Chapter 5

CMOS OPERATIONAL AMPLIFIERS

Operation amplifiers (usually referred to as op-amps) are key elements in analogue processing systems. Ideally they perform the function of a voltage controlled current source, with an infinite voltage gain. Operational amplifiers are extensively studied in basic electronic courses. Therefore, their function and operation should be well known to the reader. For this reason, this chapter deals with those circuit implementations that are specifically used to achieve the op-amp function in CMOS integrated VLSI systems. We shall learn that when used inside an integrated architecture, op-amps are mainly employed to drive capacitive loads, namely gates of transistors, capacitors or arrays of capacitors. This makes the request of having a low output impedance of little importance. Therefore, very often op-amps are replaced by operational transconductance amplifiers (OTAs) whose output resistance is quite high.

5.1 GENERAL ISSUES

From the basic courses of electronics we identify an op-amp with the basic schematic shown in Fig. 5.1. It is a four terminal block with two inputs and two

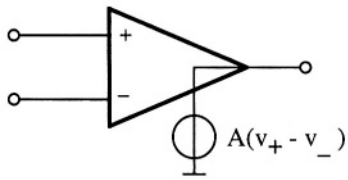


Fig. 5.1 - Symbol and equivalent circuit of the ideal op-amp

outputs. One of the outputs is the analog ground. The key function of the op-amp is to generate at the output an amplified replica of the voltage across the input terminals. Ideally, the voltage gain is infinite. Moreover, the input impedance is infinite as well and the output impedance is zero. Therefore, as we normally say, the op-amp measures the input voltage in a voltmetric fashion and generates the output in the same fashion of an ideal voltage source. From the circuit theory point of view the op-amp performs the same function as a voltage controlled voltage source (*VCVS*).

We know that an op-amp is never used as a stand alone block but has surrounding passive circuit elements to establish some feedback and achieves a given functional operation. It is because of this use that the block is called operational amplifier.

A quite general feedback connection is shown in Fig. 5.2. We have four lumped passive elements that interconnect the two inputs, the output and two input signal generators V_1 and V_2. Impedance Z_2 placed between the inverting terminal and the output ensures that the established feedback is negative. Under the assumption of ideal behaviour, the output voltage is given by

$$V_0 = V_2 \frac{Z_4}{Z_3 + Z_4} \cdot \frac{Z_1 + Z_2}{Z_1} - V_1 \frac{Z_2}{Z_1} \tag{5.1}$$

Thus, the achieved transfer function depends only on the impedances used. However if the op-amp is not ideal, the small-signal output voltage will be affected by those op-amp parameters depicting the non idealities, namely the finite gain, A_0, the finite bandwidth and the finite input and output impedance. If we consider the finite gain only, equation (5.1) is modified into

$$V_0 = \left(V_2 \frac{Z_4}{Z_3 + Z_4} \cdot \frac{Z_1 + Z_2}{Z_1} - V_1 \frac{Z_2}{Z_1} \right) \Big/ \left(1 - \frac{Z_1 + Z_2}{A_o Z_1} \right) \tag{5.2}$$

that is the output voltage calculated previously multiplied by an error term. Observe that the error vanishes when the finite gain goes to infinite. This because the error is proportional to $1/A_0$. Therefore, if we want to limit the error we should use a very large finite gain. However, we should note that the

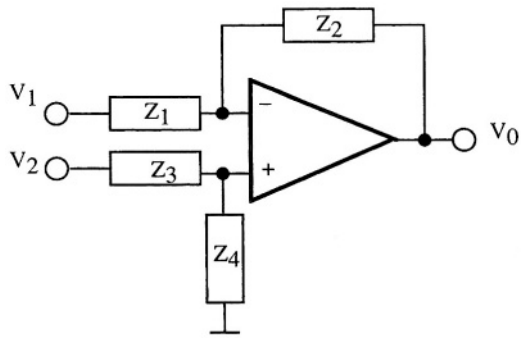


Fig. 5.2 - Typical feedback configuration used around an op-amp

input-output transfer function also depends on ratios between impedances. Possible errors affecting impedances modify too the transfer function thus contributing an additional source of error. The errors caused by the finite gain of the op-amp and that of the impedance mismatch are uncorrelated, so they must be combined quadratically. It turns out that the global error is dominated by the greater one. It follows that it is not advisable to stress the design to achieve a finite gain larger than a given amount. A proper target is to have a finite gain somewhat larger than the inverse of the impedance matching accuracy. Being the matching accuracy of integrated components around *0.1%,* a gain a bit larger than *60-70 dB* is normally enough for most applications.

The input impedance of a *CMOS* op-amp is not normally a problem. Since we use the gate of an *MOS* transistor as input terminals we have a very large input impedance over a wide frequency interval. Instead obtaining low output impedance can be problematic. Nevertheless, the typical use that we make of op-amps in integrated systems does not lead to a strict request for low output impedance. Even better, in some cases the output impedance may not be a problem at all. Let us see why.

KEEP NOTE

The finite dc gain of an op-amp matters until the error that it produces is dominant with respect to other sources of error, like the passive component mismatch.

In many systems the designer uses a capacitor as the feedback element while the output load and the input are capacitors as well, possibly connected by switches. Fig. 5.3 shows a typical situation. When the charged capacitor C_1 is switched onto the virtual ground, a transient in the circuit will take place. The output node supplies current during the transient but after the output node settles the current goes to zero. As a result of this operation the output impedance is no longer important.

