedback is when a microphone is held too closely to the speaker of a public address (PA) system. More than likely, most people have heard the loud ringing that occurs. The positive feedback causes the system to become unstable, and the undesired effect is clearly recognizable.

Positive feedback occurs when the system output is added to the system input, whereas negative feedback occurs when the system output is subtracted from its input. For our purposes in this chapter, only negative feedback will be considered. While positive feedback can be useful if controlled, its applications are discussed elsewhere in the book. We will, however, discuss methods for minimizing the effects of positive feedback.

IPR2023-00698
Theta EX2019

31.1 The Feedback Equation

Consider the feedback system shown in Fig. 31.1. The variables used in this diagram are labeled as an x value because they may be either a voltage or a current. The input signal, x_s, and the output signal x_o, may be either a voltage or a current. However, the feedback signal, x_f, must be the same type of signal as the input signal. If the input signal is a voltage, then the feedback signal must also be a voltage.

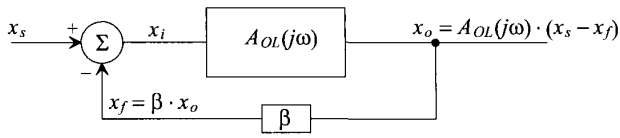


Figure 31.1 Basic block diagram to illustrate the feedback concept.

The block A_{OL} represents an amplifier's open-loop gain and is frequency dependent. The input to the amplifier, x_i, is the difference in the input source signal and the feedback signal, or

$$x_i = x_s - x_f \tag{31.1}$$

and the output is given by

$$x_o = A_{OL}(j\omega) \cdot (x_s - x_f) \tag{31.2}$$

In the following discussion, we will assume that all components of the system are ideal, meaning that the feedback network, β, will not load the amplifier. The specifications for the system can then be defined as follows:

$$A_{OL} = \frac{x_o}{x_i} \tag{31.3}$$

The *feedback factor*, β, is defined as

$$\beta = \frac{x_f}{x_o} \tag{31.4}$$

and the closed-loop gain, A_{CL}, is

$$A_{CL} = \frac{x_o}{x_s} \tag{31.5}$$

Realizing that

$$x_f = \beta \cdot x_o \tag{31.6}$$

and plugging Eq. (31.6) into (31.2) and solving for the closed-loop gain, we find that A_{CL} becomes

$$A_{CL} = \frac{x_o}{x_s} = \frac{A_{OL}}{1 + A_{OL}\beta} \tag{31.7}$$

where the frequency dependence of A_{OL} has not been shown.

IPR2023-00698
Theta EX2019

Note the dependence of the A_{CL} on the value of A_{OL}. If the value of A_{OL} becomes large (A_{OL} approaches infinity), then the value of A_{CL} approaches $\frac{1}{\beta}$. This illustrates the need for having a high-gain amplifier. If A_{OL} is very large, then the closed-loop gain becomes highly dependent on the feedback components.

The term $A_{OL}\beta$ is often referred to as the *loop gain* and will be used later sections to determine overall amplifier stability (see, also, Sec. 24.1)

31.2 Properties of Negative Feedback on Amplifier Design

When used in the design of amplifiers, feedback can provide a number of advantages. These advantages include desensitizing the gain to process parameter variation, reducing nonlinear distortion, reducing the effects of noise, extending the useful bandwidth of the amplifier, and controlling the input and output impedance levels.

31.2.1 Gain Desensitivity

Since the value of the open-loop gain, A_{OL}, is large, its value may change significantly with temperature, mismatch of devices, and other parameter variations. However, negative feedback desensitizes the closed-loop gain from changes in the open-loop gain. The following derivation illustrates this property.

Differentiating both sides of Eq. (31.7) yields

$$\frac{dA_{CL}}{dA_{OL}} = \frac{1}{(1+A_{OL}\beta)^2} \text{ or } dA_{CL} = \frac{dA_{OL}}{(1+A_{OL}\beta)^2} \tag{31.8}$$

Dividing each side of Eq. (31.8) by the corresponding factors in Eq. (31.7),

$$\frac{dA_{CL}}{A_{CL}} = \frac{1}{(1+A_{OL}\beta)} \cdot \frac{dA_{OL}}{A_{OL}} \tag{31.9}$$

where $\frac{dA_{CL}}{A_{CL}}$ represents the fractional change in A_{CL} for a given fractional change in $\frac{dA_{OL}}{A_{OL}}$.

Using Eq. (31.9), if A_{OL} = 10,000 V/V (and assuming a voltage amplifier) and β = 1/10 V/V, it can be seen that if $\frac{dA_{OL}}{A_{OL}}$ = 10 %, then the change in the closed-loop gain, $\frac{dA_{CL}}{A_{CL}}$ = 0.01 %! This can easily be verified by using Eq. (31.7) for A_{OL}=10,000 and 9,000, keeping β = 1/10 V/V and solving for A_{CL} for each case. The resulting difference proves the immunity of the closed-loop gain to changes in A_{OL}.

31.2.2 Bandwidth Extension

Negative feedback also increases the usable bandwidth of an amplifier. Assume that an amplifier has the frequency response given by

$$A_{OLH}(s) = \frac{A_{OL} \cdot \omega_H}{s+\omega_H} = A_{OL}\frac{1}{\frac{s}{\omega_H}+1} \tag{31.10}$$

A_{OLH}(s) is simply the frequency dependent version of A_{OL} and is approximated using a first-order pole at ω_H. Plugging Eq. (31.10) into Eq. (31.7) yields the high-frequency dependent closed-loop version of Eq. (31.7)

IPR2023-00698
Theta EX2019

$$A_{CLH}(s) = \frac{A_{OLH}(s)}{1 + A_{OLH}(s) \cdot \beta} = \frac{A_{OL} \cdot \omega_H}{s + \omega_H \cdot (1 + A_{OL}\beta)} \tag{31.11}$$

which can be rewritten as

$$A_{CLH}(s) = \frac{A_{OL}}{(1 + A_{OL}\beta)} \cdot \frac{1}{\frac{s}{\omega_H(1+A_{OL}\beta)} + 1} = \frac{A_{OL}}{(1 + A_{OL}\beta)} \cdot \frac{1}{\frac{s}{\omega_{HF}} + 1} \tag{31.12}$$

Note that the resulting equation is composed of two parts. The first factor should be recognized as Eq. (31.7) and is simply the expression for the closed-loop gain at midband frequencies. The second factor is the frequency dependent term. It is interesting to observe that the original –3 dB frequency ω_H in Eq. (31.10) is now multiplied by $(1+A_{OL}\beta)$. Figure 31.2 illustrates the bandwidth extension (at the cost of gain) from using negative feedback. The original open-loop frequency response is drawn in the solid line, while the closed-loop frequency response is illustrated in the dashed line. The closed-loop response shows a decrease in the gain and at the same time an increase in bandwidth.

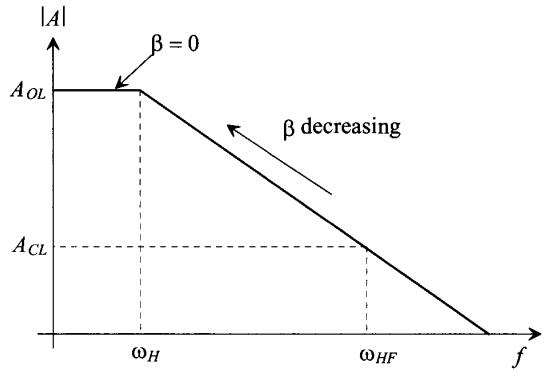


Figure 31.2 Extension of the high-frequency pole by using feedback.

If one were to decrease the value of β, from Eq. (31.11), it can be seen that two interesting effects occur. First, the closed-loop value of the gain will increase, since the first factor in Eq. (31.12) increases with a decrease in β. Second, the frequency response decreases since ω_H is multiplied by $(1 + A_{OL}\beta)$ and β is decreasing. As β decreases towards 0 the overall curve follows the original open-loop curve. It should be seen that when using feedback one trades gain for bandwidth.

The same analysis can be applied to the low-frequency response of an amplifier. Equation (31.13) approximates the low-frequency response of an amplifier with a single first-order low-frequency pole

$$A_{OLL}(s) = A_{OL} \cdot \frac{s}{s + \omega_L} \tag{31.13}$$

If we plug Eq. (31.13) into Eq. (31.7), the closed-loop low-frequency response becomes

$$A_{CLL}(s) = \frac{A_{OLL}(s)}{1 + A_{OLL}(s)\beta} = \frac{A_{OL}\frac{s}{s+\omega_L}}{1 + A_{OL}\frac{s}{s+\omega_L}\beta} = \frac{A_{OL}}{1 + A_{OL}\beta} \cdot \frac{s}{s + \frac{\omega_L}{1+A_{OL}\beta}} \tag{31.14}$$

IPR2023-00698
Theta EX2019

The result in Eq. (31.14) is comprised of the standard closed-loop gain at midband and the frequency dependent term. Note however, that with feedback, the original low-frequency pole, ω_L, in Eq. (31.13) is now divided by $1+A_{OL}\beta$ in Eq. (31.14). Figure 31.3 illustrates the effect of the feedback on the low-frequency response of an amplifier.

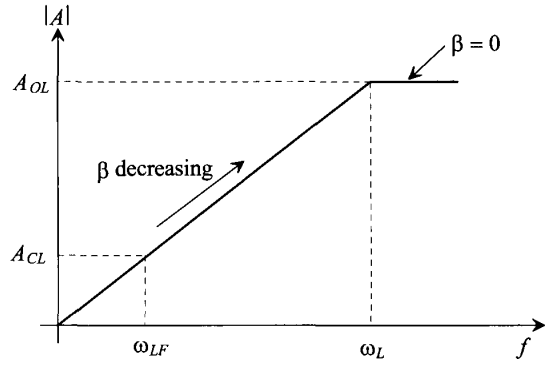


Figure 31.3 Extension of the low-frequency bandwidth by using feedback.

31.2.3 Reduction in Nonlinear Distortion

Negative feedback can also improve the nonlinear behavior of an amplifier. Examine the transfer curve of the voltage amplifier shown in Fig. 31.4 without feedback. Ideally, the amplifier should have a straight line from $-2\text{ V} < V_{in} < 2\text{ V}$. However, nonlinear behavior in the amplifier causes a different slope to occur when $V_{in} > 1\text{V}$ and $V_{in} < -1\text{ V}$. Note that the output voltage is limited by the value of the power supply (+ 15 V).

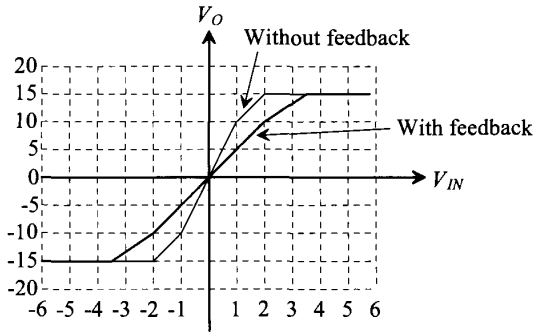


Figure 31.4 Using feedback to improve amplifier linearity.

Since the gain of the amplifier can be defined as the slope of the transfer curve, we can deduce that

IPR2023-00698
Theta EX2019

$$A_{OL} = \frac{\Delta V_O}{\Delta V_{IN}} = 10\,\text{V/V} \quad \text{for } -10\,\text{V} \le V_O \le 10\,\text{V} \tag{31.15}$$

and

$$A_{OL} = \frac{\Delta V_O}{\Delta V_{IN}} = 5\,\text{V/V} \quad \text{for } V_O > 10 \text{ and } V_O < -10\,\text{V} \tag{31.16}$$

Now suppose that feedback is applied around the amplifier with $\beta = 0.1$ V/V and the transfer curve is redrawn. The gain of the amplifier with feedback becomes

$$A_{CL} = \frac{A_{OL}}{1 + A_{OL}\beta} = \frac{10}{1 + 10(0.1)} = 5\,\text{V/V} \quad \text{for } -10\,\text{V} < V_O < 10\,\text{V} \tag{31.17}$$

and

$$A_{CL} = \frac{5}{1 + 5(0.1)} = \frac{10}{3} \quad V_O > 10\,\text{V and } V_O < -10\,\text{V} \tag{31.18}$$

The resulting transfer curve is seen in Fig. 31.4. Note that with feedback, the overall transfer curve is much more linear, resulting in an amplifier that has less nonlinear distortion than the original amplifier shown in Fig. 31.4.

31.2.4 Input and Output Impedance Control

The input and output impedances of an amplifier can also be controlled using negative feedback. As seen in Fig. 31.5, R_i is the small-signal input impedance looking into an amplifier without feedback, and R_{inf} is the input impedance with feedback applied. Similarly, R_{of} and R_o are the output impedances with and without feedback, respectively. Feedback allows us either to increase or decrease both R_{inf} and R_{of} by a factor of $(1 + A_{OL}\beta)$. Although an example of this property is difficult to derive in a general way, proofs of this property will be illustrated with actual circuit examples in a later section. For the time being, it will be sufficient to summarize the impedance control properties of feedback with Table 31.1. Note that the closed-loop impedances are completely dependent on the type of variable that is used (voltage or current) at the input and the output.

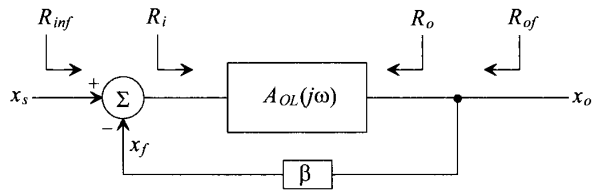


Figure 31.5 Determining the input and output impedances with and without feedback.

If the input variable, x_s, is a voltage, the closed-loop input resistance, R_{inf}, is equal to the open-loop value of the input resistance, R_{in}, multiplied by the value of $1 + A_{OL}\beta$. If the input variable is a current, R_{inf} is divided by the same factor. Alternatively, if the output variable, x_o, is a voltage, then the closed-loop output impedance is the open-loop output impedance divided by $(1 + A_{OL}\beta)$. And if the output variable is a current, the open-loop output impedance is multiplied by the same factor.

IPR2023-00698
Theta EX2019

Obviously, this concept is a powerful one. We can adjust both input and output impedances in a larger or smaller direction simply by choosing the type of feedback used in the circuit.

Table 31.1 Summary of impedances with feedback.

Input Variable, x_s	Output Variable, x_o	R_{inf}	R_{of}
V	V	$R_i \cdot (1 + A_{OL}\beta)$	$R_o / (1 + A_{OL}\beta)$
V	I	$R_i \cdot (1 + A_{OL}\beta)$	$R_o \cdot (1 + A_{OL}\beta)$
I	V	$R_i / (1 + A_{OL}\beta)$	$R_o / (1 + A_{OL}\beta)$
I	I	$R_i / (1 + A_{OL}\beta)$	$R_o \cdot (1 + A_{OL}\beta)$

31.3 Recognizing Feedback Topologies

Examine the general single-loop feedback circuit in Fig. 31.6. Remember that x_s, x_i, and x_f must all either be currents or all be voltages, and that the output variable, x_o may be either a voltage or a current. As a result of these restrictions, a total of four types of feedback are possible (Table 31.2).

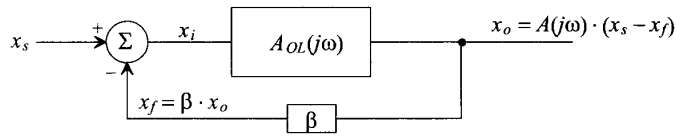


Figure 31.6 Basic block diagram to illustrate the feedback concept.

The input summation is often referred to as input mixing. If the input variables, x_s, x_i, and x_f can be written as voltages, the mixing is referred to as *series* or voltage mixing. If the input variables can be written as currents, it is referred to as *shunt* or current mixing. The type of variable used at the output determines the second term, known as sampling. If the output variable is a voltage, the sampling is then referred to as *shunt* or voltage sampling. If the output variable is a current, the sampling is known as *series* or current sampling.

Table 31.2 Feedback type as a function of system variables.

x_s, x_i, x_f	x_o	Feedback Type (mixing-sampling)
Voltage	Voltage	Series-shunt
Voltage	Current	Series-series
Current	Current	Shunt-series
Current	Voltage	Shunt-shunt

IPR2023-00698
Theta EX2019

In the analysis of feedback amplifiers, the following terminology will apply. The term *basic amplifier* will correspond to the gain circuit, A_{OL}, in the block level feedback diagrams. When determining the value of A_{OL}, it is very important to determine the loading due to the β network and any source and load resistance. The β network is also known as the feedback network. The overall circuit that includes both A_{OL} and β will be referred to as the *feedback amplifier.*

31.3.1 Input Mixing

Figures 31.7a and b illustrate the inputs to two generic feedback amplifiers. The basic amplifier in 31.7a is used in a series mixing configuration since the equation, $x_i = x_s - x_f$ can be written in terms of the voltages. Note also that the basic amplifier and the β network are in series with each other. In Fig. 31.7b, however, the basic amplifier is used in a shunt mixing configuration. Here, the summation at the input can only be written in terms of currents such that $i_i = i_s - i_f$. The feedback circuit and the basic amplifier circuit are both in parallel, or it can be said that the feedback shunts the basic amplifier.

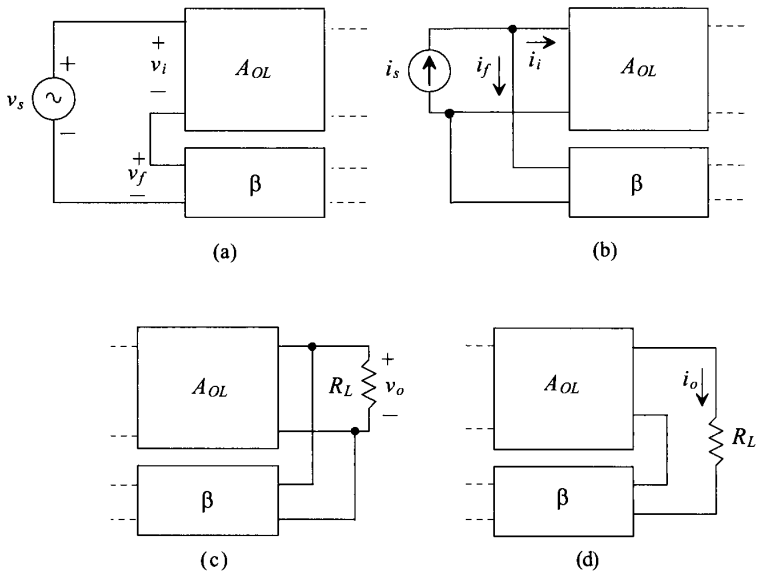


Figure 31.7 Generic block diagrams for input circuits (a) series mixing and (b) shunt mixing and output circuits, (c) shunt sampling, and (d) series sampling.

31.3.2 Output Sampling

Determining the type of variable used at the output is not an obvious endeavor. However, a general model followed by several examples will help clarify the proper procedure. Figures 31.7c and d show the outputs of two generic feedback amplifiers. Shunt sampling

IPR2023-00698
Theta EX2019

is shown in Fig. 31.7c. Note that the feedback circuit is shunting, or in parallel with the basic amplifier circuit. The output variable is considered a voltage because the feedback circuit "senses" or samples the voltage across R_L. Series sampling is seen in Fig. 31.7d. Here, the feedback network is in series with the basic amplifier circuit. The output variable is a current since the feedback circuit now senses the current through the load resistor. Two rules that help distinguish between the two types of output sampling are:

Rule 1: If one terminal of the output active device (drain or source) is driving the load and the other terminal is attached to the feedback network, the output sampling is series.

Rule 2: If the load and the feedback network are connected to the same node, the output sampling is shunt.

31.3.3 The Feedback Network

When analyzing feedback circuits, it would be wise first to recognize the basic amplifier circuit, A_{OL}, and the β network by tracing the small-signal path from the input to the output and back through the feedback. The path from the input through the A_{OL} circuit to the output will be denoted as the forward path and the path from the output back through the β network, as the feedback path. Determining which path the signal takes is not obvious when there are several devices from which to choose. Some rules that will aid in the analysis are as follows:

- The *forward path* through the basic amplifier circuit will always take the path that has the highest gain.
- The AC small-signal will always enter either a gate or a source and will always exit either a drain or a source. The gain from drain to source is very small (at least for linear applications) and is considered to be negligible for most of our applications. A small-signal will never exit through the gate.
- The feedback signal must subtract from the input signal. To ensure that this is the case, one must count the number of inversions around the loop. Every time a signal crosses a gate-to-drain junction (common source amplification), an inversion occurs. Examples of this will be discussed next.

Examine Fig. 31.8a. Note that the signal path in the forward direction progresses through the path with the highest gain—from the gate of M1 to the drain of M1, into the gate of M2 and out the drain of M2. The feedback path consists of the resistors R_1 and R_2. The feedback variable consists of the voltage that appears across R_1.

An Important Assumption

One might attempt to draw the forward path from the gate of M1 to the source of M1, through the feedback resistor, R_2, and to the output. In actuality, this would be a legitimate forward path because in some cases the feedback network is actually bidirectional. However, the signal will have very little gain since the gain from gate to source of any MOS device is a maximum of one (common drain amplification) and the gain from v_f to v_2 in the forward direction will be less than one because of the voltage divider relationship between R_5 and R_2. *Thus, for all of the analysis performed in this*

IPR2023-00698
Theta EX2019

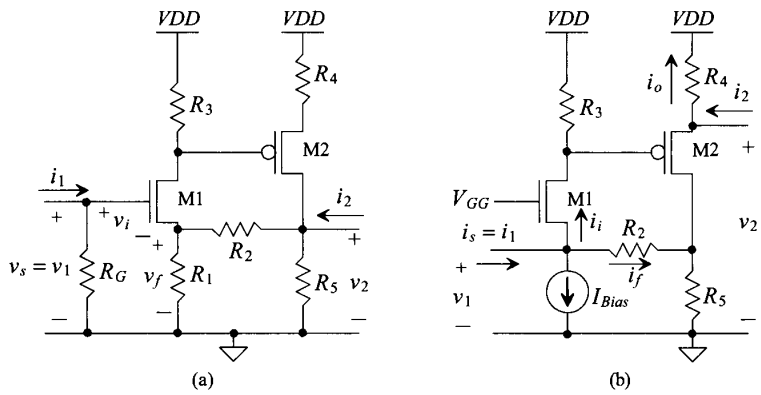


Figure 31.8 Feedback topologies (a) series-shunt and (b) shunt-series.

chapter, the forward path through the amplifier circuit dominates the expression for the total forward gain from input to output, while the forward path through the feedback is assumed to be negligible. This important assumption greatly simplifies the analysis and is a good one, since the amplifier will have very poor performance if there is not a good deal of gain through the basic amplifier. (Ideally, the value of A_{OL} is infinite.) If an active device is used in the feedback network, the forward gain through the β network is minimized.

Figure 31.8b has the same basic transistor topology, except it uses current as its input. The forward path consists of the source of M1, out the drain of M1 into the gate of M2 and out the source of M2. The feedback path consists of R_5 and R_2, and the feedback variable is the current, i_f. Note that the forward path has the highest possible gain path, since M1 acts as a common gate amplifier. The forward path could progress from the input through R_2 and into the drain of M2. However, since the small-signal gain from the drain to the source of M2 is extremely small, it will be neglected.

Counting Inversions Around the Loop

It is also necessary to check the circuit to ensure that the feedback is indeed negative. Counting the number of inversions around the loop in Fig. 31.8a may initially give the wrong impression, because the number of gate-to-drain junctions encountered by the signal around the loop is two. This implies positive feedback since the feedback variable must be inverted with respect to the input signal. However, the final mixing between v_1 and v_f will provide the additional negative behavior needed to ensure the proper feedback. Notice the relationship that exists between v_1 and v_f. Since $v_i = v_1 - v_f$ the voltage v_f has a subtractive effect on v_i. Thus, a positive change in the signal entering the gate of M1 will result in a positive change in the source of M1 and a smaller, stable voltage v_i. If v_f had been negative with respect to v_1, then v_i would be, $v_i = v_s + v_f$ and positive feedback would occur.

IPR2023-00698
Theta EX2019

Now examine Fig. 31.8b. Again, if we count the number of inversions around the loop from input back to the mixing variable, i_f, the number of inversions is odd, which is exactly what is needed, since KCL tells us that $i_i = i_1 - i_f$. The direction of the feedback current from the output is opposite the direction of i_f; however, the signal is inverted, making it equivalent to i_f, as illustrated from Fig. 31.9.

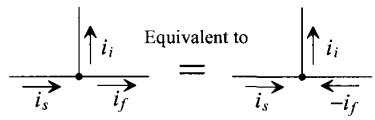


Figure 31.9 Shunt mixing illustrated with an odd number of inversions.

Examples of Recognizing Feedback Topologies

Now that the method for recognizing feedback has been discussed, examples will be presented to solidify the concepts. Again examine Fig. 31.8. Here, two types of feedback are illustrated. Transferring the block diagram concepts to transistor-level circuits can be a difficult task. However, several other rules can be applied which will reveal the type of sampling used in the circuit. The term *input active device* is used to denote the transistor that is being driven by the input source. The term *output active device* is used to denote the transistor used to drive the load.

In Fig. 31.8a, the input variables can only be written in terms of voltages such that $v_i = v_s - v_f$. We could attempt to sum currents at the node on the gate. However, the current flowing into the resistor, R_G, will not be the result of any feedback, thus making current mixing impossible. The output sampling is of the shunt type since the β network and the basic amplifier are in parallel (the β network is connected to the same node as the output) and v_o is the voltage being sampled. This amplifier employs series mixing and shunt sampling, and is referred to as a series-shunt feedback amplifier configuration.

Figure 31.8b illustrates shunt-series feedback. The DC current source is viewed as an AC open circuit, and the DC voltage source, V_{GG}, is considered as an AC short circuit. Note how the input variables can only be written in terms of currents forming the expression, $i_i = i_s - i_f$. Series sampling is used at the output since the feedback network is in series with the basic amplifier. This amplifier configuration also follows rule 2, mentioned previously; thus, the proper small-signal output variable is the current, i_o. Note that the direction of the small-signal output current, i_o, is consistent with the direction of the small-signal model, since *VDD* is considered a small-signal ground.

The variables v_1, i_1, v_2, and i_2 may or may not correspond to the defined input and output variables of the feedback circuit. For example, in Fig. 31.8b, the correct output variable is defined as i_o. However, the gain of the feedback amplifier can be ascertained in terms of v_1, i_1, v_2, and i_2, since $v_2 = i_o \cdot R_4$. Also note that v_1/i_1 and v_2/i_2 correspond to the input and output impedances, respectively, of the feedback amplifier.