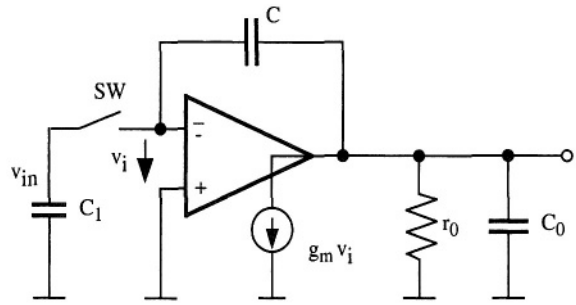


Fig. 5.3 - Typical use of an op-amp in an integrated signal processor

To verify the above statement, let us analyse in some detail the circuit shown in Fig. 5.3. It models the operational amplifier with a transconductance generator $g_m v_i$ having in parallel the output resistance r_0 and the output capacitance C_0. The capacitor C_1, charged at the input voltage V_{in} (thus, $Q_1 = C_1 V_{in}$), at time $t = 0$ is connected by the switch to the virtual ground. At time $t = 0^+$ the op-amp does not react and, assuming C and C_0 to be initially discharged (this is not really necessary since the circuit is linear and we can use the superposition principle), the charge on C_1 is shared with the series connection C - C_0. We obtain the following voltages

$$V_i(0^+) = V_{in}\frac{C_1}{C_1 + (C_0 C)/(C_0 + C)} \tag{5.3}$$

$$V_o(0^+) = V_i(0^+)\frac{C}{C_0 + C} \tag{5.4}$$

Both the virtual ground voltage and the output jump are positive (if V_{in} is positive). Then, the differential input causes a current to come from the transconductance generator that leads to the asymptotic voltages

$$v_i(\infty) = v_{in}\frac{C}{C_1 + C(1 + g_m r_0)} \tag{5.5}$$

$$v_o(\infty) = -v_i(\infty) g_m r_0 \tag{5.6}$$

Observe that the output voltage goes from positive to negative. Moreover, its final value marginally depends on the finite gain of the op-amp $A_0 = g_m r_0$. The transient of the output voltage and the one of the inverting input are described by exponential (Fig. 5.4) whose time constant is

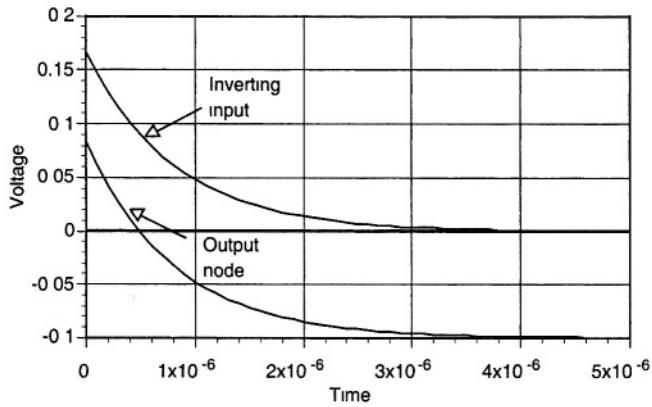


Fig. 5.4 - Transient voltage at the output node and the inverting input of the circuit in Fig 5.3. V_{in} = 1 V; C_1=0.2 pF, C=C_0=2pF; τ = 80 nsec

$$\tau = \left(\frac{1}{g_m} \cdot \frac{C}{C + C_1} \,//\, r_0\right) \cdot \left(C_0 + \frac{CC_1}{C + C_1}\right) \tag{5.7}$$

Since $1/g_m \ll r_0$, the time behaviour of the output voltage is mainly controlled by the transconductance gain and is almost independent of the output resistance. The result achieved is very important for the design of op-amps used in many analog signal processors (switched capacitor circuits, track and hold, data converters). When the feedback network comprises only capacitors (and switches) the output resistance is not relevant. We can use a special class of operational amplifiers where the output resistance can be very high and, possibly, used to enhance the voltage gain. This class of operational amplifiers is called *OTA* (operational transconductance amplifier). An *OTA* achieves a high voltage gain with a given transconductance gain, g_m, by the use of a very high output resistance.

KEEP IN MIND!

A low value of the op-amp output resitance is not essential when the load and the feedback network are made by capacitors and, possibly, switches. In such a situation we can use an *OTA* (operational transconductance amplifier).

5.2 PERFORMANCE CHARACTERISTICS

This section recalls the definition of the most important features of operational amplifiers. Moreover, we shall also consider the most useful circuit con-

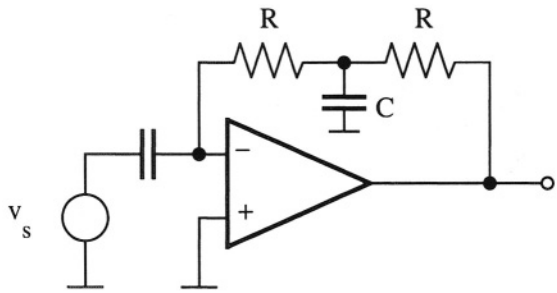


Fig. 5.5 - Schematic arrangement useful for simulating the differential gain of an op-amp.

figurations for the experimental or simulated estimation of the key parameters.

Differential gain (A_d): this is the open loop differential gain measured as a function of frequency. To estimate the differential gain we must compensate the offset (to be discussed shortly) Fig. 5.5 shows a convenient configuration for the *SPICE* estimation of the *dc* differential gain. The small signal input generator is connected between the two input terminals through a big capacitor *C,* while a *T* network made of two resistors and a capacitor establishes a feedback path around the op-amp. Observe that the *T* network acts like a unity gain configuration at very low frequencies and becomes an opened network at very high frequency. A proper choice of resistors and capacitors allows the transition between the two states to take place at a suitably frequency range. For computer simulations we can use huge elements, like *F* or many $G\Omega$ and bring the transition down to a very low frequency. Moreover, observe that the feedback network employed loads the output. Therefore, we have to use resistances *R* which value is higher than the output resistance of the op-amp.

A typical value of the differential gain, A_d, ranges from *70* to *90 dB.* For very precise functions (like high-resolution data converters), the designer needs higher gains in the *100* to *140 dB* range.

Common mode gain (A_{cm}): this is the open loop gain obtained by applying a small signal to both inputs. To measure it we can use the basic configuration shown in Fig. 5.5 with a small modification. We just have to connect the signal generator directly at the positive input and short the positive and negative terminal with a big capacitor *C.* Ideally an op-amp should amplify the differential signal only. Therefore, a low common mode gain over a wide frequency range is quite advisable. A typical value of A_{cm} at low frequency is *10 - 30 dB.*

Common mode rejection ratio (*CMRR*): this is the ratio between the differential gain and the common mode gain. A high *CMRR* is a merit factor for any op-amp.

Power supply rejection ratio (*PSRR*): if we apply a small signal in series

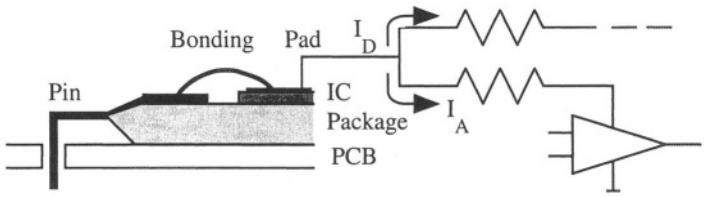


Fig. 5.6 - Typical scheme of connection between the external pin and an analog block.

with the positive or the negative power supply we obtain a corresponding signal at the output with a given amplification (A_{ps+} or A_{ps-}). The ratio between the differential gain and the power supply gain leads to two *PSRRs.* These are two merit factors showing the ability of the op-amp to reject spur signals coming from the power supply.

For mixed-signal circuits the *PSRR* is a very important issue. The power supply that we use in op-amps is not the voltage applied to external pins. As shown in Fig. 5.6, the supply voltage applied to an external pin feeds the integrated circuit through an inductive bonding from the pin to a pad. The inductance is, as a rule of the thumb, $1\ \mu H$ per *mm* of bonding length. Then, metal connections distribute the voltage to various sections of the integrated circuit. The metals display a given resistance that depends on the length of the connection. The resulting network made by a given inductance and resistances leads to a drop voltage between the external pin and the actual supply voltage. In mixed systems the current that we have to account for is not only the one in analog sections; often the dominant spur is caused by the *LdI/dt* noisy contribution related to fast current transitions in digital sections.

Having a good *PSRR* is an important merit factor. Unfortunately, especially at high frequencies, the *PSRR* achieved is typically quite poor. For computer estimation of the power supply rejection the same circuit arrangement shown in Fig. 5.5 can be used: it is required just to use a signal generator in series with a supply connections and set to zero the input signal. A typical value of *PSRR* is *60 dB* at low frequencies that decreases to *20 – 40 dB* at high frequencies.

Offset voltage (v_{os}): if the differential input voltage of an ideal op-amp is zero the output voltage is also zero. This is not true in real circuits: various reasons (that we shall study shortly) determine some umbalancement that, in turn, lead to a non-zero output. In order to bring the output to zero it is therefore required to apply a proper voltage at the input terminals. Such a voltage is the offset.

The simple unity gain configuration shown in Fig. 5.7 permits us to measure the offset. It is represented by a voltage generator in series with the non-inverting terminal. The feedback makes output and inverting input identical. Moreover, assuming that the gain is large enough, the two inputs are equal. In turn, the output measures the offset. Actually the output is not zero, as required: it is the

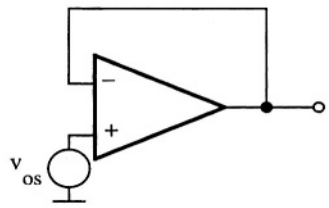


Fig. 5.7 - Schematic arrangement useful for measuring the offset.

offset. Nevertheless, assuming that the offset is relatively small the discrepancy doesn't affect the measure.

Input common mode range: this is the voltage range that we can use at the input terminal without producing a significant degradation in op-amp performance. Since the typical input stage of an op-amp is a differential pair, the voltage required for the proper operation of the current source and the input transistors limit the input swing. A large input common mode range is important when the op-amp is used in the unity gain configuration. In this case the input must follow the output. For such a conditions a so called rail-to-rail operation at the input terminals is often desired.

Output swing: this is maximum swing of the output node without producing a significant degradation of op-amp performance. Since we have to leave some room for the operation of the devices connected between the output node and the supply nodes, the output swing is only a fraction of $(V_{DD} - V_{SS})$. Typically it ranges between *60%* and *80%* of $(V_{DD} - V_{SS})$. Within the output swing range the response of the op-amp should conform to given specifications and in particular the harmonic distortion should remain below the required level.

Equivalent input noise (v_n): the noise performance of a *CMOS* operational amplifier depends on the noise of the transistors used and circuit architecture. Since the *MOS* transistor is a voltage controlled device its noise performance is well described by using an input referred voltage noise generator only. It turns out that the noise in a network caused by the interconnection of *MOS* transistors can also be represented with an input referred noise generator (see

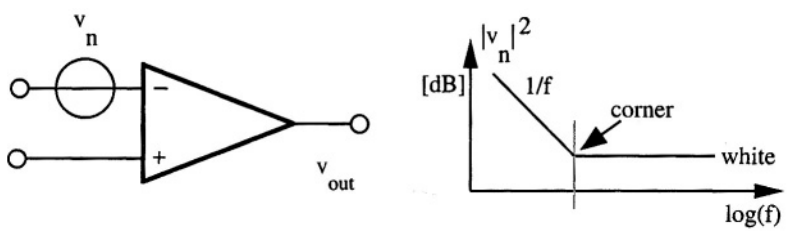


Fig. 5.8 - Equivalent input referred noise voltage generator and its typical spectrum.

Fig. 5.8) that accounts for the effect of all the *MOS* transistors of the network. The noise spectrum of an *MOS* transistor is made by a white and a *1/f* term. Thus, even the spectrum of v_n is made of the same two components: a white term and an *1/f* term. The frequency at which the *1/f* term becomes equal to the white one is called corner frequency. A typical value of the white noise is $50nV/\sqrt{Hz}$. The *corner frequency* depends on the quality of the technology and the kind of input transistors used. It is in the range of 1 *kHz* to *10 kHz*.