Series-series feedback is illustrated in Fig. 31.10a. Although the input variables are written in terms of voltages, the output variable is considered to be a current since the

IPR2023-00698
Theta EX2019

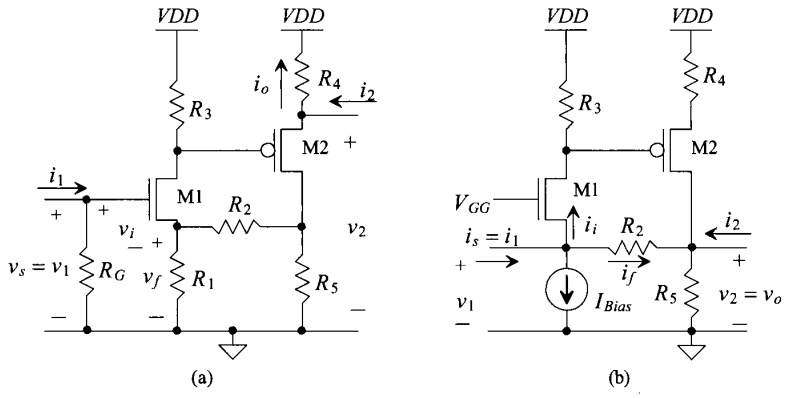


Figure 31.10 Feedback topologies: (a) series-series and (b) shunt-shunt.

feedback network is in series with the amplifier output. Note that the only difference between Figs. 31.10a and 31.8a is in placement of the output terminal. The same holds true for Figs. 31.10b and 31.8b. Lastly, Fig. 31.10b illustrates shunt-shunt feedback. Current summation occurs at the input which typifies shunt or current mixing, and the β network made up of R_5 and R_2 shunts the output.

31.3.4 Calculating Open-Loop Parameters

Once the feedback topology can be recognized, the analysis of the circuit can begin. We will perform two types of analysis: open-loop and closed-loop. Open-loop analysis entails approximately 80 percent of the circuit analysis required to solve a feedback problem. Open-loop values are then used to calculate the closed-loop values, so extreme care must be taken when analyzing the open-loop circuit. We will then "close the loop" and calculate the feedback characteristics of the overall feedback amplifier

The method of analysis for feedback amplifiers can be confusing if we do not distinguish between open- and closed-loop notation. We will denote the open-loop voltage and current variables by use of the * notation. The open-loop parameters include A_{OL}, R_i, R_o, and β and are defined as follows.

$$A_{OL} = \frac{x_o^*}{x_s^*} \text{ and is the open-loop gain of the basic amplifier} \quad (31.19)$$

$$R_i = \frac{v_1^*}{i_1^*} \text{ and is the open-loop input impedance} \quad (31.20)$$

$$R_o = \frac{v_2^*}{i_2^*} \text{ and is the open-loop output impedance} \quad (31.21)$$

$$\beta = \frac{x_f^*}{x_o^*} \text{ and is the gain through the feedback network} \quad (31.22)$$

In our discussion of general feedback principles in Sec. 31.1, it was assumed that the β network was ideal, meaning that the impedance of the β network did not load the

IPR2023-00698
Theta EX2019

amplifier circuit. However, in real-life applications, the β network can cause loading effects on both the input source and the output of the amplifier circuit. The type of feedback used will determine how the β network loading is calculated. As we progress in the discussion of each feedback topology, the method of determining the β network loading will be presented.

All variables in the following analysis are considered to be small-signal AC voltages or currents. We will perform the open-loop analysis using the following steps:

1. Replace input source with Norton or Thevenin equivalent circuit and calculate open-loop gain. The amplifier circuit will produce a different type of gain for each type of feedback. For example, the open-loop gain for an amplifier used in a series-shunt configuration will have units of V/V. This is a standard voltage amplifier since the output variable is a voltage and the input mixing sums voltages. An amplifier using series-series feedback will have a gain with units of I/V, since the output variable is now a current and the input mixing uses voltages. This type of amplifier is also known as a transconductance amplifier since the units of the gain, I/V, are equivalent to a conductance. Similarly, an amplifier used in the shunt-series configuration has a gain with units of I/I (a current amplifier), and an amplifier used in a shunt-shunt configuration will have a gain with units of V/I and is known as a transimpedance amplifier.
2. When calculating the open-loop gain of the circuit, consider the β network loading. An equivalent resistance, $R_{\beta i}$ and $R_{\beta o}$, as seen in Fig. 31.11 will represent the total input and output resistance of the β network. The method for calculating $R_{\beta i}$ involves the following steps.
   - If the output sampling is shunt (voltage), short the output node to ground.
   - If the output sampling is series (current), remove the device driving the output load as if it were taken "out-of-socket."
   - Calculate $R_{\beta i}$ as the impedance looking from the input into the β network.

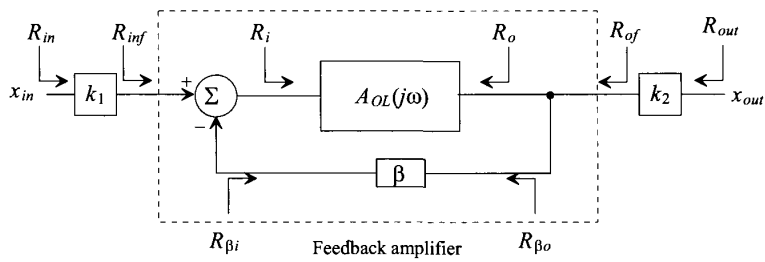


Figure 31.11 Block diagram of a generic feedback amplifier that distinguishes between open- and closed-loop impedances.

The method for calculating $R_{\beta o}$ involves similar methodology:

IPR2023-00698
Theta EX2019

- If the input mixing is shunt (current), short the input node to ground.
- If the input mixing is series (voltage), remove the input active device as if it were taken "out-of-socket."
- Calculate $R_{\beta o}$ as the impedance looking from the output into the β network.

The method for determining the loading of the β network is rooted in basic two-port theory where each type of feedback configuration corresponds to one of the four basic two-port topologies. For further information regarding two-port theory, the reader is advised to consult [1]. One method used to easily remember the above rules is to remember that if mixing or sampling is shunt, then "short" the input or output, respectively. If the mixing or sampling is series, then "sever" the input or output device, respectively, by taking it "out-of-socket."

3. Calculate feedback factor, β, for the amplifier. The open-loop circuit is analyzed to determine the gain from the output back to the point where the feedback variable mixes with the input.
4. Determine the open-loop input and output impedances, R_i and R_o. These impedances can be calculated using standard circuit analysis techniques.

31.3.5 Calculating Closed-Loop Parameters

Once all the open-loop parameters are obtained, the closed-loop values are easily calculated. The closed-loop gain is

$$A_{CL} = \frac{x_o}{x_s} = \frac{A_{OL}}{1 + A_{OL} \cdot \beta} \tag{31.23}$$

for all four topologies. The closed-loop input impedances represent only the input and output impedances of the feedback amplifier. The input impedance from the input source, R_{in}, may correspond to R_{inf} if there is no gain from the source, x_{in}, to the input of the feedback amplifier, x_s. Similarly, the value of the output resistance, R_{out}, may or may not correspond to R_{of}, depending on the type of sampling used. Figure 31.11 illustrates this distinction. Calculating the input and output impedances of the feedback amplifier is dependent on the type of mixing and sampling circuits used, respectively.

$$\text{For series input mixing, } R_{inf} = R_i(1 + A_{OL}\beta) \tag{31.24}$$

$$\text{For shunt input mixing, } R_{inf} = \frac{R_i}{(1 + A_{OL}\beta)} \tag{31.25}$$

The value of the closed-loop output impedances will be as follows:

$$\text{For series output sampling, } R_{of} = R_o(1 + A_{OL}\beta) \tag{31.26}$$

$$\text{For shunt output sampling, } R_{of} = \frac{R_o}{(1 + A_{OL}\beta)} \tag{31.27}$$

Equations (31.24)-(31.27) will be derived with each specific topology. However, it is important to note the effectiveness of using feedback to control input and output impedances. Note that if series mixing or series sampling is used, the open-loop value of the input or output impedance is multiplied by $(1 + A_{OL}\beta)$; and if shunt mixing or shunt sampling is used, the open-loop values of impedances are divided by the same factor.

IPR2023-00698
Theta EX2019

Now that the general methodology has been described, a detailed examination of the four feedback topologies will be presented. The analysis of the four basic feedback topologies begins at the discrete level and progresses into more complex integrated circuits throughout the section.

31.4 The Voltage Amp (Series-Shunt Feedback)

Consider the ideal voltage feedback amplifier shown in Fig. 31.12, with open-loop values, A_{OL}, β, R_i, and R_o already given. The basic amplifier is a voltage amplifier with gain given V/V. Since this is an ideal feedback amplifier, the β network does not load the basic amplifier, meaning that $R_{\beta o} = \infty$ and $R_{\beta i} = 0$.

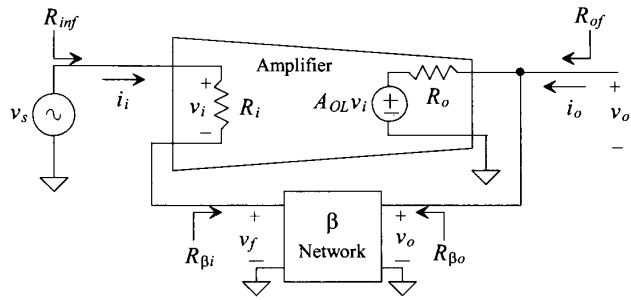


Figure 31.12 An ideal voltage amplifier (series-shunt).

To determine what type of feedback is present in the circuit, examine the input variables. Since input variables x_s, x_i, and x_f correspond to the voltages v_s, v_i, and v_f, such that the equation $v_i = v_s - v_f$ can be written the input mixing is series. The output mixing can be determined using the previous rules. Since the output of the amplifier, A_{OL}, and the β network are attached in parallel (both being connected to the load, R_L), the output mixing is considered to be shunt.

From Sec. 31.1, we already know that the closed-loop gain of the amplifier is

$$A_{CL} = \frac{v_o}{v_s} = \frac{A_{OL}}{1 + A_{OL} \cdot \beta} \tag{31.28}$$

Notice that as A_{OL} approaches infinity, Eq. (31.28) approximates to

$$A_{CL} \approx \frac{1}{\beta} \tag{31.29}$$

Equation (31.29) is important because it illustrates another powerful feedback concept. The entire gain of the feedback amplifier can be approximated as the inverse of β as the basic amplifier gain increases to higher and higher values.

We can calculate how the feedback affects the input impedance of the amplifier by applying a test voltage directly to the input of the feedback amp shown in Fig. 31.12, so that $v_s = v_{test}$ and $i_i = i_{test}$. Writing a voltage loop at the input and assuming that the feedback network does not load the amplifier gives

IPR2023-00698
Theta EX2019

$$v_{test} = i_{test} \cdot R_i + \beta \cdot v_o = i_{test} \cdot R_i + \beta \cdot \frac{A_{OL}}{1 + A_{OL}\beta} \cdot v_{test} \qquad (31.30)$$

Equation 31.30 simplifies to

$$R_{inf} = \frac{v_{test}}{i_{test}} = R_i \cdot (1 + A_{OL}\beta) \qquad (31.31)$$

The gain of the open-loop amplifier was decreased by $1 + A_{OL}\beta$, while the input impedance, the impedance the source sees, was increased by $1 + A_{OL}\beta$.

Calculation of the output impedance proceeds in the same way as the calculation of the input resistance except that the test voltage is applied to the output of the amplifier with the input shorted to ground with $v_o = v_{test}$ and $i_o = i_{test}$. Writing a loop equation at the output gives

$$v_{test} = i_{test} \cdot R_o + A_{OL} \cdot v_i = i_{test} \cdot R_o + A_{OL} \cdot (-\beta \cdot v_o) \qquad (31.32)$$

since the input is shorted to ground and $v_i = -v_f = -\beta \cdot v_o$. The output impedance is given by

$$R_{of} = \frac{v_{test}}{i_{test}} = \frac{R_o}{1 + A_{OL}\beta} \qquad (31.33)$$

or the output resistance is reduced by $1 + A_{OL}\beta$. Ideally, a voltage amplifier has infinite input resistance and zero output resistance. Adding feedback to a voltage amplifier with finite input resistance and nonzero output resistance help make the amplifier closer to the ideal.

Now that the ideal series-shunt amplifier has been examined from a block level point of view, a discussion of the nonideal series-shunt feedback amplifier at the transistor level with loading effects will be considered. The amplifier seen in Fig. 31.13 was analyzed previously and was determined to use series-shunt feedback.

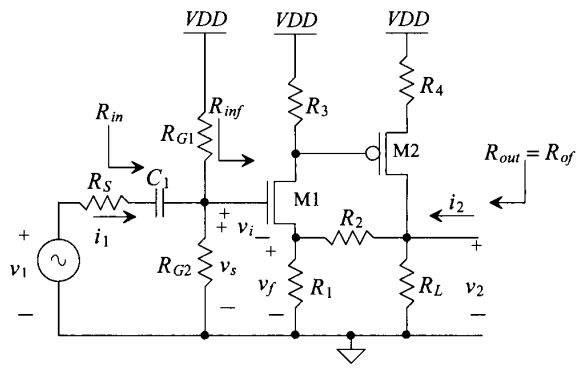


Figure 31.13 Transistor-level series-shunt feedback amplifier.

IPR2023-00698
Theta EX2019

The small-signal circuit of the feedback circuit is seen in Fig. 31.14. Note that the forward path consists of the following nodes: 1, 2, 3. The feedback path consists of nodes 3 and 4, with the feedback variable, v_f, appearing across R_1. In the previous discussion, it was assumed that the β network did not load the amplifier circuit. However, to accurately calculate the open-loop gain, A_{OL}, the loading of R_1 and R_2 on both the input and the output of the amplifier circuit needs to be considered. Note that the resistor R_S is initially ignored, since it is essentially outside the feedback amplifier.

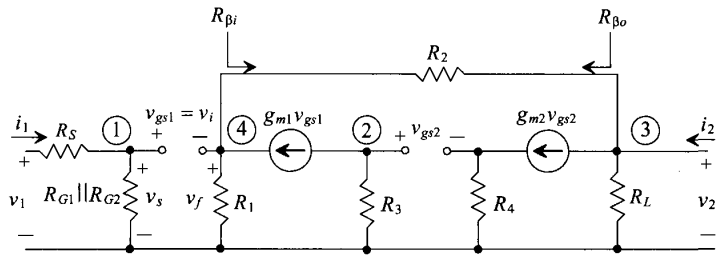


Figure 31.14 Closed-loop small-signal model of Fig. 31.13.

Since we are analyzing a series-shunt amplifier, we may determine the loading caused by the β network on the input, $R_{\beta i}$ and the output $R_{\beta o}$ in the following way (refer to Fig. 31.15). Looking into the β network from the input, we observe the resistance seen with the output terminal shorted to ground. The equivalent resistance to ground seen is the loading of the β network seen by the input of the amplifier. In this example, R_2 is seen. Therefore, in the open-loop model used to determine A_{OL}, we will include R_2 in parallel with R_1. The loading at the output is found similarly. Since the input mixing is series, we will remove M1 "out-of-socket" and look into the β network from the output. The equivalent resistance seen is then attached to the output of the open-loop model. In this example, the equivalent resistance is $R_2 + R_1$ and is attached to the output of the

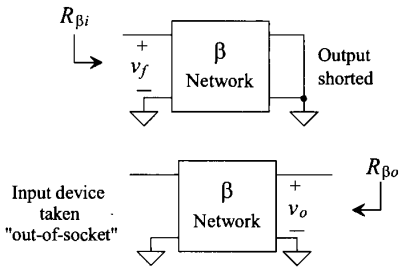


Figure 31.15 Determining the loading due to the feedback network for a series-shunt amplifier.

IPR2023-00698
Theta EX2019

open-loop model. The resulting open-loop model is seen in Fig. 31.16. We will initially assume that r_o for the MOSFETS is much larger than the discrete resistors and that the bulk and source are tied together (v_{sb} = 0). As we progress through the chapter, more difficult circuits will include drain-to-source resistances in our small-signal analysis.

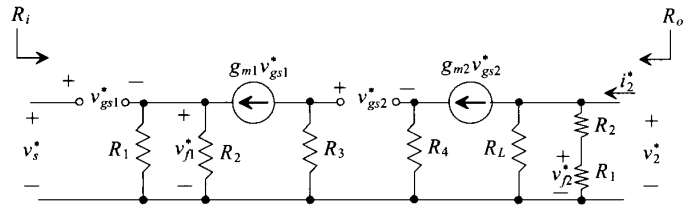


Figure 31.16 Open-loop small-signal model of Fig. 31.13.

The open-loop model is now ready to be analyzed in order to calculate A_{OL}. Since we are using a series (voltage)-shunt (voltage) feedback amplifier, the units of A_{OL} will be V/V and

$$A_{OL} = \frac{v_2^*}{v_s^*} \tag{31.34}$$

Solving for A_{OL} yields

$$A_{OL} = \frac{v_2^*}{v_s^*} = \left(\frac{v_2^*}{v_{gs2}^*}\right)\left(\frac{v_{gs2}^*}{v_{gs1}^*}\right)\left(\frac{v_{gs1}^*}{v_s^*}\right) = [-g_{m2}R_L \| (R_2 + R_1)]\left[-\frac{g_{m1}R_3}{1 + g_{m2}R_4}\right]\left[\frac{1}{1 + g_{m1}(R_1 \| R_2)}\right] \tag{31.35}$$

Next, the value of β can also be calculated from the open-loop model. Remembering that β is defined as the gain from the output back to the input mixing variable, v_f, we can write

$$\beta = \frac{v_f^*}{v_2^*} = \frac{R_1}{R_1 + R_2} \tag{31.36}$$

since the β network is simply a voltage divider relationship. Notice that the open-loop circuit now contains two values of R_2 and v_f^*. In this example, since r_o was assumed to be infinite, the gain from v_2^* to v_{f1}^* will be zero. If r_o had not been neglected, the gain from v_2^* to v_{f1}^* would have been small but finite. Therefore, it can be said that a reverse path exists through the basic amplifier as well as through the feedback network. However, the gain from v_2^* to v_{f2}^*, though less than one, will be significantly larger than from v_2^* to v_{f1}^*. *Therefore, just as the forward path through the feedback network was neglected, the reverse path through the basic amplifier is assumed to be much smaller than the reverse path through the feedback path. Therefore, the value of* β *is calculated using the resistor, R_2, closest to the output.*

Next, the value for R_i and R_o will be calculated. These values are determined using the open-loop model generated in Fig. 31.16. Since we are using MOS devices, it should

IPR2023-00698
Theta EX2019

be obvious that the input resistance to the open-loop circuit is ∞. The output resistance, however, can be calculated shorting the gate of M1 to ground and applying a test voltage to the output. Since the input is grounded, $v_{g2} = 0$, and R_o is simply the parallel combination of resistances at the output (assuming that r_{o2} is very large).

$$R_o = \frac{v_o^*}{i_2^*} = R_L || (R_1 + R_2) \tag{31.37}$$

Once the open-loop values are calculated, the closed-loop values are easily attained. The closed-loop values are

$$A_{CL} = \frac{v_2}{v_s} = \frac{A_{OL}}{1 + A_{OL}\beta} \tag{31.38}$$

$$R_{inf} = \frac{v_s}{i_s} = R_i(1 + A_{OL}\beta) \tag{31.39}$$

$$R_{out} = R_{of} = \frac{v_2}{i_2} = \frac{R_o}{1 + A_{OL}\beta} \tag{31.40}$$

Analysis of this problem has not yet been completed. Notice that we initially neglected the source resistance and the biasing resistors, R_{G1} and R_{G2}, since they played no part in the feedback analysis of the amplifier. However, they will have an effect in the overall gain and need to be considered. The total gain of the entire circuit is

$$\frac{v_2}{v_1} = \frac{v_s}{v_1} \cdot \frac{v_2}{v_s} = \frac{R_{G1} || R_{G2}}{R_{G1} || R_{G2} + R_S} \cdot A_{CL} \tag{31.41}$$

and the value of R_{in} as seen by the signal source is

$$R_{in} = \frac{v_1}{i_1} = R_{G1} || R_{G2} \tag{31.42}$$

and the analysis is now complete. Notice that the value of R_{in} is not the same as R_{inf}.

Example 31.1

For the series-shunt circuit shown in Fig. 31.17a, draw the closed-loop small-signal model and identify the forward and feedback paths, draw the open-loop model, calculate A_{OL}, β, R_i, and R_o, and calculate the closed-loop parameters v_2/v_1, v_1/i_1, and v_2/i_2. You may assume that $r_o = \infty$ for both devices and that DC analysis has already been performed with $g_{m1} = g_{m2} = 1$ mA/V.

The closed-loop small-signal model can be seen in Fig. 31.17b with the forward and feedback paths drawn. The open-loop model with loading effects is seen in Fig. 31.17c. Since the output uses shunt sampling, the value of $R_{\beta i}$ is found by shorting the output node and looking at the load resistance resulting from the feedback network. In this case, $R_{\beta i}$ is equal to R_2. Since the input mixing is series, the input device, M1, is "severed," so that only $R_1 + R_2$ is seen looking into the feedback loop from the output. Thus, $R_{\beta o} = R_1 + R_2$.

The values of r_o are considered to be infinite, so the values of R_i and R_o can be found by inspection, $R_i = \infty$, and $R_o = R_1 + R_2 = 11$ kΩ. Notice that when solving the open-loop circuit (Fig. 31.17c) v_1 and R_G are not included since we are only interested in solving the feedback portion of the circuit.

IPR2023-00698
Theta EX2019

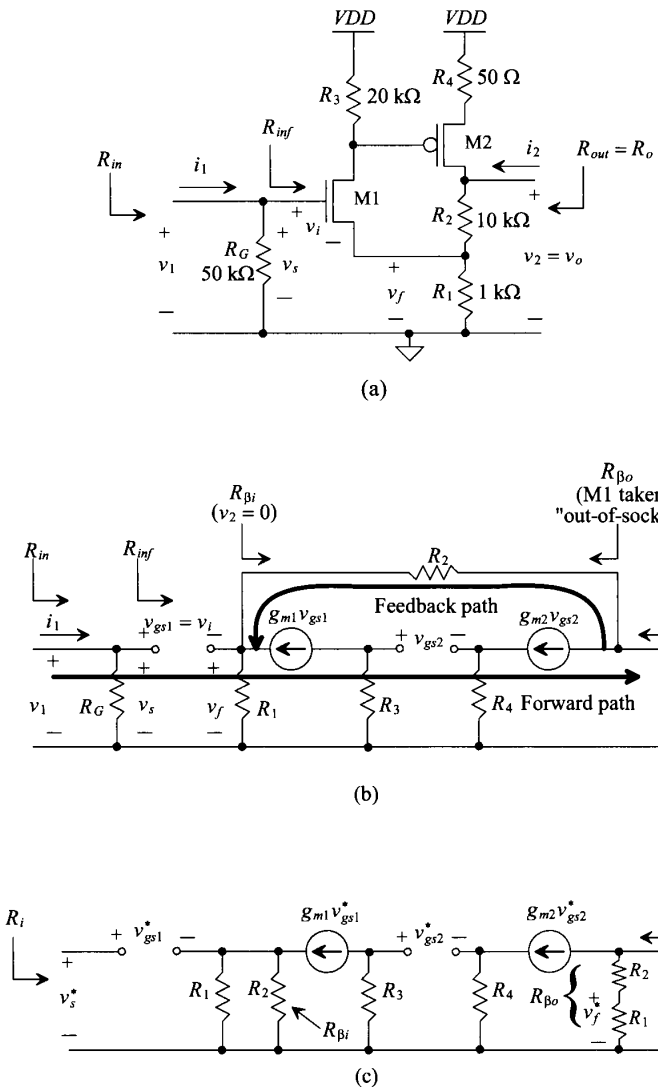


Figure 31.17 (a) Series-shunt circuit used in Ex. 31.1; (b) its closed-loop small-signal model; and (c) the resulting open-loop model.

IPR2023-00698
Theta EX2019

The open-loop gain, A_{OL}, can be found simply as the gain of two common source amplifiers with source resistance,

$$A_{OL} = \frac{v_2^*}{v_s^*} = \frac{v_2^*}{v_{g2}^*} \cdot \frac{v_{g2}^*}{v_s^*} = \left[\frac{-g_{m2}(R_2 + R_1)}{1 + g_{m2}R_4}\right]\left[\frac{-g_{m1}R_3}{1 + g_{m1}(R_1 \| R_2)}\right] = 109.8 \text{ V/V}$$

The value of β is always the gain from the output to the feedback variable. In this case, a simple voltage divider relationship exists such that

$$\beta = \frac{v_f^*}{v_2^*} = \frac{R_1}{R_1 + R_2} = 0.0909 \text{ V/V}$$

Again, notice that the product of $A_{OL}\beta$ is positive and unitless. Now that the open-loop values are calculated, the closed-loop values can be found by using Eqs. (31.38) - (31.40).

$$A_{CL} = \frac{v_2}{v_s} = \frac{A_{OL}}{1 + A_{OL}\beta} = \frac{109.8}{1 + (109.8 \cdot 0.0909)} = 10 \text{ V/V}$$

A simple check will verify that the solution is correct if $A_{CL} \approx \frac{1}{\beta}$, which is the case.

Notice in Fig. 31.17a and Fig. 31.17b that closed-loop output impedance, R_{of} is equal to the value of R_{out}. However, the value of the closed-loop input impedance, R_{inf}, is not equal to R_{in}, since the feedback amplifier itself excludes the input source and the gate resistor, R_G. The closed-loop values, R_{inf}, and R_{of}, are also easily attained, that is,

$$R_{inf} = \infty \text{ (since } R_i = \infty)$$

and

$$R_{of} = R_{out} = \frac{v_2}{i_2} = \frac{R_o}{(1 + A_{OL}\beta)} = \frac{11 \text{ k}\Omega}{10.98} = 1.002 \ \Omega$$

The last step involves finding R_{in} and v_2/v_1. The value of R_{in} can be found by examining Fig. 31.17a as,

$$R_{in} = \frac{v_1}{i_1} = R_G \| R_{inf} = 50 \text{ k}\Omega$$

The overall gain, v_2/v_1, will be equal to the value of A_{CL} since the input voltage, v_1 is equal to v_s. If we had included a value for a source resistance, then an additional voltage divider relationship would have been needed to formulate v_2/v_1 as seen in Eq. (31.41)

$$\frac{v_2}{v_1} = A_{CL} = 10 \text{ V/V} \ \blacksquare$$

31.5 The Transimpedance Amp (Shunt-Shunt Feedback)

Shunt-shunt feedback mixes current at its input and samples voltage at the output. Consider the ideal shunt-shunt feedback amplifier shown with open-loop values, A_{OL}, β, R_i, and R_o, in Fig. 31.18. In the ideal case, $R_{\beta i}$ is infinite and $R_{\beta o}$ is zero. Notice that the basic amplifier and the β network are in parallel at both the input and the output. The β

IPR2023-00698
Theta EX2019

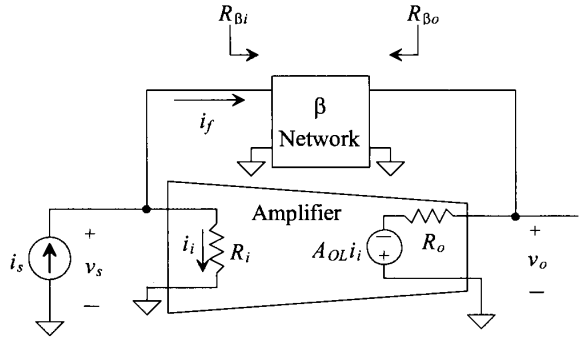


Figure 31.18 An ideal transimpedance (shunt-shunt) amplifier.

network shunts the basic amplifier; therefore, the input variable is a current. Looking into the output of the amplifier, we see that the feedback path is in parallel with or shunts the output signal. Therefore, the output signal is a voltage. Also note that since the input variable is a current and the output variable is a voltage, the basic amplifier has units of V/I, also known as a transimpedance amplifier. Since $A_{OL}\beta$ should always be unitless and positive, β will have units of I/V (mhos).