Unity gain frequency (f_T) (or gain-bandwidth product, GBW): the speed performance of the op-amp are described by small signal and large signal parameters. The small signal analysis determines the frequency response sketched by a set of zeros and poles. Since we have to ensure stability, one of the poles (f_1) must be dominant. The amplitude Bode diagram will display a *20 dB/decade* roll-off until the gain reaches *0 dB*. The frequency at which the gain becomes *0 dB* is called unity gain frequency, f_T. With a constant roll-off *20 dB/decade* of the achieved unity gain frequency equal the product of gain and bandwidth $f_1 A_o$. Therefore, f_T is also named the gain-bandwidth product, *GBW*. Other poles, $f_2, f_3 \ldots$, exceeding f_T, are named non-dominant poles.

We can measure f_T using the op-amp either in the open-loop or in the unity gain configuration. In the former case the circuit schematic shown in Fig. 5.5 must be used. In the latter we can use the configuration shown in Fig. 5.9. The output voltage follows the input ($A_0 = 0\ dB$) until frequency f_T at which point the gain starts rolling down by *20 dB/decade*.

Phase margin: this is the phase shift of the small-signal differential gain measured at the unity gain frequency. In order to ensure stability when using the unity gain configuration it is necessary to achieve a phase margin better than *60°*. A lower phase margin (like *45°* or less) will cause ringing in the output response. However, for integrated implementation it is not strictly necessary to ensure absolute stability. The use of an op-amp in specific configurations permits to know the value of the feedback factor, β. If β is lower than *1* (as it often happens) the *60°* phase margin should be fulfilled not at the unity gain frequency but at the frequency at which the gain is $1/\beta$.

Slew rate (*SR*): this is the maximum achievable time derivative of the output voltage. It is measured using the op-amp in the open loop or the unity gain con-

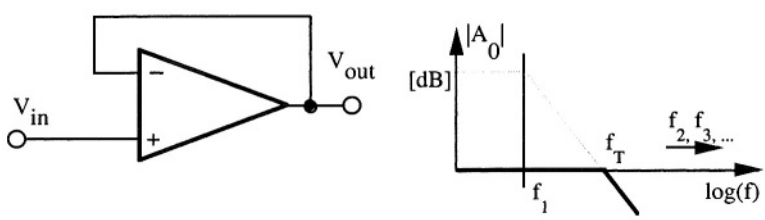


Fig. 5.9 - Schematic arrangement for estimating f_T

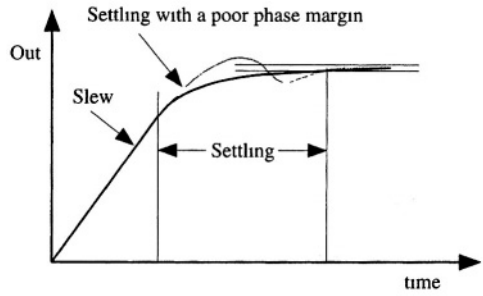


Fig. 5.10 - Typical output response of a real op-amp (slewing and settling limitations)

figuration. A large input step voltage fully imbalances the input differential stage and brings the op-amp output response into the slewing conditions. The positive slew rate can be different from the negative slew rate, depending on the specific design. Typical values ranges between *40* and *80 V/μsec*, but micro-power circuits (for whose the quiescent current available is pretty small) can show much lower figures.

Settling time: this is the time that the output voltage requires, under given operating conditions, to achieve the expected output voltage within a given accuracy (usually *0.1%* or better). The settling time is measured from the end of the slewing period. It critically depends on the phase margin: a poor phase margin leads to a ringing response that augments the settling period. Fig. 5.10 shows the combined effect of slewing and settling.

RECOMMENDATION

Define, analyse, and carefully "negotiate" with system engineers all the op-amp specifications before starting any design.

Ambiguous parameters are always source of many design iteration. By contrast, detailed and well defined specifications permit a quick, safe and successful design.

Power consumption: this is the power consumed under stand-by conditions. Observe that the power used in the presence of a large signal can significantly exceed the one required in the quiescent conditions. Moreover, the consumed power depends on the speed specifications. Typically, higher bandwidth leads to higher power consumption. Low-power operation is a very important quality factor: more and more electronic systems are powered by batteries that should supply the system for hours or days. Therefore, a key design task is to achieve the minimum power consumption for a given required speed. In many applications, we can use a dynamic biasing of the op-amp or a power down mode to reduce power consumption.

Silicon area: specific values of performances discussed above establish a given circuit configuration that, in turn, leads to a corresponding silicon area. Typically the layout of an op-amp has a rectangular shape and includes substrate and well biases. For op-amp schemes of medium or low complexity it is possible to accommodate the entire layout within $2000\ \mu m^2$ (for a $0.25\ \mu m$ CMOS technology).

Table 1 summarises the typical performance parameters achievable with a $0.25\ \mu m$ CMOS Technology. The given figures corresponds to a relatively easy design. Specific architecture permit to improve (even significantly) one or more the given parameters.

TABLE 5.1

Typical parameters of a 0.25 μm CMOS Transconductance Op-Amp

Feature	Value	Unit
DC gain	80	dB
CMRR	40	dB
Offset	4-6	mV
Bandwidth	100	MHz
Slew-rate	3	V/μs
Settling time: 1 V, C_L = 4 pF	300	nsec
PSRR @ dc	90	dB
PSRR @ 1 kHz	60	dB
PSRR @ 100 kHz	30	dB
Input referred noise (white)	100	nV/√Hz
Corner frequency	1	kHz
Supply voltage	3.3	V
Input common mode voltage	1.5	V
Output dynamic range	2.2	V_{pp}
Power consumption	1	mW
Silicon area	2000	μm^2

5.3 BASIC ARCHITECTURE

The typical gain that an op-amp (or an *OTA*) should achieve is around *80 dB*. A simple gain stage has, according to the results of a previous chapter, a gain in the order of *40 dB*. Therefore, the cascade of two stages is normally enough. Alternatively, we can use a cascode with a cascode load configuration that alone obtains around *80 dB*. The choice between a single stage and a two stage architecture depends on a number of design issues that will be considered in some detail later in this chapter. These include dynamic range, bandwidth and power consumption.

As we shall see shortly, the generic circuit schematic of an op-amp can be represented by the functional diagram shown in Fig. 5.11.

REMEMBER

Operational amplifiers that, when used in an integrated circuit, drive capacitive load don't strictly need an output stage. However, circuits with an output stage improve the output current capability, thus reducing the time that the output node takes for slewing.

The first block is a differential amplifier. It provides at the output a differential voltage or a differential current that, essentially, depends on the differential input only. The next block is a differential to single-ended converter. It is used to transform the differential signal generated by the first block into a single ended version. Some architecture don't require the differential to single ended function; therefore, the block can be excluded. Possibly, a second gain stage enhances the differential gain. Finally, we have the output stage. It typically provides a low output impedance or improves the slew rate of the op-amp. Even the output stage can be dropped: many integrated applications do not need a low output impedance; moreover, the slew rate permitted by the gain stage can be sufficient for the application. When the output stage is not used the circuit is not, strictly speaking an op-amp. It is, instead, an operational transconductance amplifier, *OTA*. The name remember us that the circuit achieves the voltage gain using an input transconductor and a relatively large output resistance. The product of transconductance and output resistance fixes upon the voltage gain.

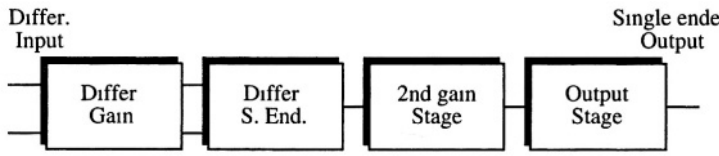


Fig. 5.11 - Functional diagram of a typical op-amp.

5.4 TWO STAGES AMPLIFIER

Fig. 5.12 shows the circuit configuration of a two stage transconductance amplifier. The scheme uses *p*-channel transistors at the input. Of course, it is possible to use a complementary scheme with *n*-channel input transistors. As specified by the name, the circuit is the cascade of two stages: the first is a differential amplifier with differential to single ended transformation, the second is a conventional inverter with active load. The circuit in Fig. 5.12 uses the same reference current for the differential amplifier and the second stage. Therefore, the bias currents in the two stages will be controlled together. Observe that the conversion from differential to single ended is achieved in the first stage with a current mirror (M_3-M_4). As a matter of fact, the signal at the output of the differential pair is current. The current from M_1 is mirrored by M_3-M_4 and subtracted from the current from M_2. The signal contributions of the two currents multiplied by the output resistance of the first stage give the single-ended first stage output voltage. The resulting signal constitutes the input of the second gain stage. Capacitor C_c (or possibly a more complex network) takes care of compensation requirements.

In the next subsection we shall study the features of the two stage op-amp in some detail. Our present target is not to perform complicated analysis but to derive equations appropriate for guiding the computer simulation design. Once again, remember that the design activity is based on computer simulations: the models used by the computer program are much more precise than the ones we use for hand calculations. Nevertheless, it is important to acquire a number of "rules of thumb" that direct the design activity and avoid using computer with a dangerous "try and see what happens" approach.

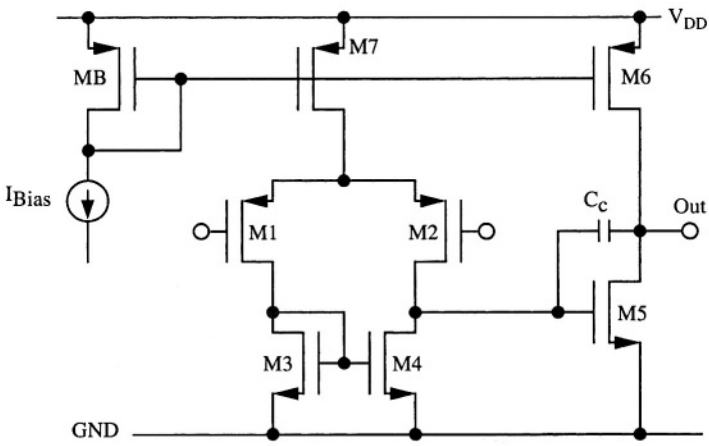


Fig. 5.12 - Basic configuration for a two stages OTA (p-channel input devices).

5.4.1 Differential Gain

Small signal differential gain results directly from the basic block analysis performed in Chapter 3. We can observe that, at low frequency, the second stage does not load the output of the first. Therefore, the low frequency gain is given by the product of the two gains. By inspection of the circuit we obtain

$$A_v = A_1 A_2 = \frac{g_{m1} g_{m5}}{(g_{ds2} + g_{ds4})(g_{ds5} + g_{ds6})} \tag{5.8}$$

$$A_v = \alpha \cdot \frac{\sqrt{\left(\frac{W}{L}\right)_1}\sqrt{\left(\frac{W}{L}\right)_5}}{\sqrt{I_7}\sqrt{I_6}} = \frac{\alpha}{I_{Bias}} \cdot \frac{\sqrt{\left(\frac{W}{L}\right)_1}\sqrt{\left(\frac{W}{L}\right)_5}\left(\frac{W}{L}\right)_B}{\sqrt{\left(\frac{W}{L}\right)_7}\sqrt{\left(\frac{W}{L}\right)_6}} \tag{5.9}$$

Where α is a proper constant that depends on mobility, specific oxide capacitance and the λ factor of the technology used. Moreover, we assume all transistors to be in saturation. Equation (5.9) states that the low frequency gain is inversely proportional to the bias current. The result is obvious: we have already seen that, for transistors in saturation, the gain of a single stage is inversely proportional to the square root of the bias current. Being the circuit the cascade of two stages we achieve immediately the result.

Looking again at equation (5.9) we observe that gain depends on the aspect ratio of four transistors and, through a, on the length of the active load devices of the first and the second stage. This is apparently a fairly large number of degrees of freedom. We shall see shortly that aspect ratios and lengths cannot be defined freely. In order to achieve given basic features (like a zero systematic offset or a symmetrical slew rate) the designer must respect certain constraints.