The closed-loop gain is

$$A_{CL} = \frac{v_o}{i_s} = \frac{A_{OL}}{1 + A_{OL}\beta} \tag{31.43}$$

Notice that the basic amplifier circuit is inverting, corresponding to an ideal op-amp circuit. Since A_{OL} is negative, then β must also be negative to ensure that negative feedback exists.

To calculate the input impedance of the transimpedance amplifier we apply a test current source to the input of the amplifier, such that $i_s = i_{test}$ and $v_s = v_{test}$. The test current and the input impedance are determined by

$$i_{test} = \frac{v_{test}}{R_i} + i_f = \frac{v_{test}}{R_i} + \beta v_o = \frac{v_{test}}{R_i} + \beta A_{OL} i_i = \frac{(1 + A_{OL}\beta) \cdot v_{test}}{R_i} \tag{31.44}$$

Assuming that the β network does not load the basic amplifier, we note that the closed-loop input impedance becomes

$$R_{inf} = \frac{v_{test}}{i_{test}} = \frac{R_i}{1 + A_{OL}\beta} \tag{31.45}$$

A similar analysis of the output impedance of the transimpedance amplifier shows that

$$R_{of} = \frac{R_o}{(1 + A_{OL}\beta)} \tag{31.46}$$

IPR2023-00698
Theta EX2019

The ideal transimpedance amplifier has zero input resistance and zero output resistance. We can see from Eqs. (31.45) and (31.46) that the addition of the feedback helps to make the basic amplifier appear closer to the ideal.

Now examine the transistor level shunt-shunt feedback circuit shown in Fig. 31.19. The closed-loop and open-loop small-signal models are seen in Fig. 31.20 and Fig. 31.21, respectively. The gate is attached to AC ground since it is attached to a DC source and the DC current source I_{SS} is an AC open circuit. Interestingly, since the values of r_{o1} and r_{o2} have been included, there is a feedback path through the basic amplifier. However, this gain is very small compared to the feedback path through the β network and is typically assumed to be negligible.

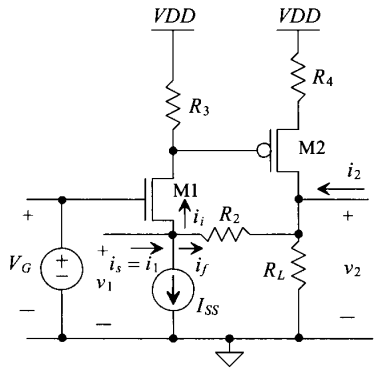


Figure 31.19 Shunt-shunt feedback amplifier.

The open-loop analysis begins by evaluating the effects of the β network on the basic amplifier using the rules stated earlier. The resistance, $R_{\beta i}$, will be the equivalent resistance seen looking into the β network from the input with the output node shorted to ground. The equivalent resistance looking into the β network from the output, $R_{\beta o}$, will be calculated with the input shorted to ground.

Since we are using shunt-shunt feedback, A_{OL} will be defined as

$$A_{OL} = \frac{v_2^*}{i_s^*} = \frac{v_2^*}{v_{g2}^*} \cdot \frac{v_{g2}^*}{v_1^*} \cdot \frac{v_1^*}{i_s^*} \tag{31.47}$$

Solving the first term in Eq. (31.47), can be quite extensive if using standard circuit analysis. A circuit technique based on two-port theory will greatly simplify the analysis. For example, the equivalent circuit for the gain, $\frac{v_2^*}{v_{g2}^*}$, can be seen in Fig. 31.22a. However, the circuit seen in Fig. 31.22b shows the equivalent circuit using an equivalent transconductance, G_M, and output resistance, R_{Leq}. The gain of the equivalent circuit, and hence the actual circuit, is

$$\frac{v_2^*}{v_{g2}^*} = G_M R_{Leq} \tag{31.48}$$

IPR2023-00698
Theta EX2019

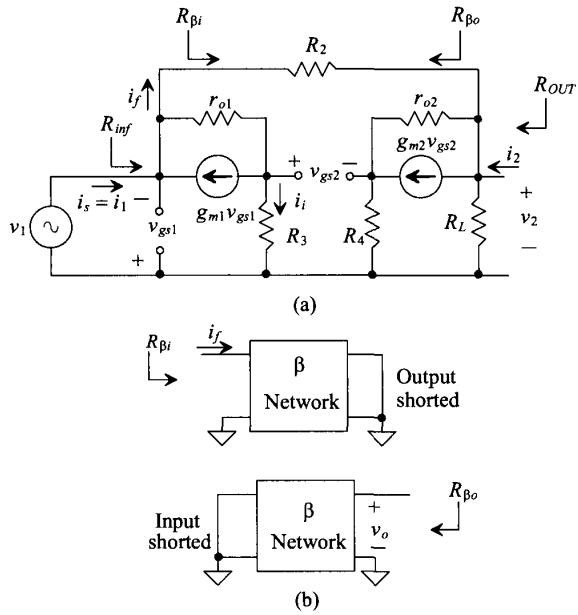


Figure 31.20 (a) Closed-loop small-signal model of Fig. 31.19 and (b) method for determining the feedback network loading.

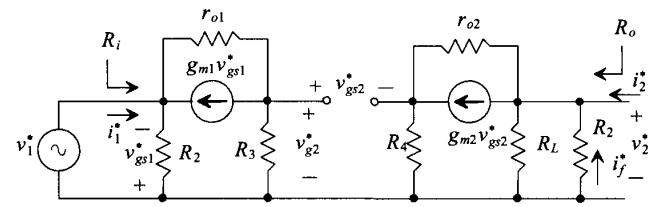


Figure 31.21 Open-loop small-signal model of Fig. 31.19.

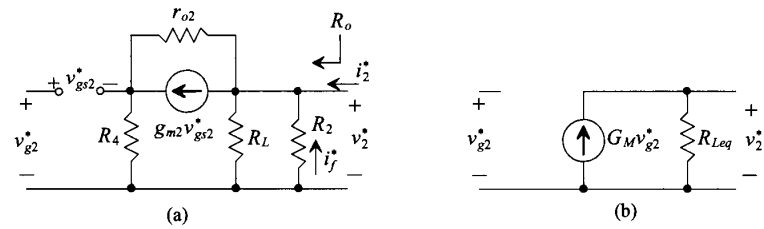


Figure 31.22 (a) Solving a portion of Fig. 31.21, including the drain-to-source resistance, and (b) the equivalent transconductance model.

IPR2023-00698
Theta EX2019

The value of R_{Leq} can easily be found as

$$R_{Leq} = R_L || R_2 || R_{inD2} \tag{31.49}$$

where R_{inD2} is the resistance seen looking into the drain of M2. From Ch. 20, we know that this resistance is

$$R_{inD2} = [(1 + g_{m2}R_4)r_{o2} + R_4] \tag{31.50}$$

The value of G_M is the short-circuit transconductance and is defined as

$$G_M = \frac{i_o^*}{v_{g2}^*} \ (R_{Leq} = 0) \tag{31.51}$$

which means that the effective transconductance can be found by shorting the equivalent load resistance, in this case $R_L||R_2$, and finding the gain from the short-circuit current to the input voltage. As seen in Fig. 31.23, the equations used to find G_M are

$$i_o^* = -g_{m2}v_{gs2} + \frac{v_{s2}}{r_{o2}} \tag{31.52}$$

$$v_{s2}^* = -i_o R_4 \tag{31.53}$$

$$v_{s2}^* + v_{gs2}^* = v_{g2}^* \tag{31.54}$$

and solving Eqs. (31.52) - (31.54) yields

$$G_M = \frac{i_o^*}{v_{g2}^*} = \frac{-g_{m2}}{1 + g_{m2}R_4 + \frac{R_4}{r_{o2}}} \tag{31.55}$$

the gain, $\frac{v_2^*}{v_{g2}^*}$, becomes

$$\frac{v_2^*}{v_{g2}^*} = \frac{-g_{m2}(R_L||R_2||[(1 + g_{m2}R_4)r_{o2} + R_4)]}{1 + g_{m2}R_4 + \frac{R_4}{r_{o2}}} \tag{31.56}$$

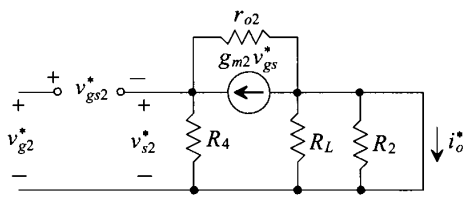


Figure 31.23 Circuit used to determine the equivalent transconductance.

Referring back to Eq. (31.47), the second factor, $\frac{v_{g2}^*}{v_i^*}$, can be found by analyzing Fig. 31.21 as

IPR2023-00698
Theta EX2019

$$\frac{v_{g2}^*}{v_1^*} = \frac{g_{m1}R_3 + \frac{R_3}{r_{o1}}}{1 + \frac{R_3}{r_{o1}}} \tag{31.57}$$

The last term in Eq. (31.47) is simply the input resistance, R_i, of the open-loop circuit shown. Using a test source, we can determine this value to be

$$R_i = \frac{v_1^*}{i_s^*} = \frac{v_t}{i_t} || R_2 = \left(\frac{1 + \frac{R_3}{r_{o1}}}{g_{m1} + \frac{1}{r_{o1}}} \right) || R_2 \tag{31.58}$$

Therefore, the entire expression for the open-loop gain, A_{OL}, becomes

$$A_{OL} = \frac{v_2^*}{i_s^*} = \frac{-g_{m2}(R_L || R_2 || ((1 + g_{m2}R_4)r_{o2} + R_4))}{1 + g_{m2}R_4 + \frac{R_4}{r_{o2}}} \cdot \frac{g_{m1}R_3 + \frac{R_3}{r_{o1}}}{1 + \frac{R_3}{r_{o1}}} \cdot \left(\frac{1 + \frac{R_3}{r_{o1}}}{g_{m1} + \frac{1}{r_{o1}}} \right) || R_2 \ \Omega \tag{31.59}$$

At first glance, this equation may appear quite daunting. However, notice that if r_{o2} is much greater than R_2, R_4, and R_L and if r_{o1} is much greater than R_3, the open-loop gain simplifies to

A_{OL} = [common source amp with source resistance][common gate amp][R_i]

$$\approx \frac{-g_{m2}(R_L || R_2)}{1 + g_{m2}R_4} \cdot g_{m1}R_3 \cdot \frac{1}{g_{m1}} || R_2 \ \Omega \tag{31.60}$$

Equation (31.60) is typically used for discrete designs in which higher values of currents are used, therefore requiring lower resistor values. However, active loads can easily be used, as seen in Fig. 31.24. Here, the resistor, r_{o3} now replaces R_3. Since R_4 is a source degeneration resistor, its value will be small. The active load that replaces R_4 is a gate-drain connected device and equal to $\frac{1}{g_{m4}} || r_{o4}$. And the resistor, R_L, is now replaced by r_{o5}. Therefore, the open-loop gain of Fig. 31.24 can be written by using Eq. (31.29) and making the proper substitutions:

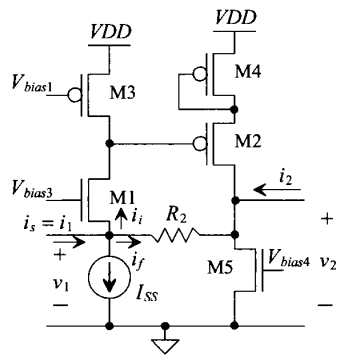


Figure 31.24 Shunt-shunt feedback amplifier using active loads.

IPR2023-00698
Theta EX2019

$$A_{OL} = \frac{-g_{m2}(r_{o5}||R_2||(1+g_{m2}(\frac{1}{gm4}||r_{o4}))r_{o2}+\frac{1}{gm4}||r_{o4}))}{1+g_{m2}\left(\frac{1}{gm4}||r_{o4}\right)+\frac{\frac{1}{gm4}||r_{o4}}{r_{o2}}} \cdot \frac{g_{m1}r_{o3}+\frac{r_{o3}}{r_{o1}}}{1+\frac{r_{o3}}{r_{o1}}}\left[\frac{1+\frac{r_{o3}}{r_{o1}}}{g_{m1}+\frac{1}{r_{o1}}}||R_2\right] \tag{31.61}$$

which can be approximated as

$$A_{OL} = \frac{v_2^*}{i_s^*} \approx \frac{-g_{m2}(R_2)}{1+g_{m2}\left(\frac{1}{gm4}\right)} \cdot \frac{1+g_{m1}r_{o3}}{2} \cdot \left(\frac{1}{g_{m1}}||R_2\right)\ \Omega \tag{31.62}$$

if it is assumed that $r_{o1} \approx r_{o3}$ and that the discrete resistor, R_2, is relatively small compared to the impedances of the active loads. An active device could have been used to substitute even the resistor, R_2. This will be discussed further later in the chapter.

Next, the value of β is calculated as

$$\beta = \frac{i_f^*}{v_2^*} = -\frac{1}{R_2}\ \text{mhos} \tag{31.63}$$

Again, note that $A_{OL}\beta$ is unitless and overall positive.

Now that all of the open-loop parameters, A_{OL}, β, R_i, and R_o, have been found, the closed-loop values are easily calculated. The value for the closed-loop gain is

$$A_{CL} = \frac{v_2}{i_s} = \frac{A_{OL}}{1+A_{OL}\beta} \tag{31.64}$$

The closed-loop input impedance is

$$R_{inf} = \frac{R_i}{1+A_{OL}\beta} \tag{31.65}$$

and the value for the closed-loop output impedance is

$$R_{of} = R_{out} = \frac{R_o}{1+A_{OL}\beta} \tag{31.66}$$

We should be able to see a trend in the feedback effects. Shunt mixing and shunt sampling cause the input and output impedances to decrease by the factor $(1 + A_{OL}\beta)$. Similarly, series input mixing and series sampling cause the input and output impedances to increase by the factor $(1 + A_{OL}\beta)$.

In many cases, the overall gain is expressed in terms of a voltage gain, $\frac{v_2}{v_1}$. Since we have calculated the transfer function in terms of the current, i_s, we can easily express this in terms of v_1 by

$$\frac{v_2}{v_1} = \frac{v_2}{i_s(R_{inf})} = A_{CL}\left(\frac{1}{R_{inf}}\right) \tag{31.67}$$

31.5.1 Simple Feedback Using a Gate-Drain Resistor

One of the most popular, and simplest, examples of shunt-shunt feedback is seen in Fig. 31.25 and consists of a simple inverting amplifier with a resistor connecting the gate and the drain. The feedback resistor, R_2, is usually a large value and serves several important

IPR2023-00698
Theta EX2019

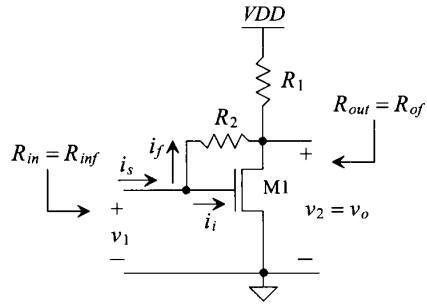


Figure 31.25 Shunt-shunt feedback using a simple gate-drain connected device.

functions. When analyzing the DC characteristics of this circuit, the voltage at the drain will be equal to the voltage on the gate, since there is no DC current flowing through R_2 (when the input is AC coupled). This ensures that the device is always in saturation and provides biasing with no other components needed. We will examine the effects of R_2 on the AC response of the amplifier and soon discover that it has little effect on the gain (if it is large) of the amplifier as well.

Figure 31.26a and b shows the small-signal models for the closed and open-loop circuits, respectively. Since the feedback resistor sums current at the gate of M1, the mixing circuit is shunt. And since the feedback is taken off of the same node as the output, the sampling circuit is shunt. Therefore, the value of $R_{\beta i}$ for the open-loop model is found by shorting the output and looking into the feedback loop from the input, and the value of $R_{\beta o}$ is found by shorting the input and looking into the feedback loop from the output.

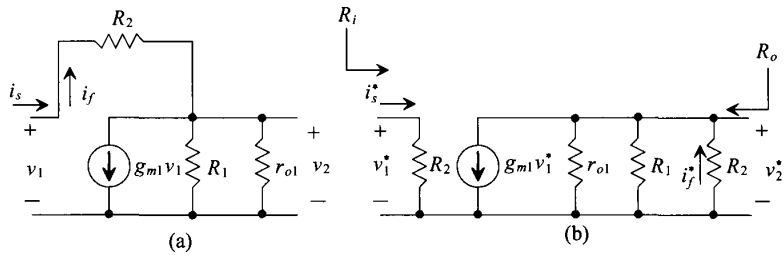


Figure 31.26 Small-signal model of Fig. 31.25: (a) the closed-loop model and (b) the open-loop model.

The open-loop values can be calculated as

$$A_{OL} = \frac{v_2^*}{i_s^*} = \frac{v_2^*}{v_1^*} \cdot \frac{v_1^*}{i_s^*} = [-g_{m1}(R_1 || R_2 || r_{o1})][R_2] \text{ V/A} \tag{31.68}$$

$$R_i = R_2 \tag{31.69}$$

$$R_o = R_1 || R_2 || r_{o1} \tag{31.70}$$

IPR2023-00698
Theta EX2019

and

$$\beta = \frac{i_f^*}{v_2^*} = -\frac{1}{R_2} \tag{31.71}$$

Using Eqs. (31.43) - (31.46), we find that the closed-loop values become

$$A_{CL} = \frac{v_2}{i_s} = \frac{A_{OL}}{1 + A_{OL}\beta} = \frac{-g_{m1}R_oR_2}{1 + g_{m1}R_oR_2\frac{1}{R_2}} \tag{31.72}$$

$$R_{inf} = \frac{v_1}{i_s} = \frac{R_2}{1 + g_{m1}R_oR_2\frac{1}{R_2}} \text{ and } R_{of} = \frac{R_o}{1 + g_{m1}R_oR_2\frac{1}{R_2}} \tag{31.73}$$

The value of R_{inf} is also equal to the value of R_{in} since no source resistance is associated with v_1. Notice that the value of the closed-loop gain is dependent on R_2. However, most applications of this amplifier use voltage as the input variable. Therefore, the value of the overall voltage gain becomes

$$\frac{v_2}{v_1} = \frac{v_2}{i_s} \cdot \frac{i_s}{v_1} = A_{CL} \cdot \frac{1}{R_{inf}} = -g_{m1}R_o = -g_{m1}(R_1 \| R_2 \| r_{o1}) \tag{31.74}$$

and if the value of R_2 is chosen to be much larger than R_1 then its effect on the AC midband gain is minimized.

Example 31.2

Calculate the gain, $\frac{v_o}{v_s}$, of the shunt-shunt amplifier in Fig. 31.27 assuming that A = 500,000 V/A, R_i = 10 Ω, and R_o = 10 Ω. Notice that the circuit is similar to a simple inverting op-amp.

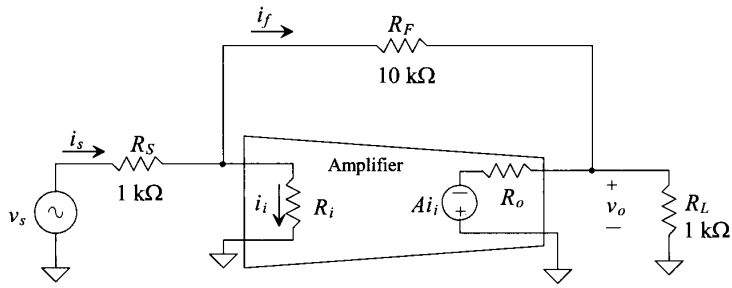


Figure 31.27 A transimpedance amplifier example.

First, we will source transform the voltage source to a current source, since the input mixing is shunt. The model of the amplifier is in terms of A, not A_{OL}, since we must now include the loading of the feedback resistor, and R_L and R_S in the calculation of A_{OL}. The basic amplifier circuit, a transimpedance amplifier with units of V/I, has an ideal input and output impedance of 0 Ω. Next, we will determine the loading of the β network consisting of R_F on the basic amplifier as seen in Fig. 31.28. A_{OL} can be determined as

IPR2023-00698
Theta EX2019

$$A_{OL} = \frac{v_o^*}{i_s^*} = -A \cdot \left(\frac{R_F || R_L}{R_F || R_L + R_o}\right) \cdot \left(\frac{R_s || R_F}{R_s || R_F + R_i}\right) \Omega$$

$$= -500,000 \cdot 0.989 \cdot 0.989 = -489,060 \text{ V/A}$$

The value of β is easily calculated as

$$\beta = \frac{i_f^*}{v_o^*} = -\frac{1}{R_F} = -0.0001 \text{ A/V}$$

And A_{CL} becomes

$$A_{CL} = \frac{v_o}{i_s} = \frac{-489,060}{1 + -489,060 \cdot -0.0001} = -9.8 \text{k}\Omega \approx R_F$$

Since $v_s = i_s \cdot R_s$, the overall voltage gain is given by

$$\frac{v_o}{v_s} = \frac{v_o}{i_s} \cdot \frac{1}{R_s} = A_{CL} \cdot \frac{1}{R_s} = -\frac{9.8\text{k}}{1\text{k}} = -9.8 \text{ V/V} \approx -\frac{R_F}{R_s}$$

Notice that this is the gain of the standard inverting op-amp configuration. ■

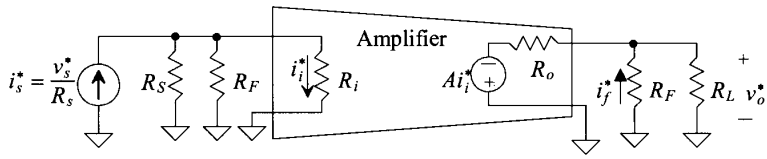


Figure 31.28 Open-loop transimpedance amplifier used in Example 31.2.

31.6 The Transconductance Amp (Series-Series Feedback)

A series-series feedback amp with open-loop values is shown in Fig. 31.29. Ideally, the values of $R_{\beta i}$ and $R_{\beta o}$ are zero. The feedback resistor R_F is in series with the input and the output of the amplifier. Therefore, the input of the amplifier is a voltage, and the output is a current. The units of A_{OL} will be I/V (a transconductance), and the units of β will be V/I (ohms). Transconductance amplifiers also have high-input and high-output impedance.

The closed-loop gain of the transconductance amplifier is

$$A_{CL} = \frac{i_o}{v_s} = \frac{A_{OL}}{1 + A_{OL}\beta} \text{ A/V} \qquad (31.75)$$

The input impedance is again given by applying a test voltage to the input of the feedback amp and calculating the current that flows into the input of the amplifier. Also, the assumption that the feedback network does not load the amplifier will be used. Setting $v_{test} = v_s$, $i_{test} = i_s$ and writing a loop at the input of the amplifier give

$$v_{test} = i_{test} \cdot R_i + v_f = i_{test} \cdot R_i + \beta \cdot A_{OL} \cdot i_{test} \cdot R_i \qquad (31.76)$$

and so the input resistance with feedback is now

IPR2023-00698
Theta EX2019

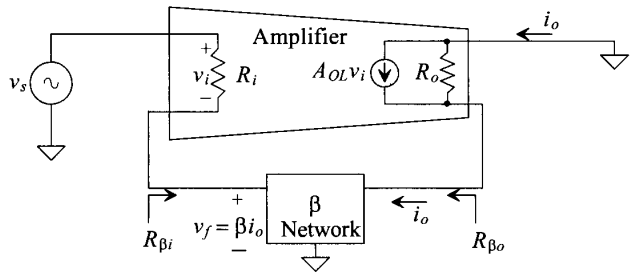


Figure 31.29 An ideal transconductance amplifier.

$$R_{inf} = \frac{v_{test}}{i_{test}} = R_i \cdot (1 + A_{OL}\beta) \text{ for } R_o \to \infty \tag{31.77}$$

keeping in mind that $i_{test} \cdot R_i = v_i$ and $A_{OL} \cdot v_i = i_o$. The output resistance is determined by applying a test current at the output with the input source shorted, and is given by

$$v_{test} = (i_{test} - A_{OL}v_i) \cdot R_o = [i_{test} - A_{OL} \cdot (-\beta i_{test})] \cdot R_o \tag{31.78}$$

since $i_{test} = i_{out}$, $v_i = -v_f$, and $R_{\beta o} = 0$. The output resistance is given by

$$R_{of} = \frac{v_{test}}{i_{test}} = R_o \cdot (1 + A_{OL}\beta) \tag{31.79}$$

The ideal transconductance amplifier has infinite output and input resistance. Again, feedback helps to make the amplifier appear closer to the ideal.

Now examine the transistor level circuit in Fig. 31.30. Notice the similarity to Fig. 31.13, the circuit used in the series-shunt example; the only difference is the location of the output connection. The feedback loops in both circuits are identical, so the feedback in Fig. 31.30 is known to be negative.

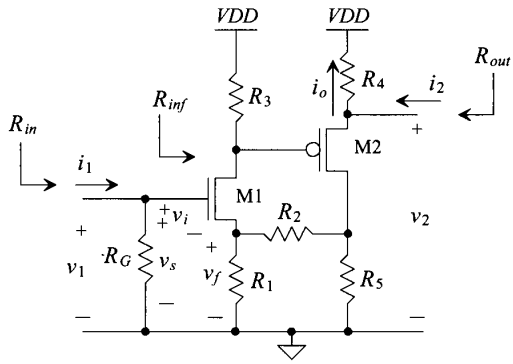


Figure 31.30 Transistor-level series-series feedback amplifier.

IPR2023-00698
Theta EX2019

Since the output and the feedback are connected to two separate terminals of the output device, the output variable is a current, sampling i_o. The small-signal model for this circuit is shown in Fig. 31.31 with the open-loop, small-signal model shown in Fig. 31.32. Since the output sampling is a current, loading of the β network will be slightly different from that of the series-shunt example. The input utilizes series mixing; therefore the loading of the β network on the output will be identical to the series-shunt example discussed previously ($R_{\beta o} = R_1 + R_2$). However, since the output sampling is series, the equivalent resistance, $R_{\beta i}$, will be the resistance seen looking into the β network from the input, with the output device taken "out-of-socket" and $R_{\beta i} = R_2 + R_5$.

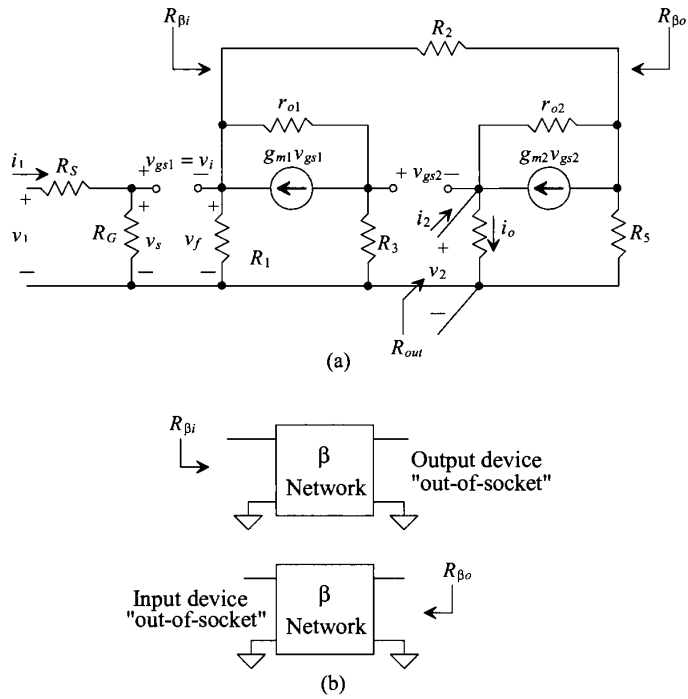


Figure 31.31 (a) Closed-loop small-signal model of Fig. 31.30 and (b) method for determining feedback loading.

Once the open-loop model has been constructed, A_{OL} can be calculated as

$$A_{OL} = \frac{i_o^*}{v_s^*} = \frac{i_o^*}{v_{g2}^*} \cdot \frac{v_{g2}^*}{v_s^*} \tag{31.80}$$

IPR2023-00698
Theta EX2019

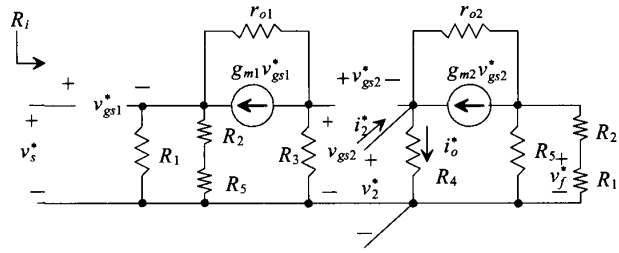


Figure 31.32 Open-loop small-signal model of Fig. 31.30.

where the term, $\frac{i_o^*}{v_{g2}^*}$, can be determined by using straightforward circuit analysis to solve $\frac{v_2^*}{v_{g2}^*}$ and then dividing the result by R_4,

$$\frac{i_o^*}{v_{g2}^*} = \frac{g_{m2}}{1 + g_{m2}R_4 + \frac{R_4 + R_5 || (R_2 + R_1)}{r_{o2}}} \tag{31.81}$$

The term, $\frac{v_{g2}^*}{v_s^*}$, is found by using the G_M method presented in the previous section on shunt-shunt feedback and is

$$\frac{v_{g2}}{v_s^*} = \frac{-g_{m1}(R_3 || [(1 + g_{m1}R_A)r_{o1} + R_A)]}{1 + g_{m1}R_A + \frac{R_A}{r_{o1}}} \text{ mhos} \tag{31.82}$$

where $R_A = R_1 || (R_2 + R_5)$. The feedback factor, β, is

$$\beta = \frac{v_f^*}{i_o^*} \approx \frac{-R_5 R_1}{R_5 + R_1 + R_2} \Omega \tag{31.83}$$

And the closed-loop gain is simply

$$A_{CL} = \frac{i_o}{v_s} = \frac{A_{OL}}{1 + A_{OL}\beta} \text{mhos} \tag{31.84}$$

The value of R_i is obviously infinite, resulting in an identical value of R_{inf}. Therefore, $R_{in} = R_{inf} || R_G = R_G$.

Calculating R_o for a series output requires some explanation. Examine Fig. 31.33. The value of R_o is the value seen looking in series with the load resistor. In this case, the value of R_o becomes

$$R_o = R_4 + \frac{\frac{R_B}{r_{o2}} + 1}{\frac{1}{r_{o2}} + g_{m2}} \approx R_4 + \frac{1}{g_{m2}} \tag{31.85}$$

where $R_B = R_5 || (R_1 + R_2)$ and the closed-loop value becomes

$$R_{of} = R_o(1 + A_{OL}\beta) \tag{31.86}$$

IPR2023-00698
Theta EX2019

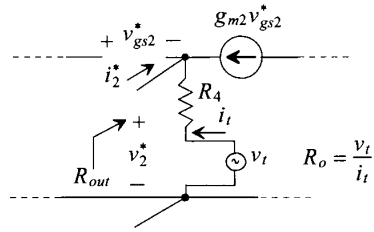


Figure 31.33 Calculation of the output impedance for the circuit in Fig. 31.30.

Notice, however, that R_{of} is not the same as R_{out}, in this case. Typically, R_{out} is designated as the resistance in parallel with the load. Taking the resistance in series with the load is not a practical specification. Therefore, the resistance R_{out} can be described as seen in Fig. 31.34. In part (a), it can be seen that $R_{of} = R_o(1+A_{OL}\beta)$ and that $R'_{of} = R_{of} - R_4$. If we want to find a value for R_{out}, using Fig. 31.34b, R_{out} is simply

$$R_{out} = R_4 || R'_{of} = R_4 || (R_{of} - R_4) \tag{31.87}$$

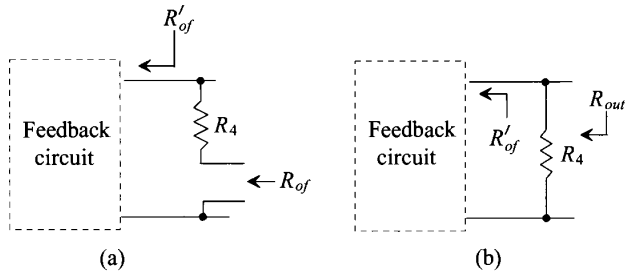


Figure 31.34 Determining the output resistance of a series sampling circuit.

31.7 The Current Amplifier (Shunt-Series Feedback)

The last feedback topology to be discussed is the shunt-series feedback amplifier, also known as a current amplifier. As can be expected, both A_{OL} and β have units of I/I, and we can expect the input impedance to be very low and the output impedance very high. Figure 31.35 illustrates the ideal shunt-series amplifier with open-loop values included. Based on past derivations, we can expect that

$$R_{inf} = \frac{R_i}{(1 + A_{OL}\beta)} \tag{31.88}$$

and R_{of} to be

$$R_{of} = R_o(1 + A_{OL}\beta) \tag{31.89}$$

The derivations of this topology will be left to the reader in the Problems section.

IPR2023-00698
Theta EX2019

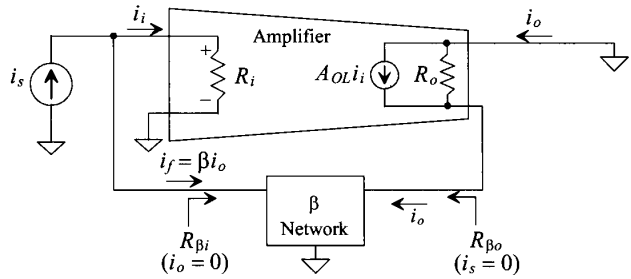


Figure 31.35 An ideal current feedback amplifier.

The transistor-level circuit shown in Fig. 31.36 is similar to the shunt-shunt topology, except for the placement of the output signal. The reader will also be asked to analyze the shunt-series circuit in the Problems section.

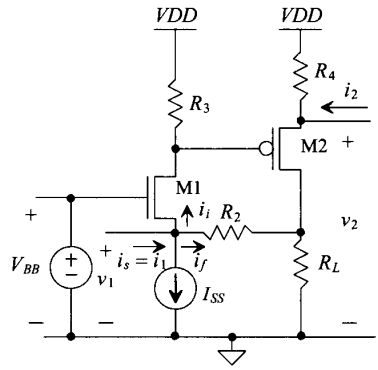


Figure 31.36 A shunt-series feedback amplifier.

Example 31.3

Examine the current amplifier seen in Fig. 31.37a. Using feedback analysis, draw open-loop small-signal models and derive values for A_{OL}, β, R_i, R_o, R_{out}, and the overall voltage gain, $\frac{v_2}{v_1}$.

Notice that although this circuit is similiar to the cascode current sink, we can also use it, though unconventionally, as a voltage op-amp. The analysis begins by identifying the feedback circuit. The current summation that occurs at the gate of M4 indicates that shunt mixing is utilized. Since the output and the feedback are taken off separate terminals of the output device, the output sampling is series. The open-loop model is seen in Fig. 31.37b. The only loading on the input due to the feedback network is r_{o2}, since M4 can be taken "out-of-socket" ($v_{gs2} = 0$). The

IPR2023-00698
Theta EX2019

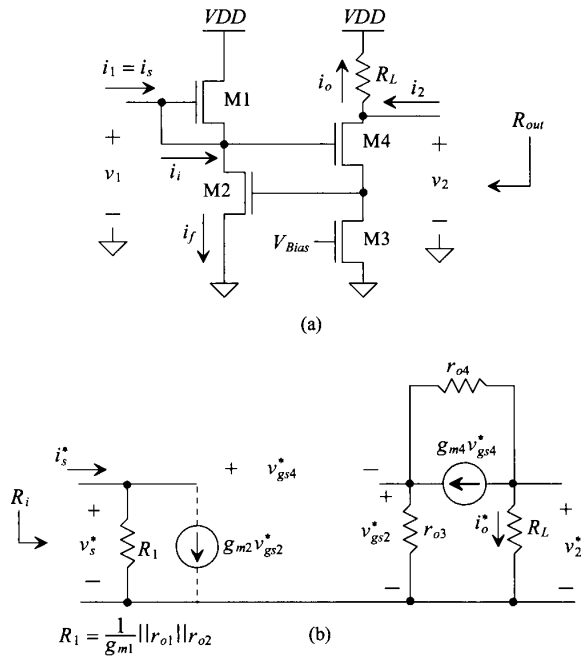


Figure 31.37 (a) Circuit used in Ex. 31.3 and (b) the open-loop model.

feedback does not load the output at all. However, we must remember to include the controlled source for determining i_f. The dependent source, $g_{m2}v_{gs}$, will not be included in calculating A_{OL} but will be needed for calculating β. Solving this open-loop circuit for A_{OL} yields (see Problem 31.28):

$$A_{OL} = \frac{i_o^*}{i_s^*} = \frac{v_2^*}{v_s^*} \cdot \frac{R_1}{R_L} = \frac{-g_{m4}(R_L || ((1+g_{m4}r_{o3})r_{o4} + r_{o3}))}{1 + g_{m4}r_{o3} + \frac{r_{o3}}{r_{o4}}} \cdot \frac{R_1}{R_L} \approx \frac{-g_{m4}R_1}{1 + g_{m4}r_{o3} + \frac{r_{o3}}{r_{o4}}} \text{ A/A}$$

The output, R_o, is found by the same manner presented in the discussion of series-series amplifiers and is

$$R_o = R_L + (1 + g_{m4}r_{o3})r_{o4} + r_{o3}$$

Similiarly, the value of R_{out} becomes

$$R_{out} = (R_o(1 + A_{OL}\beta) - R_L) || R_L$$

and R_i is simply R_1, with

$$R_{if} = \frac{R_1}{1 + A_{OL}\beta}$$

IPR2023-00698
Theta EX2019

The feedback variable is $i_f^* \approx g_{m2} v_{gs2}^* = -g_{m2} i_o^* r_{o3}$, and the value of β can be calculated as

$$\beta = \frac{i_f^*}{i_o^*} = -g_{m2} r_{o3} \text{ A/A}$$

The overall voltage gain is

$$\frac{v_2}{v_1} = \frac{i_o \cdot R_L}{i_s \cdot R_{if}} = \frac{A_{OL}}{1 + A_{OL}\beta} \cdot \frac{R_L}{R_{if}} \text{ V/V} \blacksquare$$

31.8 Stability

The previous sections illustrated the benefits of feedback and the corresponding tradeoff in gain. However, a critical concern must be examined when applying negative feedback. Some circuits will cause a phase shift in the input signal large enough that the feedback becomes positive (the output adds to the original input), resulting in an unstable system. The occurrence of instability can be minimized with some careful analysis of both the open-loop amplifier, A_{OL}, and the feedback network, β.

The *loop gain* is defined as

$$T = A_{OL}\beta \tag{31.90}$$

Remember that the product of $A_{OL}\beta$ must itself always be positive. By examining the frequency response of the loop gain, T, the overall stability of the system can be determined. The stability of the system can be summarized with the following rules and is illustrated in Fig. 31.38.

Case 1: If the change in the phase of $A_{OL}\beta$ is equal to 180° and the magnitude is below 0 dB, the system will be stable.

Case 2: If the change in the phase of $A_{OL}\beta$ is equal to 180° and the magnitude equals 0 dB, the system may or may not be stable.

Case 3: If the change in the phase of $A_{OL}\beta$ is equal to 180° and the magnitude is above 0 dB, the system will be unstable.

One might wonder why the 180° phase shift and the magnitude of 0 dB (gain of 1) are such critical factors. The answer may be better understood with a qualitative rather than a quantitative explanation. In order for positive feedback to occur, the output must be added back to its original input signal. The PA system example illustrates this concept. If the loudspeaker, which represents the output of the system, is added back to the input (the microphone), the system becomes unstable because the amplifier is attempting to amplify its own output. The result is a loud, high-pitched, ringing sound, which most people have unfortunately experienced.

In negative feedback applications, the addition of the feedback signal to the original input occurs because of the additional phase shift introduced by the frequency dependent components within A_{OL} and β. Because the feedback signal is already inverted with respect to the input, an additional 180° of phase shift will cause the feedback signal to be positive with respect to the input. It is this second 180° degrees of phase shift that becomes the important concern. Thus, the frequency response of the loop gain must be

IPR2023-00698
Theta EX2019

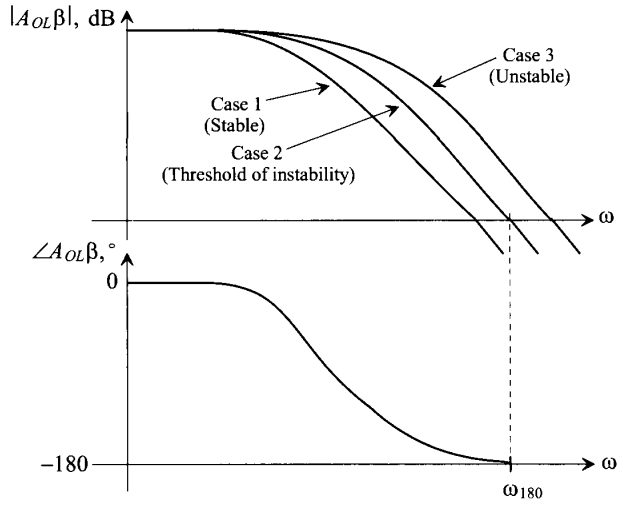


Figure 31.38 Stability analysis using the frequency response of the loop gain.

examined. The magnitude of 0 dB is also important. If the gain around the loop is less than 1, the output settles to a stable value. However, if the gain around the loop is greater than 1, the amplifier output grows and becomes unstable quickly.

Assume that A_{OL} can be described with the following frequency response:

$$A_{OL}(s) = -\frac{10}{\left(1+\frac{s}{10}\right)^2} \tag{31.91}$$

Two poles exist at $\omega = 10$ rad/s with a gain at DC equal to −10 V/V. Also assume that β is frequency dependent and has a single pole at $\omega = 10$ rad/s:

$$\beta = \frac{-1}{\left(\frac{s}{10}+1\right)} \tag{31.92}$$

The loop gain then, is

$$A_{OL}\beta = \frac{10}{\left(1+\frac{s}{10}\right)^2} \cdot \frac{1}{\left(\frac{s}{10}+1\right)} \tag{31.93}$$

The Bode plot of the loop gain can be seen in Fig. 31.39. It can be seen that since the phase plot crosses 180° slightly before the magnitude plot crosses 0 dB, the system is unstable.

IPR2023-00698
Theta EX2019

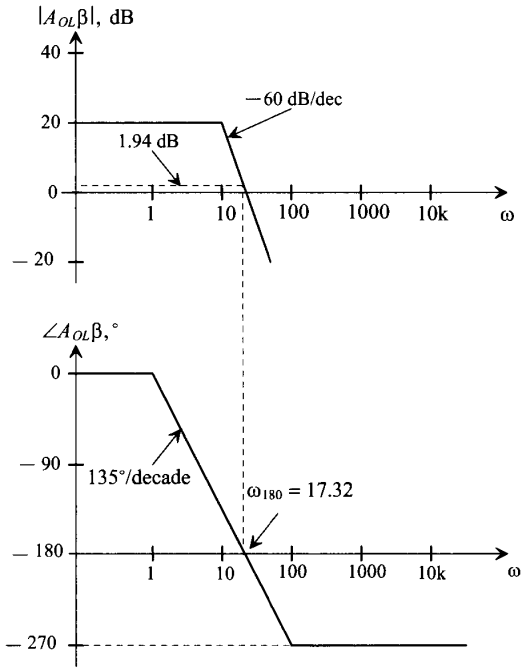


Figure 31.39 Stability analysis using the frequency response of the loop gain.

A more precise method of performing the same analysis will now be described. We can use the exact formulas to solve for the exact frequency which the phase plot crosses 180°. The phase of the loop gain, by definition, is

$$Arg[A_{OL}(j\omega)\beta(j\omega)] = \tan^{-1}\left(\frac{\omega}{z_1}\right) + \tan^{-1}\left(\frac{\omega}{z_2}\right) + ... + \tan^{-1}\left(\frac{\omega}{z_n}\right) - \tan^{-1}\left(\frac{\omega}{p_1}\right) - \tan^{-1}\left(\frac{\omega}{p_2}\right) - .... - \tan^{-1}\left(\frac{\omega}{p_n}\right) \quad (31.94)$$

where $z_1, z_2 \ldots z_n$ are the zeros in the system and $p_1, p_2 \ldots p_n$ are the poles. Since our example has three poles at the same frequency and no zeros, the phase response of the loop gain can be expressed as

$$Arg[A_{OL}(j\omega)\beta(j\omega)] = -3\tan^{-1}\left(\frac{\omega_{180}}{10}\right) = -180 \quad (31.95)$$

where ω_{180} is the frequency at which the phase response is equal to −180°. Solving for ω_{180} yields

$$\omega_{180} = 17.32 \text{ rad/s} \quad (31.96)$$

IPR2023-00698
Theta EX2019

The magnitude of the loop gain is defined as,

$$20\text{Log}|A_{OL}(j\omega)\beta(j\omega)| = 20\text{Log}(A_o) + 20\text{Log}\sqrt{\left(\frac{\omega}{z_1}\right)^2 + 1} + 20\text{Log}\sqrt{\left(\frac{\omega}{z_2}\right)^2 + 1} + \ldots$$
$$+ 20\text{Log}\sqrt{\left(\frac{\omega}{z_n}\right)^2 + 1} - 20\text{Log}\sqrt{\left(\frac{\omega}{p_1}\right)^2 + 1} - \ldots - 20\text{Log}\sqrt{\left(\frac{\omega}{p_n}\right)^2 + 1} \qquad (31.97)$$

where, $z_1, z_2 \ldots z_n$ are the zeros of the system, $p_1, p_2, \ldots p_n$ are the poles, and A_o is the midband gain. Plugging in the value for ω_{180} into Eq. (31.97), we can solve for the magnitude of the loop gain at the point at which its phase is –180°:

$$20\text{Log}|A_{OL}(j\omega)\beta(j\omega)| = 20\text{Log}(10) - 3\left(20\text{Log}\sqrt{\left(\frac{17.32}{10}\right)^2 + 1}\right) \qquad (31.98)$$

$$|A_{OL}(j\omega_{180})\beta(j\omega_{180})| = 1.94 \text{ dB} \qquad (31.99)$$

which, using rule 3, verifies that the system is unstable.

As discussed in Ch. 24 stability is typically measured using two specifications: gain margin and phase margin. Gain margin is defined as the difference between the magnitude of $A_{OL}\beta$ at ω_{180} and unity, whereas phase margin is defined as the difference between the value of the phase at the frequency at which the magnitude of $A_{OL}\beta$ is equal to unity and ω_{180}. Figure 31.40 illustrates both definitions. It should be noted that phase margin is the typical specification used for stability. Amplifiers should be designed to have a phase margin of at least 45°, though 60° phase margin is more acceptable.

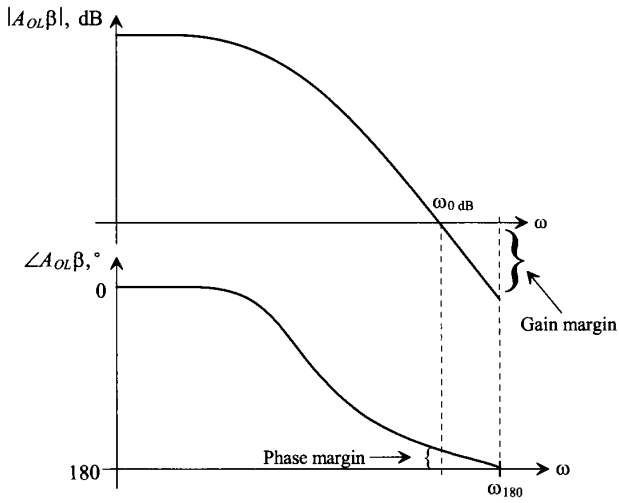


Figure 31.40 Gain margin and phase margin definitions.

IPR2023-00698
Theta EX2019

In our previous discussion, only the frequency domain analysis was considered. The effect of phase margin in the time domain is related to settling time. As the phase margin increases, less time is required for the signal to settle. As the phase margin approaches 0°, the signal will oscillate indefinitely.

31.8.1 The Return Ratio

In some cases, it may be more practical to determine the loop gain from a system point of view. One method used to determine a good approximation of the loop-gain frequency response is to break the loop, input a test signal, and determine the returned value back to the point that the loop was broken. This method is called the return ratio (*RR*) method. Consider the block diagram of the single-loop structure in Fig. 31.41 with the loop "open." The gain around the loop is,

$$RR = \frac{x_f}{x_i} = -T = -A_{OL}\beta \tag{31.100}$$

This gain represents the path from the input back around through the feedback network. It should be noted that the *RR* method may vary quite extensively from the two-port method in finding the value of $A_{OL}\beta$, but for purposes of plotting the loop gain frequency response, the *RR* should be sufficient [3].

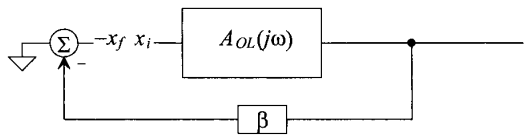


Figure 31.41 Determining the loop gain by opening the feedback loop.

The *RR* is found by first replacing all independent sources with their ideal impedances. Next, a dependent source is chosen, and the feedback loop is broken between the chosen source and the rest of the circuit. An independent test source is inserted at the node where the dependent source resided and a test signal is injected into the loop. The *RR* is then the ratio of the returned signal (which now appears across the dependent source) and the test signal, such that

$$RR = -\frac{v_r}{v_t} \tag{31.101}$$

An example will illustrate this method further.

Example 31.4

Determine the *RR* for the series-shunt op-amp circuit shown in Fig. 31.42a and compare that value to the value $A_{OL}\beta$ using the two-port method. Assume that the op-amp can be modeled as shown in Fig. 31.42b.

Since there is only one dependent source in the circuit, deciding where to break the loop is a simple endeavor. Next, the dependent source is separated from the rest of the circuit, and an independent source, v_t, is put in its original position as

IPR2023-00698
Theta EX2019

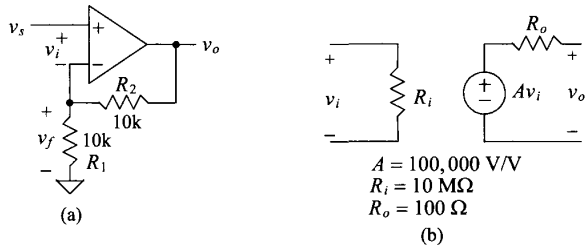


Figure 31.42 (a) A series-shunt amplifier used in Ex. 31.4 and (b) the model used for the op-amp.

seen in Fig. 31.43a. The returned signal is now across the dependent source. By inspection, the value of the *RR* can be found to be

$$RR = -\frac{v_r}{v_t} = A \cdot \left(\frac{R_I || R_1}{R_I || R_1 + R_2}\right) \cdot \left(\frac{R_2 + R_I || R_1}{R_2 + R_I || R_1 + R_o}\right) = 49,726$$

Next, the open-loop model for the original circuit is found using the two-port analysis presented in Sec. 31.2. The circuit, taking the β network loading effects into account, can be seen in Fig. 31.43b. The value of A_{OL} is easily calculated as

$$A_{OL} = \frac{v_o^*}{v_s^*} = A \cdot \left(\frac{R_1 + R_2}{R_o + R_1 + R_2}\right) \cdot \left(\frac{R_I}{R_1 || R_2 + R_I}\right) = 99,497$$

The value of β by inspection is

$$\beta = \frac{v_f^*}{v_o^*} = \frac{R_1}{R_1 + R_2} = 0.5$$

and the value of $A_{OL}\beta$ is

$$A_{OL}\beta = (99,497)(0.5) = 49,748$$

The two methods yielded very similiar answers. ■

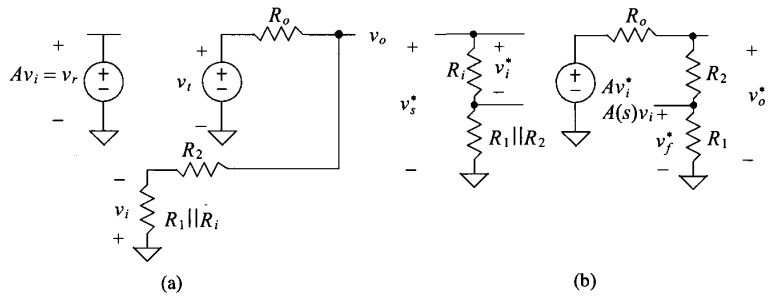


Figure 31.43 (a) The model used to calculate return ratio and (b) the model used for the two-port analysis of loop gain for Fig. 31.42.