5.4.2 Common Mode dc Gain

We determine the common mode *dc* gain by applying the same signal to both inputs. Under those conditions the electrical behaviour of the first stage becomes symmetrical. We therefore can derive the common mode *dc* gain using, for the first stage, half circuit. As shown in Fig. 5.13, transistor M_7 is divided into two parts. Moreover, because of the symmetry the drain voltage of M_4 equals the one of M_3. Therefore, we use that voltage to control the gate of M_5. The output of the first half stage is then amplified by the second stage.

The gain of the first half stage should be calculated using the small signal

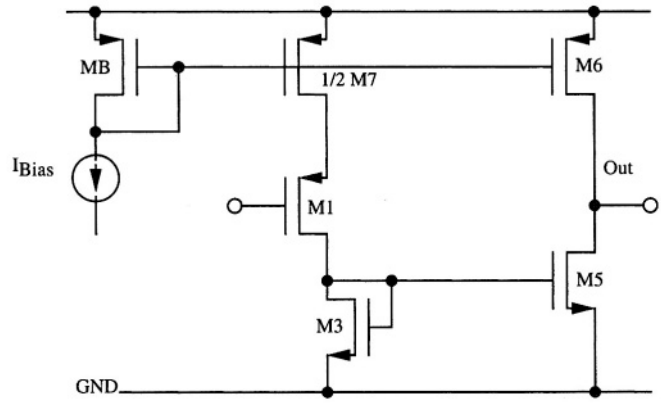


Fig. 5.13 - Circuit schematic useful for the common mode gain estimation.

equivalent circuit. However, by inspection of the circuit an approximated relationship directly results. The drain of $1/2M_7$ follows the input. Therefore, the signal current is $v_{in}/(2r_{ds7})$. Since this current flows in the diode connected transistor M_3 it results that the common-mode gain of the first stage is

$$A_{CM,1} = -\frac{1}{2g_{m3}r_{ds7}} \tag{5.10}$$

therefore, the total common mode gain is given by

$$A_{CM} = \frac{g_{m5}}{2g_{m3}r_{ds7}(g_{ds5}+g_{ds6})} \tag{5.11}$$

that leads the low frequency value of the *CMRR* to be

$$CMRR = \frac{A_d}{A_{CM}} = \frac{2g_{m3}r_{ds7}g_{m1}}{(g_{ds2}+g_{ds4})} \tag{5.12}$$

showing that the CMMR comes from the first stage design only. Such a result is obvious: the second stage amplifies in the same way any signal that it receives from the first stage; no matter if the signal results from a common mode or a differential input.

5.4.3 Offset

The supply voltages in the circuit in Fig. 5.12 are V_{DD} and ground. The analog ground, V_{AG}, is somewhere in between these two levels. Depending on the

application the designer defines the value of the analog ground used in the project. Typically it is $V_{DD}/2$. (if V_{SS} is used instead than ground, the analog ground lies midway between V_{DD} and V_{SS}). For an ideal circuit it is expected that a zero input signal (input terminals shorted) determines an analog ground voltage at the output. Instead, an improper design or possible technological mismatches in the circuit deviates the output from the expected value. Often, in open loop conditions, the output voltage saturates very close to one of the supply rails. In order to bring the output to V_{AG} it is necessary to apply a proper input signal that counterbalances the existing mismatches. This inputs signal is the offset.

We have two additive contributions to offset. They are:

- systematic offset
- random offset

The former component depends on the circuit's design. We can minimize it with suitable cell design. The second contribution comes from random fluctuation of physical and technological parameters along the chip. We can minimize it with a careful layout that minimizes the mismatch between critical components.

We estimate the systematic contribution with the following approach: the differential input is set to zero. A small signal analysis determines its effect at the output. Then, a differential input capable to counterbalance that output voltage is determined.

Connecting the inputs at the same voltage makes, as already discussed in the common mode gain evaluation, the input stage electrically symmetrical. This permits us to split the input stage into two equal parts to achieve again the circuit in Fig. 5.13. Assuming that the common mode input voltage is within the normal range of operation, *1/2* M_7 works in the saturation region and, at first approximation, its current and, consequently, the one in M_3 can be assumed a replica of the bias current, I_{Bias}.

We can regard the output stage as being made by two current generators: one (M_5) sinking the current injected by the other (M_6). Both transistors M_5 and M_6 are assumed to be in the saturation region. Since the two current sources have a high output impedance the output voltage approaches the analog ground only if the two currents I_{M5} and I_{M6} are almost equal. As M_5 mirrors the current in M_3 that in turn is a replica of the current in M_7, the following relationships must be verified

$$I_{Bias}\frac{(W/L)_6}{(W/L)_B} = \frac{I_{Bias}}{2}\cdot\frac{(W/L)_7}{(W/L)_B}\cdot\frac{(W/L)_5}{(W/L)_3} \tag{5.13}$$

that leads to the following constraint between aspect ratios of transistors

$$(W/L)_3 \cdot (W/L)_6 = \frac{1}{2} \cdot (W/L)_7 \cdot (W/L)_5 \tag{5.14}$$

The above analysis assumes the use of ideal current mirrors. Because of the transistor finite output resistance the current in one branch and its mirrored replica don't match perfectly. Thus, even if condition (5.14) is respected, the slight mismatch between the currents in M_5 and M_6 flows into the output resistance of M_5 and M_6 and causes a residual offset. The designer could trim the transistor sizing in order so as to bring the residual offset to zero with simulations. However, the practice is not advisable for two reasons: the accuracy and the process variation affecting the output resistance are both wide: zeroing results achieved with a given simulation condition are not exactly verified in the experimental verifications. A second, and more important reason, is that for optimizing the layout we need transistors with properly ratioed widths. Unrelated transistor sizing as could result from the trimming process makes the layout problematic and irregular.