IPR2023-00698
Theta EX2019

If $A_{OL}\beta$ is frequency dependent, then the phase and gain margin of the circuit can be plotted so as to determine the phase and gain margin. Now suppose that the circuit used in Ex. 31.4 was frequency dependent as seen in Fig. 31.44. Using the same strategy presented in Ex. 31.4 and using the *RR* method, the value of the loop gain becomes

$$RR = -\frac{v_r}{v_t} = \frac{49,726}{(s/200+1)^3} \tag{31.102}$$

The gain and phase margin can then be analyzed to determine if the system is stable (see Problem 31.34).

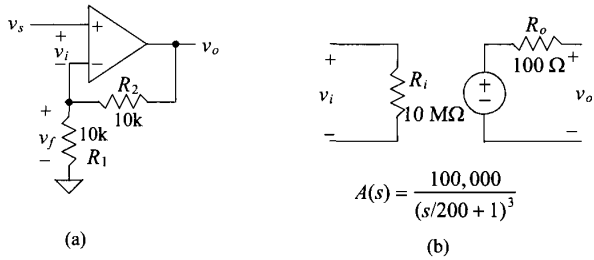


Figure 31.44 (a) A series-shunt amplifier used in Ex. 31.4 and (b) the model used for the op-amp.

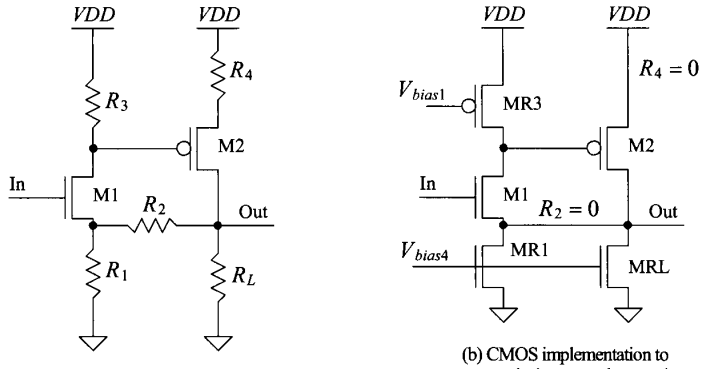
31.9 Design Examples

In this section we'll provide some design examples to further demonstrate the design of feedback amplifiers in CMOS technology. For the simulations we'll use the short-channel CMOS process with the sizes and biasing conditions seen in Table 9.2.

31.9.1 Voltage Amplifiers

Figure 31.13 showed the detailed transistor-level implementation of a series-shunt feedback amplifier (voltage in and out) using an AC coupled input. Figure 31.45a shows the implementation with a DC coupled input. Reviewing the equation for A_{OL}, Eq. (31.35), corresponding to this topology we see that to maximize the gain we can replace R_L, R_1, and R_3 with transistors MRL, MR1, and MR3 (so these resistances effectively become r_{oRL}, r_{oR1}, and r_{oR3}) and make R_2 and R_4 zero ohms. The result is seen in Fig. 31.45b. With this selection β becomes 1, Eq. (31.36), while the ideal closed gain, A_{CL}, is also one, Eq. (31.38). Note that we can further increase A_{OL} by using cascode structures. However, A_{OL} isn't the limiting performance factor in this design. The body effect of M1, as we shall shortly see, is the limiting factor. This circuit is useful as a unity voltage buffer providing large input impedance (very little capacitance) and low output resistance. It can source a relatively large current via M2 but the amount of current it can sink is limited to the biasing currents of MR1 and MRL. Note that the allowable input signal range is from $V_{GS} + V_{DS,sat}$ (= 400 mV from Table 9.2) up to $VDD - V_{SD,sat} + V_{THN}$ (= 770 mV again from Table 9.2). The output can swing from $VDD - V_{SD,sat}$ to $V_{DS,sat}$.

IPR2023-00698
Theta EX2019



(a) DC-coupled series-shunt (voltage amplifier) from Fig. 31.13.

(b) CMOS implementation to maximize open-loop gain. The topology has a closed-loop gain of 1.

Figure 31.45 (a) Series-shunt amplifier and (b) replacing resistors with MOSFETs.

Let's compare the unity buffer seen in Fig. 31.45b to a simple source-follower like the one seen in Fig. 21.39 each driving a 10 pF load. The simulated frequency responses of the amplifiers are seen in Fig. 31.46. Note the increase in bandwidth using the feedback amplifier. Also note the peaking in the series-shunt amplifier's response due to the output impedance becoming inductive at higher frequencies and forming a second-order response with the load capacitance. While the significant enhancement in bandwidth is important let's focus on why the closed-loop gain of the feedback amplifier isn't closer to 1 (0 dB).

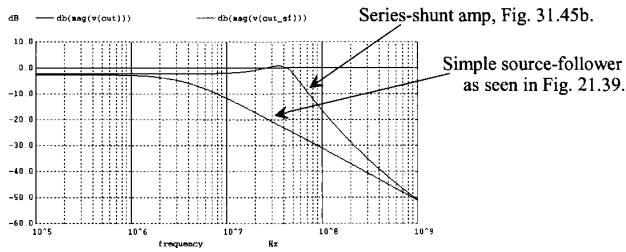
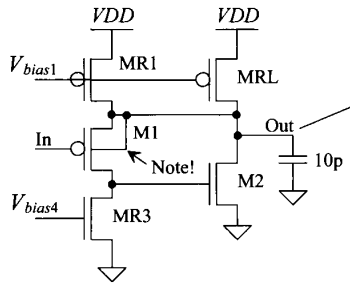


Figure 31.46 Comparing series-shunt feedback amplifier to a source-follower both driving a 10 pF load.

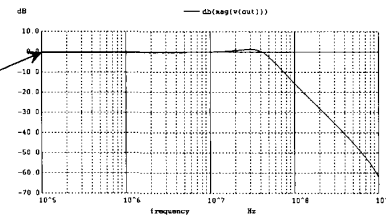
To start we might ask why the source-follower's gain isn't also closer to unity? We answered this question back in Sec. 21.2.4. The body effect causes the reduction in gain. It's easy to verify with simulations (connect M1's body to its source, that is, the output) that this is also the problem with the feedback amplifier. In an n-well process we can't make this connection since the p-substrate is at ground. We could try increasing R_2, say to 25k, to increase the amplifier's gain. However, the process and temperature variations will

IPR2023-00698
Theta EX2019

cause the gain to increase above unity in some situations. A simpler solution is to use the complement of this feedback amplifier where M1 is a PMOS device and M2 is an NMOS device, Fig. 31.47a. The simulation results are seen in Fig. 31.47b.



(a) Series-shunt amplifier with PMOS for M1 and NMOS for M2. Note M1 is in its own well.



(b) Simulating (a) with a 10 pF load.

Biasing supplied by Fig. 20.47, see Table 9.2.

Figure 31.47 (a) PMOS series-shunt amplifier and (b) simulation results.

Before moving on to another topic note that to increase the bandwidth of the amplifier we need to make the capacitive load "effectively" smaller. We know that for general design it's important to keep the overdrive voltages constant between devices by ensuring that we control the devices' biasing currents[1]. What this indicates is that we only have control over the devices' widths once an overdrive, length, and biasing current are selected for a specific transition frequency, f_T. So, to make the capacitive load "effectively" smaller we need to increase the widths of the devices (overdrive voltages remain the same while current goes up). This increases the bandwidth of the amplifier (try increasing the widths of the MOSFETs used in the simulations that generated Figs. 31.46 or 31.47 by 4 and re-simulating with the 10 pF load).

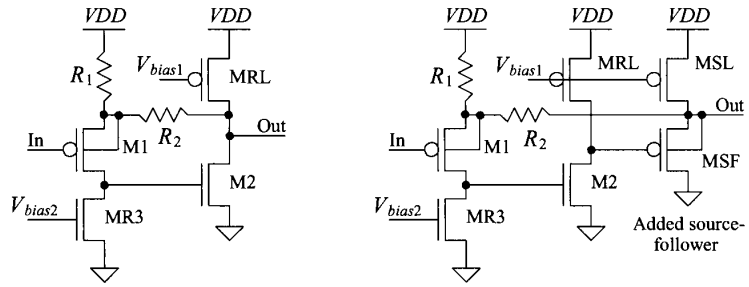
Amplifiers with Gain

Notice in the series-shunt amplifiers seen in Figs. 31.46 or 31.47 that the current flowing in M1 is provided by MR3. This is important since, as we just mentioned, we want to set the biasing current in our transistors. Further, we know the current in M2 as well since the sum of the currents flowing in M1 and M2 must be equal to the currents flowing in MR1 and MRL. In this topology it's easy, as long as we aren't driving a resistive load, to bias the transistors so they have the attributes seen in Table 9.2.

Next, consider the series-shunt amplifier seen in Fig. 31.48a. Here, as before, we've set R_4 to zero and replaced R_3 and R_L with transistors MRL and MR3. The use of transistor loads maximizes A_{OL} as discussed at the beginning of this section (again, A_{OL} can be increased further using cascoding to increase the output resistance of MRL and

[1] Contradictions to this comment occur when we increased the widths of the diff-pair back in Ch. 26 to increase their g_m, or the widths of a current mirror load to reduce input-referred offset or noise, or we are driving a resistive load and there is no way to keep the driving device's current constant (among other contradictions).

IPR2023-00698
Theta EX2019



(a) Series-shunt amplifier with gain. (b) Improved series-shunt amplifier with gain.

Figure 31.48 Using a source-follower to ensure good biasing.

MR3). The current flowing in M1 is still set by MR3. However, now the current flowing in M2 depends on the input signal and the size of the resistors R_1 and R_2. This is bad because we aren't controlling the M2's operating point (e.g., g_m). To avoid this situation consider adding the source follower to the output of the circuit as seen in Fig. 31.48b. Since the source follower's gain is near one we can continue using the equations we derived earlier. Further, now M2's current is set by MRL so its behavior is well-controlled. MSF can sink a current greater than the current supplied by MSL and the resistors since its gate is able to move freely. This is why we must use a PMOS source-follower and not an NMOS source follower, with a constant current sink load, on the output of this circuit. Since MSF is operated as a source follower its g_m and speed performance aren't as important to the overall feedback amplifier's performance.

As a design example let's set, in Fig. 31.48b, R_1 = 1k and R_2 = 9k. Also, to ensure the source-follower has plenty of drive, let's use a multiplication factor, M, of 4 in this stage. The means the widths of MSL and MSF are increased to 400 and their bias currents increase to 40 μA. The overdrive voltages are unchanged from the values listed in Table 9.2. Before simulating let's calculate the allowable input signal range for proper operation. We expect, since the gain of the amplifier is 10, that this swing should be less than *VDD*/10 or 100 mV. The minimum output voltage is set by the source-gate voltage of MSF and the minimum voltage across M2 so that

$$V_{out,\min} = V_{SG} + V_{DS,sat} = 400\ mV$$

Note that MSF's source-gate voltage can actually drop lower than 350 mV since we aren't controlling the current in this device as discussed above. Referring the minimum output voltage back to the source of M1 through the voltage divider

$$\text{Source of M1} = VDD - 400\ mV \cdot \frac{1k}{1k+9k} = 960\ mV$$

Since the V_{SG} of M1 is 350 mV, the input voltage is 610 mV or an input signal swing from (roughly) 560 mV to 660 mV (where we centered the swing on 610 mV since MSF's V_{SG} can be less than 350 mV). Setting the proper input DC operating voltage over process, temperature, and power supply variations can be a significant design concern (AC coupling can help). Figure 31.49 shows the simulation results again with a 10 pF load.

IPR2023-00698
Theta EX2019

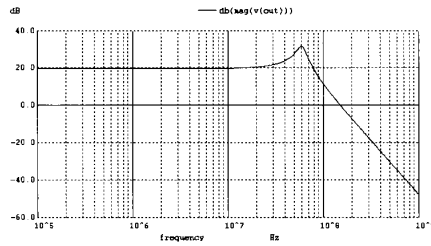


Figure 31.49 Simulating the amplifier in Fig. 31.48b in a gain of 10 configuration.

Finally, notice the increased peaking in the frequency response seen in Fig. 31.49. This additional peaking can be attributed to the added source follower. As discussed on pages 694 and 695 the impedance looking into the output of the source follower (the source of MSF) becomes inductive when it's driven with a resistive load. To remove, or reduce, the peaking note that the drains of M1 and M2 are high-impedance nodes (e.g., nodes 1 and 2 in Fig. 21.25). The compensation procedures given in Chs. 21 and 24 (e.g. using split-length devices as seen in Fig. 24.21) can be used to reduce the bandwidth and cause the amplifier's response to roll-off at – 20 dB/decade.

31.9.2 A Transimpedance Amplifier

The transimpedance, or transresistance, (shunt-shunt) amplifier was discussed back in Sec. 31.5 (and Ex. 8.17). One application of this feedback amplifier is seen in Fig. 31.50. The feedback resistor seen in Fig. 8.34 is removed to increase the gain (to lower the input-referred noise current). *Reset* goes low to enable sensing the diode's current, i_d. The RMS output noise, if C_F is small, is approximately $\sqrt{kT/C_F}$. This topology may not be useful in a communication receiver application where data are continually present on the input of the amplifier; however, it can be very useful in imaging applications. We should also point out why we want the input resistance of the transimpedance amplifier (TIA) to be zero, that is, so that all of the reverse-biased photodiode's current is input to the TIA. Both sides of the diode are at AC ground so the diode's capacitance doesn't steal some of the diode's current (the capacitance doesn't discharge). Note that ideally

$$|A_{OL}| = |-R_m| \rightarrow \infty \tag{31.103}$$

The (ideal) closed-loop gain is

$$A_{CL} = \frac{1}{\beta} = \frac{-1}{j\omega C_F} = \frac{v_{out}}{-i_d} \tag{31.104}$$

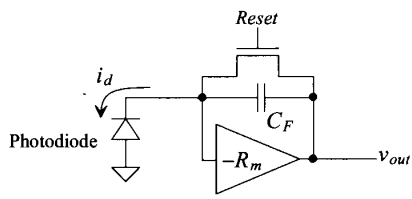


Figure 31.50 Using a transimpedance amplifier (TIA) to sense a photodiode's current.

IPR2023-00698
Theta EX2019

Let's get an idea for the size of C_F for a particular application. Suppose the diode generates one electron every 100 ns. This is equivalent to an average diode current of $1.6 \times 10^{-19} C/100\ ns$ or 1.6 pA. If we set C_F to 10 fF then the output of the TIA will appear (ideally, neglecting noise and photodiode dark current, i.e., leakage current) as steps, when the electron is generated, with heights of

$$V_{step} = 1.6 \times 10^{-19}/C_F = 16\ \mu V \tag{31.105}$$

Clearly increasing C_F reduces the sensitivity of the TIA. Why not reduce C_F then? The answer is that the device and interconnect parasitics are comparable to 10 fF (if careful) so reducing C_F results in non-ideal behavior. Realistically, 10 fF is questionable if care isn't exercised when designing the TIA.

Returning to the calculations, the output of the TIA rises, when *Reset* is low, at a rate of

$$\frac{dv_{out}}{dt} = \frac{16\ \mu V}{100\ ns} = \frac{i_d}{C_F} = \frac{1.6\ pA}{10\ fF} = 160\ mV/ms \tag{31.106}$$

For an imaging application at 30 images per second we may have upwards of 30 ms to sense the diode's current before resetting the diode. In this case this slow-rate of ascent doesn't appear to be a problem (but it depends on the application). We'll return to these calculations in a moment.

Next we need to select a shunt-shunt feedback amplifier topology. It's important to note that when *Reset* goes high the TIA is in the unity-follower configuration and so negative feedback must be employed (the gain from the input to the output has to be negative). This means that we can't use the topology shown in Fig. 31.19. We can, however, use the topology seen in Fig. 31.25. If we replace R_2 in Eqs. (31.68) - (31.73) with $1/j\omega C_F\ (= Z_2)$ and assume Z_2 is much smaller than R_1 or r_{o1} (and $g_{m1}Z_2 >> 1$) then both R_{inf} and R_{of} approach $1/g_{m1}$ (a relatively small value which is what we want). Also note that by replacing R_1 with a PMOS cascode structure and cascoding M1 we can increase A_{OL} with little effect on the noise performance of the amplifier (see Fig. 21.44). The resulting amplifier is seen in Fig. 31.51. Note that the 10 pF load doesn't affect the performance of the TIA (much) since we are sensing for very long periods of time and the transistors M1-M5 are biased with 10 μA. Also, the reset transistor, M5, is made as small as possible so that the charge injection and capacitive feedthrough (when M5 shuts off) are both minimized.

Figure 31.52 shows the simulation results when the input current, i_d, is 10 pA. We expect the output of the TIA to rise at 160 mV/ms as calculated above. However, we see that the output rises at a rate of, roughly, 500 mV/25 μs considerably faster than our calculated value. *What current is the TIA integrating that could cause this error?* Looking at the simulation results we see that M5 is leaking around 13 pA (source of M5 to ground) into the TIA's input. However, the big problem is the current flowing into the gate of M1. This current is roughly 190 pA! The direct tunneling gate current in this book's short-channel CMOS process, from Table 9.2, is around 5 A/cm^2 (generally measured with *VDD* on the gate and the source/drain/substrate grounded). We aren't seeing this much gate current since the gate voltage of M1 is only around 350 mV (see page 475 and the associated discussion for additional information). The question is how do we modify the TIA's design to reduce this error? The simple solution is to reduce the

IPR2023-00698
Theta EX2019

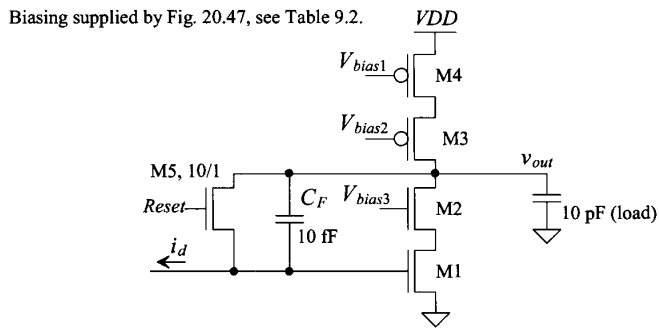


Figure 31.51 A transimpedance amplifier.

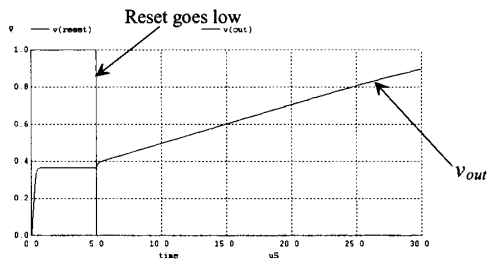


Figure 31.52 Simulating the TIA seen in Fig. 31.51 with 10 pA input current.

widths of M1-M4. Let's drop the NMOSs' widths to 10 and the PMOSs' widths to 20. The actual sizes of the NMOS and PMOS are now 500n/100n and 1μ/100n respectively. Using the same bias circuit, Fig. 20.47, our overdrives remain unchanged and the drain current drops to (roughly) 1 μA. This change also reduces the input-referred noise by increasing A_{OL} and reducing the size of the flicker and thermal drain noise currents. The simulation results, again with a 10 pA input current and 10 pF load, are seen in Fig. 31.53. The leakage current from M5, i_{S5}, is still around 13 pA while the leakage current, i_{G1}, from the gate of M1 dropped to 36 pA. The net current integrated by the TIA is then the 49 pA leakage current from M1 and M5 added to the 10 pA of signal current for an output change of

$$\frac{dv_{out}}{dt} = \frac{i_d + i_{S5} + i_{G1}}{C_F} = \frac{59\,pA}{10\,fF} = 5.9\ V/ms = 59\ mV/10\ \mu s \qquad (31.107)$$

which matches fairly well with the simulation results seen in Fig. 31.53. Note that one might be tempted to compensate for M1/M5's leakage current by adding another, opposite direction, leakage path (another MOSFET connected to the input). This is generally not the best solution since this approach increases the TIA's input capacitance and noise. Further the leakage varies with process shifts and *VDD*. A better solution is to use an older CMOS process, if possible/available, with considerably less gate leakage current.

IPR2023-00698
Theta EX2019

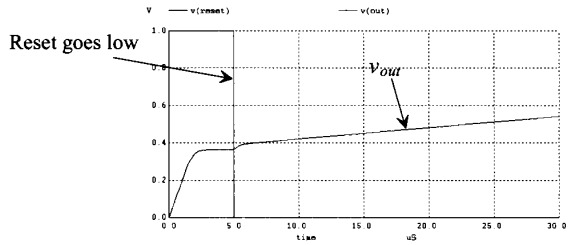


Figure 31.53 Reducing the widths of the MOSFETs in the TIA to lower the integrated gate current from M1.

REFERENCES

[1] A. S. Sedra and K. C. Smith, *Microelectronic Circuits, Sixth Edition,* Oxford University Press, 2010. ISBN 978-0-19-532303-0.

[2] W. M. C. Sansen, *Analog Design Essentials,* Springer, 2006. ISBN 978-0-387-25746-4.

[3] P. J. Hurst, "Exact Simulation of Feedback Circuit Parameters," *IEEE Transactions on Circuits and Systems*, Vol. 38, No. 11, pp. 1382-1389, November 1991.

PROBLEMS

31.1 An op-amp is designed so that the open-loop gain is guaranteed to be 150,000 ± 10 percent V/V. If the amplifier is to be used in a closed-loop configuration with $\beta = 0.1$ V/V, determine the tolerance of the closed-loop gain.

31.2 What is the maximum possible value of β using resistors in the feedback loop of a noninverting op-amp circuit? Sketch this op-amp circuit when $\beta = 1/2$.

31.3 Examine the feedback loop in Fig. 31.54. A noise source, v_n, is injected in the system between two amplifier stages. (a) Determine an expression for v_o which includes both the noise and the input signal, v_s. (b) Repeat (a) for the case where there is no feedback ($\beta = 0$). (c) If $A_1 = A_2 = 200$, and feedback is again applied around the circuit, what value of β will be required to reduce the noise by one-half as compared to the case stated in (b)?

31.4 An amplifier can be characterized as follows:

$$A(s) = 10,000 \cdot \frac{100}{s + 100} \text{ V/V}$$

A series of these amplifiers are connected in cascade, and feedback is used around each amplifier. Determine the number of stages needed to produce an overall gain of 1,000 with a high-frequency rolloff (at – 20 dB/decade) occurring at 100,000 rad/sec. Assume that the first stage produces the desired high-frequency pole and that the remaining stages are designed so that their high-frequency poles are at least a factor of four greater.

IPR2023-00698
Theta EX2019

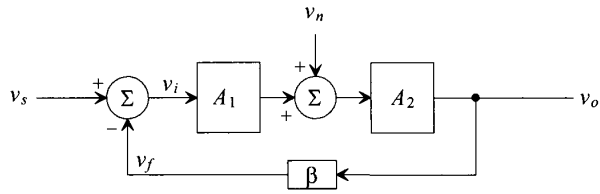


Figure 31.54 Problem 31.3, noise injected into a feedback system.

31.5 An amplifier can be characterized as follows:

$$A(s) = 1{,}000 \cdot \frac{s}{s+100} \text{ V/V}$$

and is connected in a feedback loop with a variable β. Determine the value of β for which the low-frequency rolloff is 50 rad/sec. What is the value of the closed-loop gain at that point?

31.6 Make a table summarizing the four feedback topologies according to the following categories: input variable, output variable, units of A_{OL}, units of β, method to calculate $R_{\beta i}$ and $R_{\beta o}$, and expressions for A_{CL}, R_{if}, and R_{of}.

31.7 Using the two n-channel common source amplifiers shown in Fig. 31.55a and the addition of a single resistor, draw (a) a series-shunt feedback amplifier, (b) a series-series feedback amplifier, (c) a shunt-shunt feedback amplifier, and (d) a shunt-series amplifier. For each case, identify the forward and feedback paths, ensure that the feedback is negative by counting the inversions around the loop, and label the input variable, the feedback variable, and the output variable. Assume that the input voltage has a DC component that biases M1.

31.8 Repeat problem 31.7 using the two-transistor circuit shown in Fig. 31.55b.

31.9 Repeat problem 31.7 using Fig. 31.55c.

31.10 Repeat problem 31.7 using Fig. 31.55d.

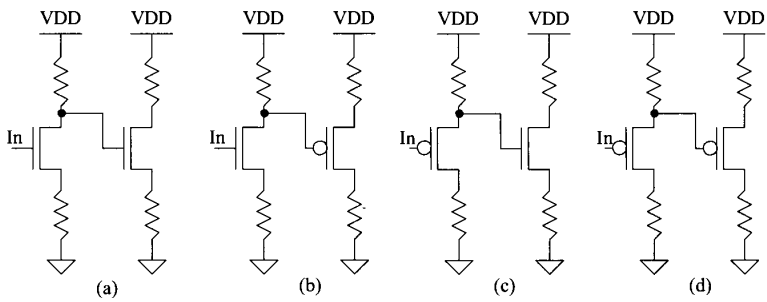


Figure 31.55 Two-transistor feedback topologies.

IPR2023-00698
Theta EX2019

For each of the following feedback analysis problems, assume that the circuit has been properly DC biased and that MOSFETs have been characterized. The n-channel devices have $g_m = 0.06$ A/V and $r_o = 70$ kΩ. The p-channel devices have $g_m = 0.04$ A/V and $r_o = 50$ kΩ.

31.11 Using the series-shunt amplifier shown in Fig. 31.56, (a) identify the feedback topology by labeling the mixing variables and output variable, (b) verify that negative feedback is employed, (c) draw the closed-loop small-signal model, and (d) find the expression for the resistors $R_{\beta i}$ and $R_{\beta o}$.

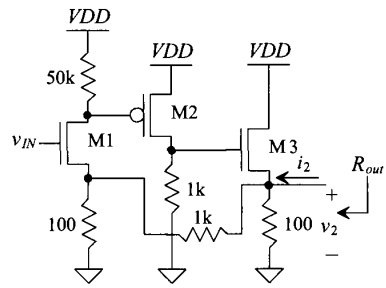


Figure 31.56 A series-shunt feedback amplifier with source-follower output buffer.

31.12 Using Fig. 31.56 and the results from problem 31.11, (a) draw the small-signal open-loop model for the circuit and (b) find the expressions for the open-loop parameters, A_{OL}, β, R_i, and R_o and (c) the closed-loop parameters, A_{CL} and R_{out}. Note that finding R_{in} is a trivial matter since the signal is input into the gate of M1.

31.13 Using the series-shunt amplifier shown in Fig. 31.57, (a) verify the feedback topology by labeling the mixing variables and the output variable closed-loop small-signal model and (b) find the values of $R_{\beta i}$ and $R_{\beta o}$.

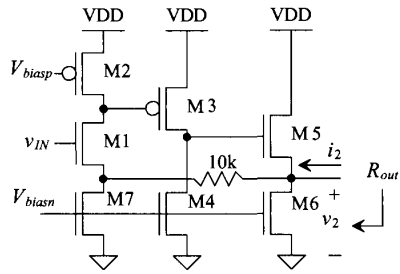


Figure 31.57 A series-shunt amplifier with source-follower output buffer.

IPR2023-00698
Theta EX2019

31.14 Using the series-shunt amplifier shown in Fig. 31.57 and the results from problem 31.13, (a) draw the small-signal open-loop model for the circuit and (b) calculate the open-loop parameters, A_{OL}, β, R_i, and R_o and (c) the closed-loop parameters, A_{CL}, and R_{out}. Note that Fig. 31.57 is identical to Fig. 31.56 except that the resistors have been replaced with active loads.

31.15 Using the principles of feedback analysis, find the value of the voltage gain, $\frac{v_2}{v_{in}}$ and $\frac{v_2}{i_2}$ for the series-shunt circuit shown in Fig. 31.58.

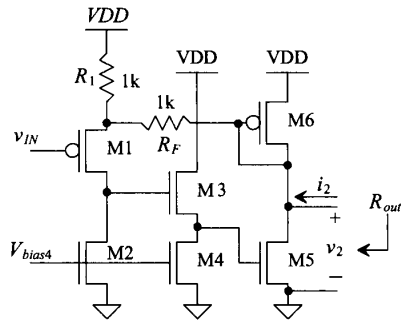


Figure 31.58 Feedback amplifier used in problem 31.15.

31.16 A shunt-shunt feedback amplifier is shown in Fig. 31.59. (a) Identify the feedback topology by labeling the input mixing variables and the output variables, (b) verify that negative feedback is employed, (c) draw the closed-loop small-signal model, and (d) find the values of $R_{\beta i}$ and $R_{\beta o}$.

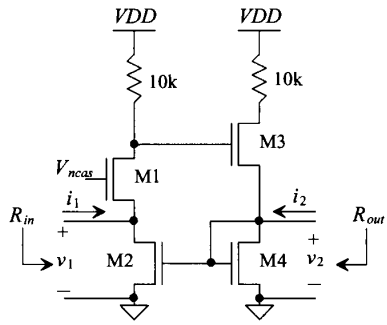


Figure 31.59 A shunt-shunt feedback amplifier.

31.17 Using the shunt-shunt amplifier shown in Fig. 31.59 and the results from problem 31.16, (a) draw the small-signal open-loop model for the circuit and (b) calculate expressions for the open-loop parameters, A_{OL}, β, R_i, and R_o and (c) the closed-loop parameters, A_{CL}, R_{in}, and R_{out}.

IPR2023-00698
Theta EX2019

31.18 Using the principles of feedback analysis, find the value of the voltage gain, $\frac{v_2}{v_1}$, $\frac{v_1}{i_1}$, and $\frac{v_2}{i_2}$ for the shunt-shunt circuit shown in Fig. 31.60.

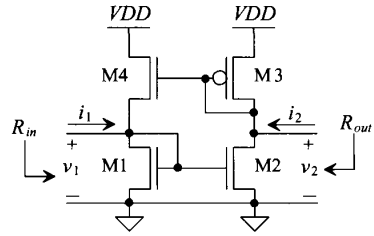


Figure 31.60 A shunt-shunt feedback amplifier, see problem 31.18.

31.19 Using the series-series feedback amplifier shown in Fig. 31.61, (a) identify the feedback topology, (b) verify that negative feedback is employed, (c) draw the closed-loop small-signal model, and (d) find the values of $R_{\beta i}$ and $R_{\beta o}$.

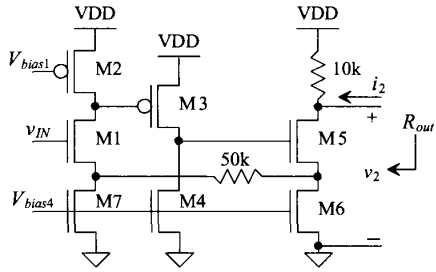


Figure 31.61 Series-series feedback amplifier with source-follower output buffer.

31.20 Using the series-series amplifier shown in Fig. 31.61 and the results from problem 31.19, (a) draw the small-signal open-loop model for the circuit and (b) calculate the open-loop parameters, A_{OL}, β, R_i, and R_o and (c) the closed-loop parameters, A_{CL}, R_{in}, and R_{out}.

31.21 Using the shunt-series amplifier in Fig. 31.35, derive the expressions for A_{OL}, R_{if}, and R_{of}.

31.22 Convert the shunt-shunt amplifier shown in Fig. 31.59 into a shunt-series feedback amplifier without adding any components. (a) Identify the feedback topology, (b) verify that negative feedback is employed, (c) draw the closed-loop small-signal model, and (d) find the values of $R_{\beta i}$ and $R_{\beta o}$.

IPR2023-00698
Theta EX2019

31.23 Using the shunt-series amplifier from problem 31.22, (a) draw the small-signal open-loop model for the circuit and (b) calculate the open-loop parameters, A_{OL}, β, R_i, and R_o and (c) the closed-loop parameters, A_{CL}, R_{in}, and R_{out}.

31.24 A feedback amplifier is shown in Fig. 31.62. Identify the feedback topology and determine the value of the voltage gain, $\frac{v_2}{v_1}$, R_{in}, and R_{out}.

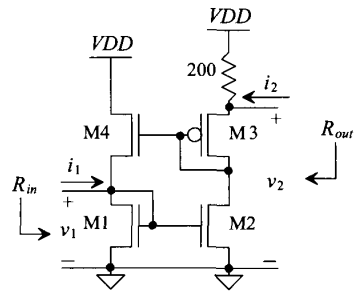


Figure 31.62 Feedback amplifier used in problem 31.24.

31.25 Notice that the amplifier shown in Fig. 31.63 is a simple common source amplifier with source resistance. Explain how this is actually a very simple feedback amplifier and determine the type of feedback used. Determine A_{OL} and β.

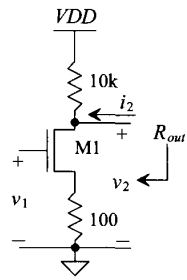


Figure 31.63 Common-source amplifier with source degeneration.

31.26 A feedback amplifier is shown in Fig. 31.64. Identify the feedback topology and determine the value of the voltage gain, $\frac{v_2}{v_1}$, R_{in}, and R_{out}.

31.27 A feedback amplifier is shown in Fig. 31.65. Identify the feedback topology and determine the value of the voltage gain, $\frac{v_2}{v_1}$ and R_{out}.

31.28 Prove that the expression for the open-loop gain derived in Ex. 31.3 is correct.

31.29 Determine if the amplifier seen in Fig. 31.44 is stable.

IPR2023-00698
Theta EX2019

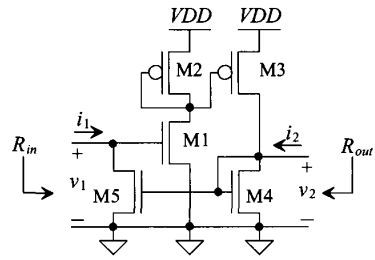


Figure 31.64 Feedback amplifier used in problem 31.26.

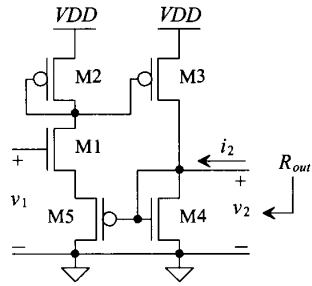


Figure 31.65 Feedback amplifier used in problem 31.27.

31.30 The op-amp shown in Fig. 31.66a can be modeled with the circuit of Fig. 31.66b. With a feedback factor, $\beta = 1$, determine if the op-amp is stable (and the corresponding phase and gain margins) for the following transfer function and $\omega_2 = 10^5, 10^6, 10^7$, and 5×10^6 rad/sec.

$$A_{OL}(j\omega) = \frac{10,000}{\left(1 + j\frac{\omega}{100}\right)\left(1 + j\frac{\omega}{\omega_2}\right)}$$

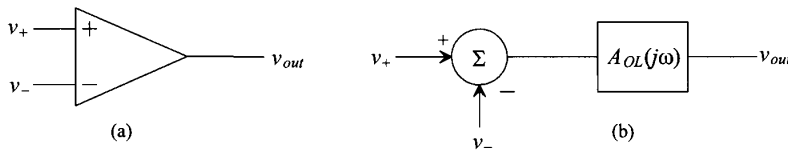


Figure 31.66 Modeling the op-amp.

31.31 The phase plot of an amplifier is shown in Fig. 31.67. The amplifier has a midband gain of −1,000 and 3 zeros at $\omega = \infty$ and three other unspecified poles. If the amplifier is configured in a feedback configuration and β is frequency independent, what is the exact value of β that would be necessary to cause the amplifier to oscillate?

IPR2023-00698
Theta EX2019

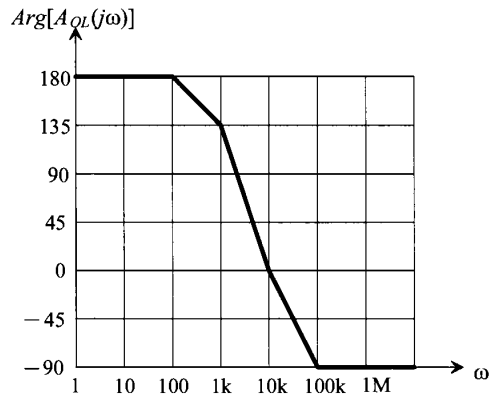


Figure 31.67 Phase response used in problem 31.31.

31.32 You have just measured the gain of the op-amp circuit shown in Fig. 31.68. You know from basic op-amp theory that the gain of the circuit should be $-R_2/R_1$ V/V. However, your measurements with R_2 = 10 kΩ and R_1 = 1 kΩ revealed that the gain was only –5 V/V. What is the open-loop gain of the op-amp?

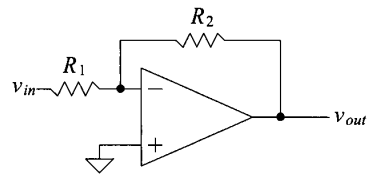


Figure 31.68 How finite open-loop gain effects closed-loop gain, problem 31.32.

31.33 Using the circuit shown in Fig. 31.69 and the *RR* method, find a value of R_1 and A_o which will cause the phase margin to equal 45° at ω = 8,000 rad/sec. The amplifier can be modeled as having an infinite input impedance and zero output resistance and has a frequency response of

$$A(s) = \frac{-A_o}{(s/200 + 1)(s/10,000 + 1)}$$

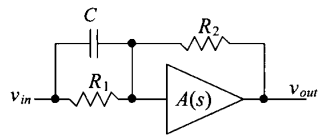


Figure 31.69 Amplifier used in problem 31.33.

IPR2023-00698
Theta EX2019

31.34 Determine if the system with a return ratio as described by Eq. (31.102) is stable.

31.35 Redesign the transimpedance amplifier seen in Fig. 31.51 using the long-channel CMOS process discussed in this book.

31.36 How would the input-referred noise for the TIA design presented in Sec. 31.9.2 be reduced?

IPR2023-00698
Theta EX2019

Index

A

ABSTOL, 28
Abut, 98, 427, 429-430
AC analysis in SPICE, 19-20, 285
AC small-signal analysis, 280
Accumulation, 132-134, 136, 159, 514
Active area, 83-99
  design rules, 99
Active-PI loop filter, 573-574
Active load, 657, 872
Active pixel sensor, 504-505
Adaptive biasing, 920-923
ADC, see Analog-to-digital converter
Adder, 367-369, 373, 395, 407-410, 412, 419-421
Aliasing, see Analog-to-digital converter
Amplifier,
  adaptive biasing, 920-923
  biasing, 291-292, 297-302, 613, 621, 624-625, 635-652, 699, 863-866
  body effect 691-692, 709
  cascode, 686-689, 696-697
  class A, 672, 688, 691, 697, 701, 727
  class AB, 672, 697-701, 709-710, 727-728, 800, 802-805, 810, 814, 817, 819-820, 865, 878
  common-gate, 671, 689-690, 708-709, 732, 783
  common-source, 657-658, 666-667, 671-673, 698-699, 717, 751, 756, 773-774, 793, 800, 814, 820-823
  current source load, 657, 665, 671-674, 686, 689-691, 698-699, 704, 719, 731-733, 774, 876
  distortion, 701, 704-706, 710, 737-740, 790, 807, 831
  efficiency, 699-702
  feedback, see Ch. 31
  frequency response, 661-666, 674-685, 687, 696-697, 708, 721
  gain, 279, 291, 657-673, 686-692, 719-720
  input capacitance, 693, 1141
  noise, 219-262, 669-670, 686, 688, 694
  noise figure, 233-240
  overshoot, 811, 895, 936
  phase response, 662-665, 675-688
  pole splitting, 676-685
  push-pull, 698-703
  slew rate, 672-673
  source follower, 670, 690-698
Amplifier (continued),
  transimpedance, 666-667
Analog-to-digital converter (ADC), 931, 947-957, 985-1015
  0.5 LSB, 943
  1.5 bits/stage, 1068-1075
  aliasing, 953-955, 1008
  aperture error, 956-957
  architectures, 985-1016
  binary search, 1003-1005
  calibrating, 1079
  charge redistribution, 1005-1007
  comparator placement, 1061, 1070, 1082
  coarse conversion, 991
  cyclic, 1059-1066
  DFT, see discrete Fourier transform
  differences of oversampling and Nyquist rate converters, 1008-1009
  differential nonlinearity (DNL), 950-951, 989-990, 998, 1007
  digital error correction, 1068
  discrete Fourier transform (DFT), 218, 956
  fine conversion, 991
  flash, 985-990
  flash accuracy issues, 988-990
  folding frequency, 955
  gain error, 953, 992-994
  higher-order sigma-delta modulation, 1014
  implementing, 1052-1095
  integral nonlinearity (INL), 951-952, 989, 997-998, 1007
  least significant bit (LSB), 940
  missing codes, 951
  noise shaping, 1011-1012
  Nyquist rate, 1007
  offset, 953, 989-990, 996-998, 1007
  operational amplifier issues, 992-994
  oversampling, 1007-1015
  parallel, feed-forward, 990
  pipeline, 994-998, 1067
  pipeline accuracy issues, 996-998
  pulse-density modulation, 1009
  quantization error, 948-949
  quantization levels, 947-949
  quantization noise, 956
  residue, 994
  resolution versus oversampling, 1014-1015
  sampling function, 954-955
  SAR, 1003-1005
  scale factor error, 953
  sigma-delta modulation, 1009
  signal-to-noise ratio (SNR), 956
  specifications, 947
  successive approximation, 1003-1007
  successive approximation accuracy issues, 1007
  successive approximation register, see SAR, 1003-1005
  transfer curve, 948-953, 987, 994

IPR2023-00698
Theta EX2019

Analog-to-digital converter (ADC) (continued),
two-step, 990-992
two-step accuracy issues, 992
Analog models, 269-310
body effect transconductance, 287-288
transconductances, 279-280, 288
Analog signal characteristics of, 932
Anisotropic etch, 171
Anode, 43, 757, 760
AOI logic, 364-369
Aperture error, see sample-and- hold
Arbiter, 383
Astronomy, 246
Autozeroing, 1006, 1051, 1053
Avalanche breakdown, 143, 154, 253-254
Averaging noise,
flicker, 246
thermal (white), 240-241
Averaging circuit, 539-540, 873-874, 1078-1079

B

β (transconductance parameter), 142
Back end of the line (BEOL), 90, 177-178, 199-208
Backend processes, 208-211
Bandgap energy of silicon, 759
Bandgap voltage references (BGR), 745, 761-770
Battery, 695, 730
BCD, see binary coded decimal
BEOL, see back end of the line
Beta-multiplier reference (BMR), 624-630, 647, 650, 750-754
compensation, 630
Biasing,
general long channel analog design, 291-293
general short channel design, 297-300
high-speed design, 299-302, 863-866
long-channel circuit, 647
push-pull amplifier, 699
short-channel circuit, 650
Binary coded decimal, 965
Bi-phase, 584-585
Bipolar junction transistor (BJT), 47, 150
lateral, 758
parasitic, 2, 37, 56, 341-342, 757-758
temperature behavior, 758-759
vertical, 757
Bird's beak, 181
BJT, see bipolar junction transistor
Bloats, 35
BMR, see beta-multiplier reference
Body effect, 139, 148-149, 276-277, 691-692, 709
gamma, γ, 139, 276
Body transconductance, 287-288
Bonding pad, 4, 59
layout, 59-62, 100-102
Bootstrapped inverter, 546-547
Bottom plate capacitance, 115, 839-840
Bottom plate sampling, 839-842, 1052
Breakdown,
diode model in SPICE, 58
oxide, 154
voltage, 58, 143
BSIM4, 154-157
Buffers,
class AB output, 672, 697-701, 709-710, 727-728, 800, 802-805, 810, 814, 817, 819-820, 865, 878-891, 900-901
digital, 344-348
fully differential input, 920
input, 534-542
NMOS only, 350, 546-547
op-amp, 793-810, 865, 878-891, 900-901
source follower, 696-698
Bulk CMOS process, 31, 180
Buried channel, 304
Bus, 68, 100-103, 321, 351, 412-419, 427, 527

C

C'_{ox}, 114-115, 123-124, 132-135
Calibration, 1023-1024, 1035-1038, 1049-1050, 1061, 1079
Calma stream format, see GDS, 36
Caltech intermediate format, see CIF, 36
Capacitance, 44, 135
bottom plate capacitance, 115, 839-840
bottom and sidewall, 117
depletion, 43-45, 47-49, 58, 94, 100, 117, 122, 133, 147, 244, 315, 342-343, 398, 478, 546, 618, 844
diffusion, 45-47, 117
metal layer parasitics, 61
MOSFETs 116-118, 132-135
sidewall, 44-45, 117
Capacitive feedthrough, 449, 831-832, 1050
Capacitor
common-centroid layout, 980
depletion, 43
poly-poly, 113-115
layout, 113-115, 851-852
parasitics, 61, 115
temp co, 116
voltage coefficient, 116
Capacitor error averaging, 1068, 1075-1083
Carrier lifetime, τ_T, see diode, minority carrier lifetime
Carrier transit time, τ_T, see diode, minority carrier lifetime
Cascode,
amplifier, 686-689, 696-697
current mirror, 636-652
differential amplifier, 733-735, 793-794, 877-891
origin of the name, 636

IPR2023-00698
Theta EX2019

Cathode, 43, 757
CGBO, 135
CGDO, 123-124, 135
CGSO, 123-124, 135
Channel hot electron (CHE), 468-469
Channel-length modulation (CLM), 143-147, 273, 277, 280-281, 289, 613-617, 639, 856
Charge injection, 830-840, 856, 863, 937, 1050
Charge pumps, 329, 461, 542-546, 558-561, 578-580, 591-593, 600, 603
CHE, see channel hot electron
Chemical/mechanical polishing (CMP), 163, 170, 173, 182, 185, 187, 201-207, 211
Chemical vapor deposition (CVD), 176-177, 179, 182, 184-185, 187, 190, 193, 196, 199, 202-207
Chip, 3, 32, 36
  cross-sectional view, 4, 32
  design, 1
  IC design flowchart, 2
  packaging, 4, 211
  layout view, 4, 32
  organization, 447
CIF (Caltech Intermediate Format), 36, 53
Class AB, 672, 697-701, 709-710, 727-728, 800, 802-805, 810, 814, 817, 819-820, 865, 878
CLM, see Channel length modulation
Clock feedthrough, 449, 456, 830, 832-834, 838-840, 860, 863, 937
Clock generation, 577, 1082-1084
  nonoverlapping, 401-402, 843, 850, 853-856, 874, 896, 1084
Clock recovery circuit, 551-552, 584, 588-589, 591, 602-607, 608, 611
  self-correcting, 588-594
Clock synchronization, 383, 551, 591, 596
Clocked logic, 397, 403-408
CMOS, 1
  active pixel sensor, 504-505
  IC design process, 1-2
  passive elements, 105-106
  process flow, 91
  process integration, 161
  scaling, 54, 152
  trends, 152
  twin-tub, 31
  unit processes, 161
CMOS fabrication, 161-212
  APCVD, 176
  back end of the line (BEOL), 90, 177-178, 199-208
  backend processes, 208-211
  BEOL, see back end of the line
  boule, 163
  burn-in, 208, 211
  chemical/mechanical polishing (CMP), 163, 170, 173, 182, 185, 187, 201-207, 211
  chemical vapor deposition (CVD), 176-179, 182, 184-187, 190, 193, 196, 199, 201-207
CMOS fabrication (continued),
  critical dimension (CD), 179
  contact module, 202-203
  contaminates, 171, 208
  Czochralski (CZ) growth, 162-163
  degree of anisotropy, 170-171
  deposition, 90, 161, 173-177, 179, 191-192, 196, 200-208, 210-211
  depth of focus (DOF), 168-169, 201, 205
  die seperation, 3, 211
  diffraction effects, 168-169
  diffusion, 117, 163, 165-167, 211
  doping processes, 52-53, 161, 165-167
  dry etching, 170-171, 208
  electrical coupling, 202, 205
  electronic grade silicon (EGS), 162
  electromigration problems, 203
  etching, 33, 90-91, 93, 170-172, 175-176, 183, 190, 192, 200, 202, 208, 211, 444
  etch rate, 170-171
  etch profile, 171
  EGS, see electronic grade silicon
  epitaxial wafers, 31-32, 180
  FEOL, see front end of the line
  field oxide (FOX), 38, 83-84, 86-87, 90-91, 116-117, 181
  front end of the line (FEOL), 90, 177-178, 180-199
  final test, 208, 211
  gate oxides, 163, 191-192
  IMD, see intra-metal (layer) dielectric
  intra-metal dielectric (IMD), 69, 178, 205-208
  ion implantation, 140, 165-167, 170, 184, 211
  ion implanter, 165
  implant profile, 165-166
  isolation, 90-91, 173, 177, 181-183, 205
  lattice damage, 165, 184
  LOCOS, 181
  mask alignment, 84, 86, 94, 170
  masks, 34-35, 53, 69, 84, 86, 163, 167-168, 170, 173, 179
  metallization, 178, 203-205
  metallurgical grade silicon (MGS), 162
  MGS, see metallurgical grade silicon
  n-wall, 184-186
  n-well, 179, 181, 184, 187-190
  numerical aperture, 168
  p-wall, 184-186
  p-well, 179, 181, 184, 187-191
  packaging, 3-5, 59, 208, 211
  passivation, 62, 173, 178-179, 208-210
  pattern transfer, 168
  PECVD, 176-177
  PECVD reactor, 177
  photolithography, 161, 167-171, 173, 182-193, 196, 202-203, 206, 208-211
  physical vapor deposition (PVD), 175-176, 200, 204, 208

IPR2023-00698
Theta EX2019

CMOS fabrication (continued),
plasma etching, 173-177
polysilicon, 166, 178-179, 191-194
pre-metal dielectric (PMD), 178, 200-204
process description, 89-91, 178-179
process integration, 177-208
rapid thermal annealing, 199-201
rapid thermal processing (RTP), 164
reactive ion etching (RIE), 171-173
registration errors, 170
resolution, 168-169, 171
reticle, see mask
salicide module, 199
scribe lines, 209, 211
selectivity, 170
shallow trench isolation (STI), 90-91, 177-178, 181-188
silica, 161-162
silicide, 83, 88-92, 94, 199-201
SOI wafers, 180
source drain module, 193-199
starting material, 180-181
step coverage, 174
stepper, 34, 167-168
straggle, 165
streets, see scribe, 209, 211
technology computer assisted design (TCAD), 179
thermal oxidation, 34, 161, 163-164, 182, 184, 190, 192, 211
thin film deposition, 173-177
thin film removal, 161, 170-171
trench liner, 184
twin tub (well) module, 187-190
unit processes, 161-177
via 1 module, 205-207
wafer manufacture, 161-163
wafer probe, 208-211
wet etching, 170-171
yield, 152, 209-211
CMOSedu design rules, 53-55, 69
CMFB, see common-mode feedback
CMRR, see common-mode rejection ratio
Common-centroid, 111-113, 618, 724, 851, 916, 980
Common drain amplifier, see source follower
Common gate amplifier, 671, 689-690, 708, 732, 783
Common-mode feedback (CMFB), 836-837, 863, 869-875, 881-885, 888-904, 906-908, 1057-1058
compensating, 871-873
dynamic, 1089-1091
settling time, 895
switched-capacitor, 874-875, 896-904, 907-908, 1089-1091
Common-mode rejection ratio (CMRR), 721-724, 742, 773, 789-790, 823, 825, 833
Common-mode voltage (V_{CM}), 837
Common source amplifier, 657-658, 666-667, 671-673, 698-699, 717, 751, 756, 773-774, 793, 800, 814, 820-823
Communication system, 582-583
Comparators, 909-930
block diagram, 910
characterizing, 915-918
clamped input stage, 917
clocked, 918-919
dc performance, 915-916
decision circuit, 910-914
dynamic, 854-856
gain and offset, 916, 1061-1062, 1068, 1075
hysteresis, 912-914
input slew rate, 918
inverter based, 854-856
kickback noise, 448-449, 456, 910, 918, 930
operation, 909
output buffer, 913-915
placement, 1061, 1082-1084
pre-amp, 910-913
propagation delay, 918
self-biased, 920
sensitivity, 909-910, 930
transient response, 916
Compensation, 773-792, 794, 814-815, 819-820, 892, 895
beta-multiplier reference, 630
common mode feedback (CMFB), 871-873
constant-g_m, 738-740
high-speed, 783-787
indirect, 783-787, 825
gain enhancement, 809-812
lead, 782
Miller, 774, 788
nested Miller, 820
pole splitting, 674-679, 685, 780, 783, 814, 821
slew-rate limitations, 787-789
Conductivity, 39-41
silicon-dioxide, 468
Constant-g_m,
bias circuit, 625
diff-amp, 738-740
Contact resistance, 47, 64, 70-71, 80, 147, 209, 962
Contact potentials, 42-43, 124, 137-138
Convergence help in SPICE, 28, 401
Complementary to absolute temperature (CTAT), 745, 759, 761-762, 765, 768
Counter, 329, 485-486, 506, 998-1002
as a digital filter, 506-507
Counter-doped, 52-53
Coupled noise, 1092-1093
rejection, 838
Critical dimension (CD), 179
Cross-sectional view, 3-5
Crosstalk, 59, 71-72, 427