If we use the condition (5.14) in (5.9) we have

$$A_v = \frac{\alpha}{I_{Bias}} \cdot \frac{2\sqrt{(W/L)_1\sqrt{(W/L)_3}}}{(W/L)_7}(W/L)_B \tag{5.15}$$

Therefore, when we take care of the systematic offset the *DC* gain can be no longer adjusted by operating over the second stage of the op-amp.

The second contribution to offset is random. Since it comes from unavoidable mismatches, we have to understand where the critical points of the design are and, possibly, concentrate on the sensitive points to improve the matching performance.

LESS *DESIGN FREEDOM*

In order to achieve a nominally zero systematic offset the relationship (5.14) between transistor sizing must be fulfilled. The condition costs the designer one of the degrees of freedom available.

The circuit has two stages, therefore, any possible mismatch will determine offset at the input of the first and the second stage (Fig. 5.14). The two offsets are uncorrelated and are combined quadratically. However, because of the gain of the first stage, when we refer the

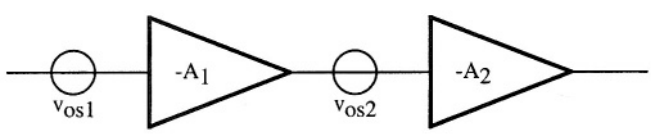


Fig. 5.14 - Input referred offset generators in a two stages amplifier

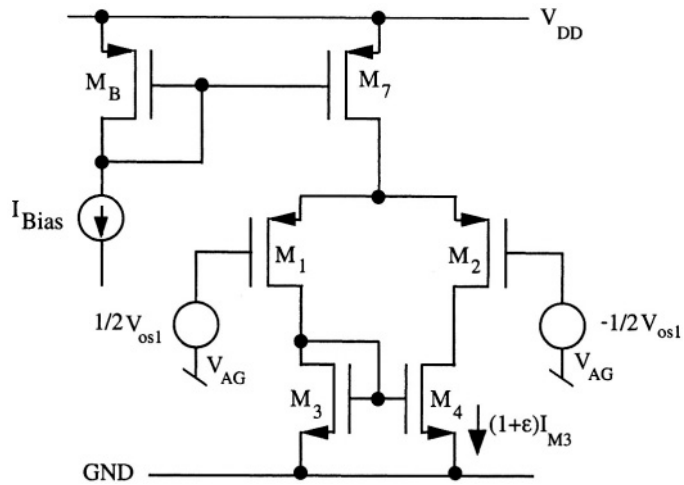


Fig. 5.15 - A mismatch in the mirror factor M3-M4 causes random offset.

offset of the second stage at the input terminal we have to divide it by the gain of the first stage to obtain

$$V_{os} = \sqrt{V_{os1}^2 + (V_{os2}/A_1)^2} \tag{5.16}$$

Assuming that the two offset contributions have similar value, since A_1 is pretty large, the input stage practically determines the random offset.

Nominally, the transistors M_1 and M_2, as well as M_3 and M_4, in Fig. 5.15 are equal. However, because of fabrication errors they show differencies. To simplify the analysis we assume that a possible mismatch between M_1 and M_2 can be transferred in a mismatch between M_3 and M_4. Therefore, we assume that $M_1$and M_2match perfectly and we study the effect on a mirror factor different to *1*, say $(1+\varepsilon)$, in the pair M_3-M_4. The current in M_4 is a mismatched replica of the current in M_1. Therefore, to make equal the current in M_4 with the one in M_2 it is necessary to suitably imbalance (by the offset voltage) the differential input. Assuming that the offset is small and that the total current in the input pair is I_{Bias} we have

$$\left(\frac{I_{Bias}}{2} - g_{m1}\frac{V_{os1}}{2}\right)(1+\varepsilon) = \left(\frac{I_{Bias}}{2} + g_{m2}\frac{V_{os1}}{2}\right) \tag{5.17}$$

that becomes, being $g_{m1} = b_{m2}$

$$V_{os1} \cong \frac{I_1}{g_{m1}} \cdot \varepsilon \tag{5.18}$$

Therefore, the voltage offset is proportional to the mismatch ε through the multiplying coefficient I_1/g_{m1}. For the input pair in saturation and in the sub-threshold conditions we have, respectively

$$\begin{aligned} \frac{I_1}{g_m} &= \frac{V_{gs1}-V_{Th}}{2} \qquad \textit{in saturation} \\ \frac{I_1}{g_m} &= nV_T = \frac{nkT}{q} \qquad \textit{in sub-threshold} \end{aligned} \tag{5.19}$$

Therefore, for a given mismatch the random offset is lower for input stages operated in sub threshold. A possible global mismatch is $\varepsilon = 0.02$ and a typical overdrive voltage in saturation is $(V_{gs} - V_{Th}) = 300\ mV$. Thus, the offset in saturation can be as large as $3mV$.

From (5.19) we note that the offset in *CMOS* circuits is always worse than for their bipolar counterpart. Since in many design the transconductance of CMOS transistors in saturation is from *5* to *10* times smaller than for equivalent bipolar counterpart, we have to expect in bipolar circuits a random offset which is *5* to *10* times smaller than *CMOS* circuits.

DESIGNER, TAKE NOTE

A large systematic offset discloses an inadequate circuit design.

A large random offset marks an improper layout.

5.4.4 Power Supply Rejection

Fig. 5.16 shows possible spur generators affecting a *p-channel* input two-stages amplifier. We assume that the source of M_3-M_4-M_5 and M_7-M_8 are physically one next to the other and a low impedance achieves their common connection. Therefore the sources of M_3-M_4-M_5 and those of M_7-M_8 are at the same voltage. By contrast, the transistor M_B can be located somewhere far from the amplifier (transistor M_B can be, for example, used to provide the bias voltage of many op-amps on the same chip). The wire connecting the source of M_B and the source of M_6-M_7 can collect some noise as represented by Δv_n^+. In addition, the spur signal on the supply lines possibly alters the current reference, I_{Ref}. The spur current source, $i_{n,Ref}$, models this effect.