IPR2023-00698
Theta EX2019

CTAT, see complementary to absolute temperature
Current differential amplifier, 737, 921-922
Current mirrors, 273-274, 510, 512, 563, 613-656
  biasing, 621-626, 647, 650, 863-866
  cascode, 273-274, 636-646
  design, 615
  dynamic, 856-858
  floating, 647, 651-652, 698-699, 703-704, 710, 804-805, 808, 820, 865
  layout, 616-620, 642
  low-voltage, wide-swing, 639-651
  matching, 616-620
  output resistance, 288, 298, 613, 622, 636-644
  regulated drain, 645-646
  SPICE statement, 620
  subthreshold operation, 635
  supply independant biasing, 624
  temperature coefficient, 632
  V_{THN} mismatch, 616
  KP mismatch, 616-617
  wide swing, 639-651
Current mirror load, 715-717, 731
Current-mode DAC, 1018, 1024-1030, 1042
Current steering DAC, 1042-1044
Current starved VCO, 329, 561-565, 571, 574
CVSL logic, 369-370
Cyclic ADC, 1058-1066
Czochralski (CZ) growth, 162-163

D

DAC, see Digital-to-analog converter
Damascene process, 90, 207
Data conversion
  signal characteristics, 931-935
DC analysis in SPICE, 13-15
DC sweep convergence in SPICE, 29
Decades, 20
Decoder, 329, 430, 433, 447, 449, 457-462, 480, 966, 985-987
Decoupling capacitor, 73-74, 81, 417, 419
Decibels, 20
Delay time, 49-52, 65, 89, 317-318, 348, 358-359, 362, 526-527, 531-533
  NAND gate, 392
  phase shift, 665
  TG, 376
  word line, 443
Delay elements, 408, 595-596
Delay locked loop (DLL), 592-602
Delay & transition times 317-320
Delta-sigma modulation, see Sensing using delta-sigma modulation (DSM), see also Sigma-delta modulation
Depletion capacitance, 43-45, 47-49, 58, 94, 100, 117, 122, 133, 147, 244, 315, 342-343, 398, 478, 618, 844
Depletion device, 140
Depletion region, 43
Depth of focus (DOF), 168-169, 201, 205
Design rules, 53-54, 69, 75, 98-99
  active and poly, 98-99
  CMOSedu, 53-55, 69
  n-well, 36-37, 53-54
  metal layers, 69
DFT, see discrete Fourier transform
DIBL, see drain induced barrier lowering
Dickson's charge pump, 544
Dielectric constants, 114
Die, 3-4, 32, 59-60, 62, 68, 75, 110, 116, 209, 211, 411-412, 427, 429, 443, 746, 958, 960, 998
Die seperation, 3, 211
Differential amplifier, 711-744
  ac operation, 718-721
  adaptive biasing, 920-922
  body effect, 730, 742
  cascode, 733-735
  class AB, 727-733
  common-centroid layout, 724
  common mode range (CMR), 713-716, 732-733, 735
  common mode rejection ratio (CMRR), 721-723
  constant transconductance, 738-740
  current, 737, 921-922
  current mirror load, 715-717, 731-733
  dc operation, 711
  gain, 719, 722, 733, 736
  low voltage, 736
  matching, 724-726
  minimum power supply voltage, 717
  noise, 726-727
  output resistance, 719, 733-734
  self-biased, 534-538, 920
  slew rate, 727
  source cross coupled pair, 727-733
  wide swing, 736-740, 918-919
Differential nonlinearity (DNL), 941-943, 950-951, 963, 967-971, 976-980, 989-990, 998, 1007, 1029-1034, 1040-1041
  improving using segmentation, 1032-1034, 1043
Differential output op-amp, 836-842, 863-908
  benefits, 838
  simulating, 841, 878-904, 927
Diffusion,
  capacitance, 45-47, 117
  process, 117, 163, 165-167, 211
Digital error correction, 1068 , *see also* Analog-to-digital converter (ADC), 1.5 bits/stage
Digital-to-analog converter (DAC), 931, 938-947, 965-985
  0.5 LSB, 943
  accuracy, 941
  binary switch array, 966-967
  binary weighted, 971, 974-975, 977-980
  calibration, 1035-1038
  charge scaling, 978-982, 1005-1007

IPR2023-00698
Theta EX2019

Digital-to-analog converter (DAC) (continued),
current source mismatch, 976-977
current steering, 973-975, 1042-1044
cyclic, 982-983, 1019
decoder, 966
differential nonlinearity (DNL), 941-943, 963, 967-971, 976-980
dynamic range, 947, 963
full scale voltage, 939, 941, 963
gain error, 945-946, 953, 1019
input combinations, 938
integral nonlinearity (INL), 943-945, 967-971, 976-977, 979, 980
ladder, 967, 971
latency, 945
least significant bit (LSB), 940-941
LSB, see least significant bit
monotonicity, 943
most significant bit (MSB), 940
MSB, see most significant bit
nonmonotonic, 943
offset, 943-946
op-amp issues, 992-994
pipeline, 984-985, 1098
R-2R, 971-973, 1024-1035, 1038-1040
reference voltage, 938-941, 956
resistor string, 966-971
resistor string mismatch, 967-970
resolution, 933-934, 937-938, 940-943, 946-947
settling time, 945
segmentation, 1032-1034
signal-to-noise ratio (SNR), 945-946, 956
specifications, 938-947
split-array, 980-981, 1019
transfer curve, 940
W-2W, 1043-1044
wide-swing, 1026-1032
without an op-amp, 1038-1044
Digital models, 311-319
Digital phase-locked loop (DPLL), 551-612
active-PI loop filter, 573-574, 581, 588, 591
block diagram, 552
charge pump, 558-561, 578-580, 591-593, 600, 603
dead zone, 577, 600, 609
delay elements, 595-596
examples, 596-607
jitter, 557, 564, 571, 579, 582, 591-592, 594, 597, 600
lock range, 567, 569-570, 573-574, 576-577, 579-580, 592
loop filter, 552, 554-557, 559-561, 567-582
PFD, 552-553, 557-561, 575-580, 591, 600-603, 608
pull-in range, 567-569, 571, 574, 576-577, 579, 592
static phase error, 572-573, 577, 579, 600, 604, 610
Digital phase-locked loop (DPLL) (continued),
system concerns, 582
tristate output, 559-560, 575-579, 581
XOR DPLL, 568-574
XOR Phase Detector, 552-557, 559, 561
Diode,
anode, 43, 757, 760
breakdown voltage, 58, 253
carrier transit time, τ_T, see minority carrier lifetime
cathode, 43, 757
depletion capacitance, 43-45, 117
diffusion capacitance, 45-46, 117
drain or source to substrate, 117, 143, 398, 403, 543, 618
electrostatic discharge, 100-102
ESD, see diode, electrostatic discharge
long base, 45
minority carrier lifetime, τ_T, 45-48, 242-243
n-well/p-substrate, 31-32, 39-45, 53, 57-58, 94, 96, 543
noise, 242-244, 249, 252-253
parasitic BJT, 757-759
photodiode, 249, 504
reverse recovery time, 46-49
saturation (scale) current, I_S, 39, 47, 181, 398
Schottky, 92, 760
short base, 45
SPICE modeling, 14-15, 47-49
storage capacitance, see diffusion capacitance
storage time, 46-49
temperature behavior, 759
t_{rr}, see reverse recovery time
t_s, see storage time
τ_T, see minority carrier lifetime
Zener, 253-254
Discrete Fourier transform (DFT), 218, 956
Distortion,
amplifier, 701, 704-706, 710, 737-740, 807, 993-994
common-mode level variation, 740, 790, 807
multiplier, 923
pulse width, 531
signal-to-noise plus distortion ratio (SNDR), 956
SPICE modeling, 705-706
switches, 831
total harmonic distortion (THD), 704-706, 923
Distortionless transmission, 582
DLL, see Delay-locked loop
DNL, see Differential nonlinearity
Dominant pole, 680
Domino logic, 405-407, 409
Doping processes, 52-53, 161, 165-167
Doublet, 810
DPLL, see Digital phase-locked loop
Drain induced barrier lowering (DIBL), 154
DRAM, see memory, DRAM

IPR2023-00698
Theta EX2019

Drawn length, 87, 116-117, 132, 144
Drawn width, 68, 87, 116-117, 132, 315
DRC (design rule check), 36
Drive current, see on current
Drivers, 344, 449, 459-463, 481, 546
  distributed, 347-348
DSL logic, 370-372
DSM, see Sensing using delta-sigma modulation (DSM)
Dummy elements, 113, 618, 851
Dynamic circuits, 829-862
  amplifiers, 858
  capacitive feedthrough, 67, 830-832
  charge injection, 830, 856, 830-833
  clock feedthrough, see capacitive feedthrough
  comparator, 854-856
  current mirrors, 856-858
  dummy switch, 832
  IOS, 856-857
  kT/C noise, 833-834, 856
  logic, 397-410
  OOS, 856-857
  op-amp settling time, 852-853
  reducing charge injection & clock feed-through, 832
  reducing op-amp offset-voltage, 853-854
  sample-and-hold, 838-840
  simulating, 401
  storage capacitors, 852-853, 856
  switch, 829-834
  switched capacitor integrator, 845-850
  switched capacitor resistor, 843-845
Dynamic logic, 397-410
  charge leakage, 399-400
  clocked CMOS, 403-408
  domino logic, 405-406
  PE logic, 405
  NP logic, 407
  shift register, 402
  simulating, 401
  zipper logic, 407
Dynamic power dissipation, 339-340
Dynamic range, 490, 506, 833, 839, 947, 963, 1009, 1014

E

Edge detector, 374, 587
Effective channel length, 132
Effective switching resistance, 312-315, 317, 320, 335, 344, 390, 392
Electrode layer, 115
Electrical length, 144
Electromigration, 59, 68, 70
Electrostatic discharge (ESD),
  diode, 59, 100-103, 530-531, 549
Energy,
  bands in silicon, 39-43
Energy (continued),
  electrical, 213-215, 217
  implant, 166
  thermal, 225-226, 236
Energy spectral density (ESD), 217
Enhancement device, 140
Epi layer, 31
Epitaxial wafers, 31-32, 180
ESD, see electrostatic discharge, see also energy spectral density
Etching, 33, 90-91, 93, 170-172, 175-176, 183, 190, 192, 200, 202, 208, 211, 444
Equalizer, 539, 582-583, 610
Excess gate voltage, see overdrive voltage, 271
Excess phase shift, 665

F

Fall time, 317-318
Fermi energy level, 42-43
FEOL (front end of the line), 90, 177-178, 180-199
Feedback, 1099
  amplifier, 1106
  beta network, 1107
  bandwidth extension, 1101-1103
  closed-loop impedances, 1112
  closed-loop parameters, 1112
  determining signal path, 1107
  equivalent transconductance, 1121-1123
  factor β, 776, 1100
  forward path, 1107
  gain desensitivity, 1101
  gain margin, 781, 1138
  input/output impedance control, 1104
  input mixing, 1106
  inversions around the loop, 1107-1108
  open-loop impedances, 1110
  open-loop parameters, 1110
  output sampling, 1106
  phase margin, 781, 1138
  properties, 1104-1105
  recognizing topologies, 1105
  reduction in nonlinear distortion, 1103
  rules for output sampling, 1107
  series-series, 1109-1110, 1128-1132,
  series-shunt, 1141-1143
  series mixing, 1105
  series sampling, 1105
  shunt-series, 1108-1109, 1132-1135
  shunt-shunt, 1109-1110, 1119-1128, 1145
  shunt mixing, 1105
  shunt sampling, 1105
  stability, 777
  using a gate-drain resistor, 1125-1128
Field device, 38, 181, 184-185
Field implant, 38-39
Field oxide, see FOX (field oxide)

IPR2023-00698
Theta EX2019

Flash memory, 113, 463, 469-476, 481-482, 487-489
Flatband voltage, 139
Flicker noise, see noise
Flip-flops, 380-389
  data (D), 386-388
  hold time, 388-389
  metastability, 383-385, 396
  setup time, 388-389
Floating current source, 647, 651-652, 698-699, 703-704, 710, 804-805, 808, 820, 865
Floating gate, 113, 466
Floor plan, 412, 958
FNT, see Fowler-Nordheim tunneling
Folded-cascode, 641, 648-649, 714-715, 802-810, 827-828, 872
Fowler-Nordheim tunneling (FNT), 469-474, 487
FOX (field oxide), 38-39, 83-87, 181
FPGA (field programmable gate arrays), 1, 412
Frequency response,
  amplifier, 661-666, 674-688, 696-697
  doublet, 810
  noise, 220-223
  op-amp, 680, 792-793
  pole splitting, 674-679, 685, 780, 783, 814, 821
  transmission channel, 538-541
Frequency synthesis, 577, 591
f_T, unity current gain frequency, 290-292, 297-302, 309, 316, 863, 903, 906
Fully-differential circuits, 833, 836-842, 863-930
  benefits, 833, 838
  bottom plate sampling, 839-842
  common-mode feedback, 836-837, 863, 869-875, 881-885, 888-895, 896-904, 906-908
  delay elements, 595-602
  gain, 836
  noise rejection, 838
  sample and hold, 838-842, 880-881, 901-904
  SPICE model, 841-842
f_{un}, op-amp unity-gain frequency, 260-261, 680-685, 740, 779-781, 785-786, 792-793, 797-799, 852-853, 993
  gain-enhancment, 808-812
  plot (general), 680

G

Gain-enhancement (GE), 808-812, 889-890
Gain-f_T product (GFT), see MOSFET, GFT (gain-f_T product)
Gain margin, 781
Gamma, γ, 139, 276
Gaussian probability distribution function, 248-249
Gate current, see Tunneling, gate current
GDS, 36, 53, 78
Generation, 40
GFT, see MOSFET, GFT (gain-f_T product)
GIDL, Gate-Induced Drain Leakage, 154
g_m, 279-280
  long-channel equation, 280
  short-channel, 293, 297-299
  subthreshold, 292-293
Glitch area, 1041-1042
GM, see Operational amplifier, gain margin
g_{mb}, 289-290
GMIN, 29, 401
GOX (gate oxide), 86-87, 114
Gradual channel approximation, 140, 151
Graphical design format, see GDS, 36
Gray code, 965-966
Ground bounce, 71-74, 81
Guard rings, 110, 343, 757, 851, 957, 960

H

Harmonic distortion, see Distortion, 704-706, 831, 923, 993-994
Hi-Z, 321-322, 350-351, 370, see also Tri-state outputs
High-speed op-amp, 783-787, 891
High impedance node, 397, 629-630, 673, 696, 721, 796, 819
Hold time, 388-389
Hot-carrier effect, 151
Hysteresis, 523-529, 537-538, 547-548, 910-916, 918

I

I_{drive}, see on current
$I_{D,sat}$, 144-145, 158, 271-272
INL, see Integral nonlinearity
Imaging, 504-505
  noise, 246
  sensing, 504-519, 1145-1148
Impact ionization, 153
Inductance, 72
Incomplete settling, 496, 502-503, 520
Input buffers, 531-542, 920
Input offset storage (IOS), 856-857
Input-referred offset, see Offset voltage
Instanced (or placed), 429
Integral nonlinearity (INL), 943-945, 951-952, 967-971, 976-977, 979-980, 989, 997-998, 1007, 1029-1032, 1034-1038, 1040-1041
Integrator,
  ADC, 998-1002
  delta-sigma, 492, 1010-1014
  noise, 245, 267
  offset, 846
  PLL, 573, 575
  switched capacitor, 845-850, 861, 901, 907-908
Inter layer dielectric (ILD), 69, 81
Interconnect burden, 429
Interdigitated layout, 110-113, 960
Intrinsic,
  carriers, 40-41

IPR2023-00698
Theta EX2019

Intrinsic (continued),
propagation delay, 65-66, 81
silicon, 40-42, 52-53
Inverter, 331-352
buffer design, 344-347
dynamic power dissipation, 339-340
layout, 341-343, 413, 422-425
noise margins, 333
other configurations, 349-351
power delay product, 340
switching characteristics, 337-338
sizing, 344
transfer characteristics, 332-336
tri-state output, 351
VTC, see transfer characteristics
I_{off}, see off current
I_{on}, see on current
Ion implantation, 165
I_S, see Diode, saturation current
Isotropic etch, 171

J

Jagged plots, 29
Jitter, 557, 564, 571, 579, 582, 591-592, 594, 597, 600, 1000
peaking, 582
J_s, scale current density, 57
Junction capacitance, see depletion capacitance, 43

K

Keeper MOSFET, 406-407, 461
Kickback noise, 448-449, 456, 910, 918, 930, 1061
KP, 142
kT/C noise, 229, 240, 244, 250-251, 517, 833-834, 846, 856, 896

L

Lambda, λ,
channel-length modulation (CLM), 144, 147, 289, 298, 617, 620
layout (in the MOSIS rules), 54
wavelength, 168
LASI, see http://CMOSedu.com
Latch, 380-386, 400, 403-404
Latch-up, 2, 46, 180, 341-343, 349-350
Lateral diffusion,
MOSFET source/drain, 116-117, 124, 144, 315, 616, 620
n-well fabrication, 35-36, 56
SPICE, 147
Layout,
active, 83-85
adder, 420-421
binary-weighted capacitor array, 980
bipolar junction transistor, 757-758
Layout (continued),
bonding pad, 59-60, 62, 75-77
capacitor, 113-116, 419, 851,
capacitor bottom plate, 115, 840
common-centroid, 111-113, 724
corners, 38
design rules, 37, 68, 99
diff-pair, 724
diode, 757-760
DRAM cell, 443-446
drawing order, 4
dummy elements, 113, 618, 851
electrode, 115
flash memory cell, 466-471
guard rings, 110, 343, 757, 851, 957, 960
inverter, 341-343, 413, 422-425
large-width MOSFETs, 121
lateral bipolar, 757-758
long-length MOSFETs, 120
matching (to improve), 617-620, 724
MOSFET, 86-87, 95 (NMOS), 96 (PMOS) 116-124, 617
MUX/DEMUX, 422
n-well, 32, 36-37, 94
NAND and NOR gates, 358, 415
pad, 60, 75, 102
parasitic pnp, 757
power and ground, 417
resistor, 32, 94, 109-113
Schottky diode, 760
serpentine pattern, 57
SR latch, 416-417
SRAM, 463-464
standard-cell, 97-98, 343, 413, 427-431
standard cell frame, 97, 343
stick diagram, 422
TG (transmission gate), 415
unit cell, 109, 851-852
view, 3-5
VLSI, 1, 411
Layout view, 3-5
LDD, see lightly-doped drain
L_{diff}, see lateral diffusion
L_{drawn}, see drawn length
L_{elec}, see electrical length
L_{eff}, see effective channel length
Leakage current, 93, 181, 184, 200, 244, 249, 397-403, 547
Level shifting, 547-548, 692
pumped output voltage driver, 546
Lifetime of a minority carrier, τ_T, see diode, minority carrier lifetime
Lightly-doped drain (LDD), 90-91, 116, 151-152, 179, 193-196
Linear region, see Triode region
LOCOS, 181
Long-channel,
analog models, 269-296

IPR2023-00698
Theta EX2019

Long-channel (continued),
digital models, 312-316
MOSFETs, xxxi, 132-150
tables, 292, 317, 320
Low impedance node, 673
Low power, 149, 299, 451, 476, 563, 635, 727, 729, 737
Low voltage, 476
current mirror, 637-639
differential amplifiers, 727, 729, 736-740
reference, 760, 768-770
LTspice, see http://CMOSedu.com

M

Manchester NRZ, 584
Masks, 34-35, 53, 69, 84, 86, 163, 167-170, 170, 173, 179, 851
aligning, 84, 86, 94, 170
CMOS fabrication, 182-209
Matching, 8
capacitor, 116, 1075-1081
common-centroid, 111-113, 618, 724, 851, 916, 980
current mirror, 615-618, 636, 642
diff-pair, 724-726
interdigitated layout, 110
layout, 105, 617-620
NMOS vs PMOS, 440-441
resistors, 110-113, 1040-1041
R-2R DAC, 1029-1030
threshold voltage, 616, 726, 856, 867
transconductance, 616
Measurements, 246, 267
currents in SPICE, 10, 274-275
noise, 213, 219-224
probe, 326-327
Memory circuits, 433-482
array architectures, 434-447
buried capacitor cell, 446
chip organization, 447
decoders, 457-461
DRAM, 348, 438-446
EEPROM, 469
efficiency, 447
EPROM, 468
erased, 466, 471
F, feature size, 443-445
flash, 113, 463, 469-476, 481-482, 487-489
folded bit line, 441-446
floating gate, 466-474
global decoders, 458
local decoders, 458
memory cells, 463
NAND flash, 471
NSA, 435-439
open array, 436
organization, 457
Memory circuits (continued),
peripheral circuits, 448
PE decoder, 461
PROM, 464-465
programmed, 466, 471
PSA, 440-441
RAM block diagram, 433, 457
refresh operation, 441
ROM, 464
row drivers, 461
sense-amp design, 448-456
sensing basics, 435-441
SRAM, 463-464
trench capacitor cell, 444-445
Metal layers, 59-74
bond pad, 4, 59
capacitive feedthrough, 66
cross-sectional view, 64
delays, 65
design rules, 69
electromigration, 59, 68, 70
inter layer dielectric (ILD), 69
intrinsic propagation delay, 65
parasitics, 61, 64-65
sizing, 68-74
Metallurgical grade silicon (MGS), 162
Metastability, 383-385, 396, 450, 452, 478, 918
Miller,
capacitance, 311-312, 661-662, 664, 674, 684
compensation, 774, 788
effect, 311, 687, 693
eliminating, 686
neglected zero, 662-665
nested, 820
theorem, 660-661, 676
Minority carrier lifetime, τ_T, see diode, minority carrier lifetime
Mixed signal, 706, 842, 957, 960
Mobility, 40, 106, 144, 153
Monotonic, 943, 966, 1041
Moore's law, xxxi
MOSFET, 131
accumulation, 132-134, 136, 159, 514
adjusting threshold voltage, 140
capacitances, 117,123, 135
capacitors, 544
C_{gs} calculation, 145
channel-length modulation, 144
depletion, 133
depletion device, 140
diode-connected, 272, 534, 542, 545, 579, 624, 627, 629, 639, 648, 673, 708, 719, 737, 819-820, 872
dummy poly strips, 618
effective width, 116
enhancement device, 140
excess gate voltage, see overdrive voltage
field device, 38, 181, 184-185

IPR2023-00698
Theta EX2019

MOSFET (continued),
flatband voltage, 139
flicker noise, 302-303, 833
f_T, unity current gain frequency, 290-292, 297-302, 309, 316, 863, 903, 906
gain-f_T product (GFT), 300-302
gate tunnel current, 154, 242, 403, 474-475
$I_{D,sat}$, 144-145, 271
I_{off}, 150, 294, 398, 408
I_{on}, 150, 152-153, 157-158, 297, 314, 316
IV characteristics, 140, 157
lateral diffusion, 116-117, 124, 144, 315, 616, 620
layout and cross-sectional views, 86-87, 95-96, 118-122
linear region, see MOSFET, triode region
long channel, xxxi, 132-150, 269-296, 312-317, 320
NMOS layout, 95
ohmic region, see MOSFET, triode region
output resistance, 288
overdrive voltage, 271, 297-302, 616, 800, 802, 804, 863-864, 884, 892, 896, 903, 906
overlap capacitances, 132
oxide encroachment, 116-117, 132, 181-182
pad oxide, 182
parallel connection, 358
parasitic resistances, 118
pinch-off, 143
PMOS layout, 96
punch-through, 143
saturation region, 143-145, 271
scaling, 152
series connection, 359
short channel, xxxi, 151-158, 297-302, 314-317, 320
SPICE statement, 119-120
source/drain depletion capacitances, 117
subthreshold slope, 150
symbols, 96-97, 131
temperature effects, 293-295
threshold voltage, 135-140, 293
triode region, 141, 271, 278
$V_{DS,sat}$, 144-145, 271
MOSIS, 3, 53-55
multiproject wafer, 3
web address, 3
Multipliers, 923-929
multiply-by-2, 992, 994, 1060, 1062-1064, 1071, 1080
multiplying quad, 924
simulating, 926-928
squaring circuits, 928-929
Multivibrator circuits, 529-531, 565-567
MUX/DEMUX, 378-381, 422

N

N-well, 31-58
cross-section, 32, 36, 39
design rules, 36, 54
patterning, 35
resistor, 32, 38, 94
n-well substrate diode, 39
N-select (n+ implant), 84-85, 179
NAND gate, 353-356
layout, 358, 414-415
switching characteristics, 358-363
switching point voltage, 354-355
NBTI, Negative Bias Temperature Instability, 153
Negative frequency, 218
Netlists, downloadable at CMOSedu.com
n_i, see Intrinsic, carrier
NMOS, 95-97, see also MOSFET
NMOS clock driver, 546-547
NMOS inverter, 349, 546-547
Noise, 213-268
1/f, see flicker
amplifier, 220-223
analysis, 235
avalanche, 252-254, 268
averaging thermal (white) noise, 240-241
averaging flicker noise, 246
burst (popcorn), 252-253
correlation, 256-258
differential amplifier, 726-727
diode, 242-244, 249, 252-254
equivalent bandwidth, 220-223
estimating RMS value from the time-domain amplitude, 249
excess, 253
feedback, 259-261
figure, 233-240
flicker, 213, 244-253, 261-262, 267, 302-303, 833
gate tunnel current, 220, 242
Gaussian probability distribution function, 248-249
generator, 253-254
imaging chip, 245-246
input-referred, 220-224, 226-240, 247, 249, 251, 256-261
integrator, 245, 267
kT/C, 229, 240, 244, 250-251, 517, 833-834, 846, 856, 896
margins, 333
mean squared, 214-215, 262
MOSFET, 302-304
modeling, 219
noiseless resistor model in SPICE, 236-238
op-amp, 247-252, 259-261
phase noise, 592
photodiode, 249-250

IPR2023-00698
Theta EX2019

Noise (continued),
pink, 244
popcorn, 252
power spectral density, 215-218
random telegraph signal, 252-253
red, 245
resistors, 225-229
RMS, 215, 217, 219-224, 248-249
root mean squared, see RMS
shot, 242-244, 268
signal-to-noise ratio, 230-234, 246, 266-267, 945-946, 956, 964
spectrum analyzer, 216-218
SPICE modeling, 227-228, 237-238, 243-244, 251-252, 302-304
noiseless resistor model, 236-238
standard deviation, 248-249
substrate, 851
temperature, 239-240
thermal, 225
transimpedance amplifier, 249-251
variance, see Noise, mean squared
white, 219-221, 225, 240-242, 253-254
Zener, 253-254
Nonoverlapping clocks, 401-402, 843, 850, 853-856, 874, 896, 1082-1085
NOR gate, 353-357
layout, 358, 414-415
switching characteristics, 358-363
switching point voltage, 356-357
NRD and NRS, 118-120, 813
Nyquist criterion, 934

O

Octave, 20
Off currrent (I_{off}), 150, 294, 398, 408
Offset binary, 965
Offset voltage,
ADC, 953, 1000, 1005-1007, 1014
adding to a sense amp, 498
charge injection, 831, 836, 841
CMRR, 723, 789-790
common-mode voltage, 792
constant-g_m, 740, 792
DAC, 945-946
delta-sigma ADC, 1014
diff-amp, 725-726, 867-868, 875
flash ADC, 988-990
gain error, 936
integrator, 846
modeling in an op-amp, 776, 827, 867-888
modeling with SPICE, 654, 867-868, 870-871
multiplier, 923
op-amp, 775-776, 804, 813, 818, 825, 853-857, 863, 867-868, 870-871, 892-895, 902-903
pipeline, 996-998
random, 775
Offset voltage (continued),
removal, 518-519, 853-857
sensing, 497-500, 502
systematic, 775, 818, 916, 930
Ohmic contact, 83
Ohmic region, see Triode region
On current (I_{on}), 150, 152-153, 157-158, 297, 314, 316
One shot, 529
Open circuit gain, 298, 300-302, 668, 673, 678, 876
Operating point analysis in SPICE, 282-283
Operational amplifiers, 773-828, 863-908
biasing, 865-866, 876
CMFB, see common mode feedback
CMRR, see common mode rejection ratio
common-mode output voltage, 838
common mode feedback (CMFB), 836-837, 863, 869-875, 881-885, 888-895, 896-904, 906-908
common mode rejection ratio (CMRR), 721-724, 742, 773, 789-790, 823, 825, 833
folded-cascode, 641, 648-649, 714-715, 802-810, 827-828, 872
fully-differential, 836-842, 863-904
f_{un}, op-amp unity-gain frequency, 260-261, 680-685, 740, 779-781, 785-786, 792-793, 797-799, 808-812, 852-853, 993
gain-enhancement, 808-812, 889-890
gain margin (GM), 781
high-speed design, 783-785, 806-812, 863
input common-mode range (CMR), 476, 730, 732, 741, 743, 774, 787, 701-702, 802, 807, 870, 872, 889
layout of differential op-amps, 865
offset-voltage, 775, 853, 867, 893
open-loop response, 780, 781, 795, 799, 801,804, 805, 806, 811, 812, 822, 823
OTA with buffer, 800-808
output voltage swing, 775, 880, 884, 886, 891
phase margin (PM), 781
power amplifier, 807-808
power dissipation, 775
PSRR, 790-791, 838
offsets, 775, 892
RHP zero, 662, 782
settling time, 852, 895, see also Settling time
simulation results, 886, 902
single-ended to differential, 894
slew-rate limitations, 787, 800
SPICE (ideal) model, 628, 841
start-up problems, 887
switched capacitor CMFB, 874-875, 896-897
voltage regulator, 812-817
Operational transconductance amplifier (OTA), 796-808
Optical Proximity Correction (OPC), 79

IPR2023-00698
Theta EX2019

Oscillator,
astable, 530-531
ring, 339-341, 547
Schmitt trigger, 527-529
Voltage controlled, 561-567
OTA, see operation transconductance amplifier
Output buffer, see Buffers
Output Offset Storage (OOS), 856-857
Overdrive voltage, 271, 297-302, 616, 800, 802, 804, 863-864, 884, 892, 896, 903, 906
Overglass layer, 62-63, 75-77, 80, 83
Overlap capacitance, 123-124, 132
Oxide,
breakdown, 154
capacitance, 114-115, 132
encroachment, 116-117, 132, 181-182
growth, 34, 163-164

P

P-select (p+ implant), 84-85, 179
P-wall, 184-186
P-well, 31, 52-54, 179, 181, 184, 187-191, 469-474
Packaging, 3-5, 59, 208, 211
Pad layer, 62
Parasitics, 1-2
Pass gate (PG), 321-326, 357
Pass transistor, see Pass gate (PG)
Passivation, 62, 178-179, 208-210
Path selector, 378
Patterning, 32-35
PCE, see power conversion efficiency
PE logic, 404-405
Peak detector, 539-540
PECVD (plasma enhanced CVD), 176-177
PG, see Pass gate (PG)
Phase-locked loops, see Digital phase-locked loop (DPLL)
Phase margin, 781
Phase shift, 665
Photodiode, 249, 1145
Photolithography, 161, 167-171, 173, 182-193, 196, 202-203, 206, 208-211
Pinch-off, 143
Pipelining, 407
Pitch, 443
Plasma etching, 173-177
PLL, see Digital phase-locked loops
PM, see Operational amplifier, phase margin
PMD, see pre-metal dielectric
PMOS, 96-97, see also MOSFET
PN junction physics, 39-43
PNP, lateral, 757
model, 758
Pole splitting, 674-679, 685, 780, 783, 814, 821
Poly (polysilicon), 83, 166, 178-179, 191-194
design rules, 99
poly2, 113-115
Polycide, 88
Power and ground, 72-74, 417
Power conversion efficiency (PCE), 699-702
Power delay product (PDP), 340
Power dissipation,
inverter, 339-340
Pre-metal dielectric (PMD), 178, 200-204
Probe, 326-327
Process characteristic time constant, 316
Process description, 89-91, 178-179
Process flow, 89-91, 161
Process integration, 177-208
Process, voltage, and temperature (PVT), 299, 531, 745, 867
Proportional to absolute temperature (PTAT), 745, 762
PSRR, 790-791, 838
PTAT, see proportional to absolute temperature
Pulse density modulation, 1008
Pulse statement, 21
Pumps, 542-550
driver, 546
Punch-through, 143
Push-pull amplifier, 698-704
PVT, see process, voltage, and temperature

R

Rail, 698
Random offsets, 775
RC circuit, 17-20, 50
distributed, 49-52, 65, 89, 348
Reactive ion etching (RIE), 171-173
Recombination, 40, 252
References, 745-772
bandgap, 757-772
CTAT, 745, 761-762
diode based, 757-772
low voltage, 760, 770
PTAT, 745, 762-765
thermal voltage, 762-764
Registration errors, 170
Regularity, 412
Regulated drain CM, 645
RELTOL, 28
Resistance, 37
calculation, 37
looking into the drain, 668
looking into the source, 669
resistivity, 37
sheet, 37
Resistor,
common-centroid layout, 111-113
delay through, 49-52
dummy elements, 113
guard rings, 110
interdigitated layout, 110
layout, 38, 56-57, 109-113

IPR2023-00698
Theta EX2019

Resistor (continued),
layout of corners, 38
n-well, 32, 38-39, 94
properties, 88
sheet resistance, 37
switched-capacitor, 843
temp co, 88, 106, 116, 631,
unit elements, 109
Reticle, see mask
Return Ratio, 1139-1141
Reverse recovery time, 46-49
RHP zero, 662, 685
Ring oscillator, 339, 547
Ringing, 695
Rise time, 50, 317-320
Routing, 430
ROX (recessed oxide), see FOX (field oxide)
RSHUNT, 29, 401

S

Saturation current, see Diode, saturation (scale) current, I_S
Salicide, 199-200
Saturation region, 143-145, 271
S/H, see Sample-and-hold
Sample-and-hold, 834-836, 838-842, 880-881, 901-904, 935-938, 1052-1058
acquisition time, 936
aperture error, 937
aperture jitter, 937
aperture uncertainty, 937
droop, 937
gain error, 936
hold mode, 937
linearity, 936
offset, 936
overshoot, 936
sampling error, 937
settling time, 936
single-to-differential, 1054-1058
Scale current, see Diode, saturation (scale) current, I_S
Scale factor, 36-37, 44, 53-56, 60-62, 75, 89, 114, 117, 119-120, 132, 134, 146-147, 154, 157-158
Scaling theory, 152
Scanning Electron Microscope (SEM), 55, 79, 124
Scattering, 52, 106
Schmitt Trigger, 523-529
applications, 527-529
Schottky diode, 92, 760, 772
Scope probe, 326-327
Scribe, 75, 209, 211
Select, 84-85, 179
Self-aligned gate, 86-88
Self-biased diff-amp, 534-538, 914, 920
Self-biasing, 624-626, 750-770
start-up circuit, 625, 629, 752-754
Self-correcting phase detector, 588-590
Sense amplifier design, 448-456
clock feedthrough, 449
contention current, 450
creating an imbalance, 451-455
kickback noise, 449-451, 456
increasing input range, 454, 918-919
removing sense-amp memory, 451
reducing power, 453
simulation examples, 454
Sensing using delta-sigma modulation (DSM), 483-519
dynamic range, 490
examples, 484-486
feedback signal, 492-496
Flash memory, 487-496
incomplete settling, 496
imagers, 504-519
low pass filter, 507
mismatches, 517-519
noise issues, 506-507, 517
offset, 497, 517-519
parasitic capacitance, 492
pixel, 504
precision of sense, 485
programmed state, 487
resistive memory, 497-504
sampling, 505
sensing circuit, 488, 499, 511
sigma-delta modulation, 484
voltage to current conversion, 508-510
Sensitivity analysis, 623, 626
Series connection of MOSFETs, 322
delay, 325-326
Series-shunt feedback, 1045, 1113, 1141-1145
Series-series feedback, 1128
Settling time, 663, 782, 788-789, 800, 802, 804, 810, 819, 840, 852-853, 880, 892, 895, 901-904, 936, 945, 990, 993, 1000
incomplete, 496, 502-503, 520
Setup time, 388-389
Shallow trench isolation (STI), 90-91, 177-178, 180-188
Sheet resistance, 37
Shielding, 961
Shift register, 402
Shotgun shell, 242
Short-channel,
analog models, 297-302
digital models, 314-316
effects, 153-154
MOSFETs, xxxi, 151-158
tables, 300, 317, 320
Short circuit protection, 826-827
Shot noise, 242-244
Shunt-series feedback, 1132
Shunt-shunt feedback, 1046-1047, 1119, 1145-1148

IPR2023-00698
Theta EX2019

Sigma-delta modulation, 1007-1015, see also
Sensing using delta-sigma modulation
first-order, 1010-1014
higher order, 1014
noise shaping, 1014-1015
resolution verses oversampling, 1015
Signal to noise ratio (SNR), 230-234, 246, 266-267, 945-946, 956, 964
Signal to noise and distortion ratio (SNDR), 956
Silicide, 83, 88-92, 94, 199-201
Silicide block, 89
Silicon bandgap energy, 759
Silicon on insulator (SOI), 180
SIN SPICE source, 16-17, 282-287
Single-ended to differential conversion, 894
Single-ended to differential converter,
Skew, 531-537
Slew rate, 727, 787, 800, 853, 918, 922
Small-signal,
AC analysis, 280
models, 279-288
signal parameters, 280, 287-289
SNDR, see signal to noise and distortion ratio
SNR, see signal to noise ratio
SOI, see silicon on insulator
Source-coupled multivibrators, 565
Source-coupled pair, see differential amplifier
Source-coupled VCO, 565
Source degeneration, 667
Source follower, 670, 690-698
body-effect, 691
input capacitance, 693, 1141
output stage, 696-698
Space-charge region, see depletion region
Spectral density of noise, 216-217
Spectrum Analyzer, 216
SPICE, 1, 8-29, 67, 78, 145
ABSTOL, 28
AC analysis, 19-20, 285
bipolar junction transistor, 15
breakdown, 58
common mistakes, 29
convergence, 28, 401
DC analysis, 13-15
diode breakdown voltage, 58
diode model, 14, 47-49, 758
distortion, 705-706
generating netlist, 8
GMIN, 29, 401
initial conditions, 24-25, see also UIC
integrator, 26-28
jagged plots, 29
level, 145
measuring currents, 274-275
mistakes, 29
models, parameters for long channel CMOS process, 145-147
SPICE (continued),
models, parameters for short channel CMOS process, 154, 158
MOSFET models, 302-304
noise analysis, 227-228, 237-238, 243-244, 251-252, 302-304
noiseless resistor model, 236-238
op-amp, 12, 841
opening a netlist, 8
operating point analysis, 9-10, 282-283
piece-wise linear (PWL) source, 23
problems, 29
pulse statement, 21
Q of an LC circuit, 25-26
RELTOL, 28
RSHUNT, 29, 401
Schottky diode, 772
simulating dynamic circuits, 401
SIN source, 16-17, 282-287
subcircuit (subckt), 13
switches, 24
.tran statement, 15, 286, 665-666, 705-706
transient analysis, 15-19, 21-25, 27-28, 286
transfer function analysis, 10-11
UIC, 24-25, 29
Voltage-Controlled Voltage Source (VCVS), 11-12, 628, 841
Voltage-Controlled Current Source (VCCS), 12
VNTOL, 27
Squaring circuits, 928
Square-law equations, 142-144, 271-272
Stability, 776-777
Standard-cell, 97
Standard deviation, 248-249
Start-up circuit, 625, 629, 752-754
Static logic, 353
Stepper, 34, 167-168
STI, see shallow trench isolation
Stick diagram, 422
Storage node, 397
Storage capacitance, 45-46
Storage time, 46-49
Streaming out, 36
Streets, see scribe, 209, 211
Strong inversion, 133
Student projects, 329
Substrate, 31
Substrate pump, 547
Subthreshold,
current, 149, 278
current source/sink, 635
g_m, 292-293
output resistance, 278
slope, 150
transconductance, 292
Switch logic, 369-370
Switched capacitor circuits, 492-494, 843-853
capacitor layout, 851-852

IPR2023-00698
Theta EX2019

Switched capacitor circuits (continued),
filter, 848-849
integrator, 845-848
resistor, 488-489, 843-848
slew-rate requirements, 862
Switches, 311-317, 829-834
Symmetry, 618
Synchronization, 383, 551, 591, 596
Systematic offsets, 775

T

Tail current, 712, 720-722, 798, 803, 865, 868, 873, 882, 894, 898, 903-904, 907, 920-923
Technology computer assisted design (TCAD), 179
Telescopic diff-amp, see Cascode diff-amp
Temperature coefficient,
bandgap references, 761-770
diode, 759
diode self-biasing, 761-762
MOSFET-only, 633-635
MOSFET voltage divider, 749-750
positive, 106-107
resistors, 88, 105-107
resistor-MOSFET divider, 631-632, 746-749
thermal voltage self-biasing, 762-765
threshold voltage, 293-296, 750
transconductance multiplier, 634-635, 754-755
Temp co, see Temperature coefficient
Temperature effects,
bandgap of silicon, 293
capacitor, 116
KP (transconductance parameter), 295, 633
MOSFETs, 293-296
thermal voltage, 294
threshold voltage, 293-296
TG (transmission gate), 324, 375, 829-832
THD, see Total harmonic distortion
Thermal (white) noise, 225-229
Thermal oxidation, 34, 161, 163-164, 182, 184, 190, 192, 211
Thermal voltage, V_T (kT/q), 39, 758
change with temperature, 294
Thermometer code, 965-966, 974-975, 985-986, 988, 992
Threshold voltage, 135-140, 276, 467
adjusting, 140, 190-192
determining, 293
matching, 616
NMOS, 292, 300
PMOS, 292, 300
temperature behavior, 293-296
Timing errors, 583
Total harmonic distortion, 704-706
t_{ox}, 86-87, 114, 292, 300
.Tran statement, 15-16, 286
Transient SPICE analysis, 48, 52, 66-67, 79, 286, 705-706
Transconductance, 280, 297
parameter, β, 142
variation with frequency, 695
Transconductance amplifier, 796-798, 835-836
Transimpedance amplifier, 249-250, 261, 267, 666-667, 688-689, 1145-1148
Transmission channel, 538-541
Transmission gate, 324, 375, 402
applications, 378
delay through, 376
layout, 415
series connection, 377
static gates, 379
Transit time, 45-47
Transition frequency, f_T, 290-292, 297-302, 309, 316, 863, 903, 906
Tri-state outputs, 351, 370
buffer, 372-373
phase detector output, 559
Trimming, 476, 753-755, 766, 769-770, 1007
Triode region, 141, 271, 278
channel resistance, 278
t_{rr}, see diode, reverse recovery time
True-single phase clocked logic (TSPC), 409-410
t_s, see Diode, storage time
TSPC, see true-single phase clocked logic
τ_T, see Diode, minority carrier lifetime
Tub, 52
Tunneling,
Fowler-Nordheim, 469-470, 487
gate current (direct), 153-154, 220, 242, 300, 403, 468-476, 1146-1148
zener diode, 253
Two's complement, 965-966
Twin tub, 31
(well) module, 187-190

U

UIC (SPICE use intital conditions), 24-25, 28
Unbalanced input signals, 894-895
Unity-gain frequency, see f_{un} op-amp unity-gain frequency
Unit processes, 161-177

V

V_{CM}, see common-mode voltage
VCO, see voltage controlled oscillator
$V_{DS,sat}$, 144-145, 271
Velocity,
carriers, 106-107, 142, 151, 153
overshoot, 151, 153, 297, 302
saturation, 151, 291, 299, 634
speed of light, 65
VLSI, 1, 411
VNTOL, 27

IPR2023-00698
Theta EX2019

Voltage coefficient, 107, 116
Voltage-controlled oscillator (VCO), 552, 561, 595
  characteristics, 561
  current starved, 561
  for use with the PFD, 561
  logic based, 567
  source coupled, 565
  use with XOR phase detector, 557
Voltage dividers, 746-756
Voltage follower, 1143
Voltage generators (charge pumps), 542-548
  example, 547-549
  higher voltages (Dickson), 544-545
Voltage references, 745-772
Voltage regulator, 812-819
V_{ov}, see Overdrive voltage
V_T, see Thermal voltage, kT/q
VTC, see Inverter, transfer characteristics
V_{THN}, see Threshold voltage, NMOS
V_{THP}, see Threshold voltage, PMOS

W

Wafer, 3, 32, 161, 180
  manufacture, 161-163
  probe, 208-211
  production, 162
Waterfowl, 242
W_{drawn}, see drawn width
Weak inversion, 132-133, 149-150, 301
Well, 31-58, 53
  contact, 96
  triple, 53, 187, 278
White noise, 219-221, 225, 240-242, 253-254
Wide swing diff-amp, 736-740
Workfunction,
  gates, 191, 304
W-2W current mirror, 1043-1044

X

X_d, 136-137
X_{dl}, 144
XOR gate, 366, 380
  phase detector, 553, 555-556

Y

Yield, 152, 209, 412

Z

Zener diode, 253-254, 268
Zero-nulling resistor, 685, 774-782
Zipper Logic, 407

IPR2023-00698
Theta EX2019

About the Author

R. Jacob (Jake) Baker is an engineer, educator, and inventor. He has more than 20 years of engineering experience and holds over 200 granted or pending patents in integrated circuit design. Jake is the author of several circuit design books. For a detailed biography see: http://cmosedu.com/jbaker/jbaker.htm

IPR2023-00698
Theta EX2019

Square-Law Equations

For a triode-operating long-channel NMOS device

$$I_D = KP_n \cdot \frac{W}{L} \cdot \left[(V_{GS} - V_{THN})V_{DS} - \frac{V_{DS}^2}{2} \right]$$

for $V_{GS} \geq V_{THN}$ and $V_{DS} \leq V_{GS} - V_{THN}$

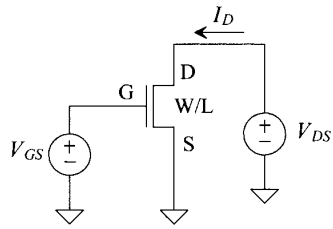
For a long-channel NMOS device operating in the saturation region:

$$I_D = \frac{KP_n}{2} \cdot \frac{W}{L}(V_{GS} - V_{THN})^2[1 + \lambda(V_{DS} - V_{DS,sat})]$$

for $V_{GS} > V_{THN}$ and $V_{DS} \geq V_{GS} - V_{THN}$

On the border between saturation and triode:

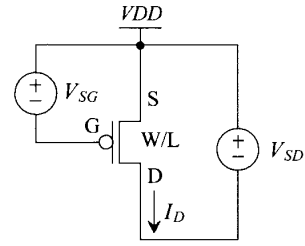
$V_{DS,sat} = V_{GS} - V_{THN}$ and the drain current is called $I_{D,sat}$, see Fig. 6.11



For the PMOS device equations make the following substitutions in the equations listed above

$$V_{DS} \rightarrow V_{SD},\ V_{GS} \rightarrow V_{SG},\ \text{and}\ V_{THN} \rightarrow V_{THP}.$$

All of the voltages and currents in the PMOS and NMOS equations are **positive**. For example, for the PMOS device to conduct a drain current requires $V_{SG} > V_{THP}$. For the NMOS to conduct a drain current requires $V_{GS} > V_{THN}$.



IPR2023-00698
Theta EX2019

Long-channel MOSFET parameters for general analog design in this book VDD = 5 V and a scale factor of **1 μm** ($scale$ = 1e–6)			
Parameter	**NMOS**	**PMOS**	**Comments**
Bias current, I_D	20 μA	20 μA	Approximate
W/L	10/2	30/2	Selected based on I_D and $V_{DS,sat}$
$V_{DS,sat}$ and $V_{SD,sat}$	250 mV	250 mV	For sizes listed
V_{GS} and V_{SG}	1.05 V	1.15 V	No body effect
V_{THN} and V_{THP}	800 mV	900 mV	Typical
$\partial V_{THN,P}/\partial T$	−1 mV/C°	−1.4 mV/C°	Change with temperature
KP_n and KP_p	120 μA/V²	40 μA/V²	t_{ox} = 200 Å
$C'_{ox} = \varepsilon_{ox}/t_{ox}$	1.75 fF/μm²	1.75 fF/μm²	$C_{ox} = C'_{ox}WL \cdot (scale)^2$
C_{oxn} and C_{oxp}	35 fF	105 fF	PMOS is three times wider
C_{gsn} and C_{sgp}	23.3 fF	70 fF	$C_{gs} = \frac{2}{3}C_{ox}$
C_{gdn} and C_{dgp}	2 fF	6 fF	$C_{gd} = CGDO \cdot W \cdot scale$
g_{mn} and g_{mp}	150 μA/V	150 μA/V	At I_D = 20 μA
r_{on} and r_{op}	5 MΩ	4 MΩ	Approximate at I_D = 20 μA
$g_{mn}r_{on}$ and $g_{mp}r_{op}$	750 V/V	600 V/V	Open circuit gain
λ_n and λ_p	0.01 V^{-1}	0.0125 V^{-1}	At L = 2
f_{Tn} and f_{Tp}	900 MHz	300 MHz	For L =2, f_T goes up if L = 1

Models for digital design using the long- and short-channel processes discussed in this book.				
Technology	R_n	R_p	**Scale factor**	$C_{ox} = C'_{ox}WL \cdot (scale)^2$
1 μm (long-channel)	$15k\frac{L}{W}$	$45k\frac{L}{W}$	1 μm	$(1.75fF) \cdot WL$
50 nm (short-channel)	$\frac{34k}{W}$	$\frac{68k}{W}$	50 nm	$(62.5\ aF) \cdot WL$

IPR2023-00698
Theta EX2019

Short-channel MOSFET parameters for general analog design in this book *VDD* = 1 V and a scale factor of **50 nm** (*scale* = 50e–9)			
Parameter	**NMOS**	**PMOS**	**Comments**
Bias current, I_D	10 μA	10 μA	Approximate, see Fig. 9.31
W/*L*	50/2	100/2	Selected based on I_D and V_{ov}
Actual *W*/*L*	2.5μm/100nm	5μm/100nm	L_{min} is 50 nm
$V_{DS,sat}$ and $V_{SD,sat}$ V_{ovn} and V_{ovp}	50 mV 70 mV	50 mV 70 mV	However, see Fig. 9.32 and the associated discussion
V_{GS} and V_{SG}	350 mV	350 mV	No body effect
V_{THN} and V_{THP}	280 mV	280 mV	Typical
$\partial V_{THN,P}/\partial T$	– 0.6 mV/C°	– 0.6 mV/C°	Change with temperature
v_{satn} and v_{satp}	110 x 10^3 m/s	90 x 10^3 m/s	From the BSIM4 model
t_{ox}	14 Å	14 Å	Tunnel gate current, 5 A/cm^2
$C'_{ox} = \varepsilon_{ox}/t_{ox}$	25 *f*F/μm^2	25 *f*F/μm^2	$C_{ox} = C'_{ox}WL \cdot (scale)^2$
C_{oxn} and C_{oxp}	6.25 *f*F	12.5 *f*F	PMOS is two times wider
C_{gsn} and C_{sgp}	4.17 *f*F	8.34 *f*F	$C_{gs} = \frac{2}{3}C_{ox}$
C_{gdn} and C_{dgp}	1.56 *f*F	3.7 *f*F	$C_{gd} = CGDO \cdot W \cdot scale$
g_{mn} and g_{mp}	150 μA/V	150 μA/V	At I_D = 10 μA
r_{on} and r_{op}	167 kΩ	333 kΩ	Approximate at I_D = 10 μA
$g_{mn}r_{on}$ and $g_{mp}r_{op}$	25 V/V	50 V/V	!!Open circuit gain!!
λ_n and λ_p	0.6 V^{-1}	0.3 V^{-1}	$L = 2$
f_{Tn} and f_{Tp}	6000 MHz	3000 MHz	Approximate at $L = 2$

Effective digital switching resistances and oxide capacitances using the drawn sizes seen for both the long- and short-channel processes.					
Technology	**Drawn**	**Scale factor**	**Actual size**	$R_{n,p}$	$C_{ox,n,p}$
NMOS (long-channel)	10/1	1 μm	10 μm by 1 μm	1.5k	17.5 *fF*
PMOS (long-channel)	30/1	1 μm	30 μm by 1 μm	1.5k	52.5 *fF*
NMOS (short-channel)	10/1	50 nm	0.5 μm by 50 nm	3.4k	625 *aF*
PMOS (short-channel)	20/1	50 nm	1 μm by 50 nm	3.4k	1.25 *fF*

IPR2023-00698
Theta EX2019