We estimate the output noise voltage due to the spur generators by a small signal analysis. The various contributions are considered separately and are properly combined afterwards. Fig. 5.16 considers two possible outputs: the first, $v_{o,loc}$, applies when the output is used locally: it is referred to the source of M_5. The second one, $v_{o,far}$, is when the output signal is used far away from the op-amp location. In this case, the voltage spur affecting the wire connection

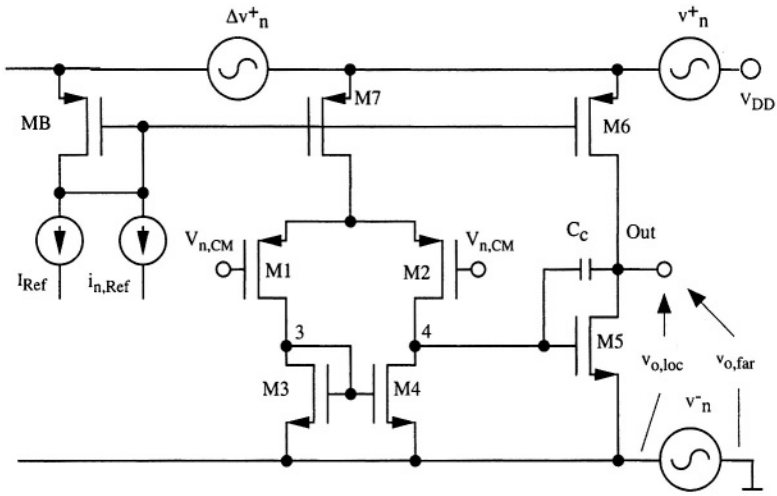


Fig. 5.16 - Two stages amplifier with possible voltage and current spur generators

from the output to the far away location should supplement the result achieved. Here we consider the local output only: the far away output differs from it by just v^-_n.

We can observe that the noise component Δv^+_n modifies the gate-to-source voltage originated by the diode connected transistor M_B

$$V_{GS6} = V_{GS7} = V_{GS,M_B} + \Delta v^+_n \tag{5.20}$$

As a result, some noise will affect the currents of M_6 and M_7. Assuming the spur Δv^+_n small enough, the noise currents are given by

$$\frac{i_{n,6}}{(W/L)_6} = \frac{i_{n,7}}{(W/L)_7} = \mu C_{ox}(V_{GS,M_B} - V_{Th})\Delta v^+_n \tag{5.21}$$

Moreover, since transistors M_6 and M_7 mirror the current in M_B, the noise terms in (5.21) are added to the ones caused by $i_{n,Ref}$. Therefore, we can combine the effect of Δv^+_n with the one of $i_{n,Ref}$. Assuming those two contributions to be uncorrelated the combination should be quadratic. However, the resulting currents affecting M_6 and M_7 are fully correlated.

The current of M_7 is equally split by M_1 and M_2. Moreover, because of the symmetry of the circuit, at low frequency, the voltage signal at node *4* equals the one at node *3*. Therefore M_5 mirror half of the spur current in M_7. The spur currents from M_5 and M_6 flow in the output node. Their superposition multiplied by the small signal output resistance lead to the low frequency output

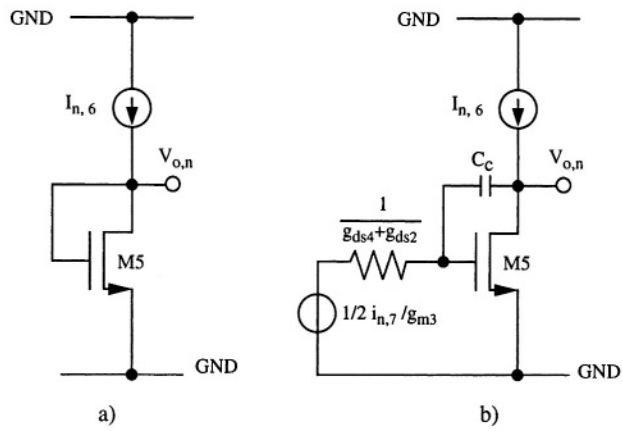


Fig. 5.17 - Small signal circuits for calculating the output spur produced by noise current

noise caused by noise on the current reference

$$v_{o,n,1} = i_{n,tot}\left[\frac{(W/L)_6}{(W/L)_B} - \frac{1}{2}\frac{(W/L)_5(W/L)_7}{(W/L)_4(W/L)_B}\right]\frac{1}{g_{ds6}+g_{ds7}} \tag{5.22}$$

Observe that if we use in (5.22) the same condition derived to minimize the systematic offset (equation (5.14)) the output noise becomes zero. Therefore, the rule for nulling the systematic offset also permits us to reject the low frequency components of $i_{n,Ref}$ and Δv^{+}_{n}. Unfortunately the compensation almost vanishes at high frequencies. Capacitor C_c, whose function will be discussed shortly, at high frequencies connects the gate and drain of M_5 making the second stage of the op-amp like the schematic shown in Fig. 5.17 a). The contribution of the first stage fades out and the second stage behaviour fully controls the output voltage. Therefore, we have

$$v_{o,n,1} = i_{n,Ref}\frac{(W/L)_6}{(W/L)_B}\cdot\frac{1}{g_{m,5}} \tag{5.23}$$

The transition between the results given by equation (5.22) and that in (5.23) occurs at frequencies for which the impedance of C_c, amplified by the Miller effect, becomes comparable to the output resistance of the first stage. Fig. 5.17 b) shows the circuit that we can use to study the issue. A voltage source and its output resistance depicts the small signal operation of the first stage. Remember that the spur currents $i_{n,6}$ and $i_{n,7}$ are fully correlated being produced by the same noise generator, $i_{n,Ref}$ (in Fig. 5.16). Analysis of the circuit leads to the transition frequency mentioned above. We don't provide here the details of calculations. They can be a useful exercise for the reader.

BE AWARE

The power supply rejection is a very difficult issue. Suitable circuit solutions provide acceptable protections at low frequency. Unfortunately, at very high frequency the circuit defences are almost vanquished because of unpredictable capacitive couplings.

In order to study the low frequency effect of v^+_n and v^-_n we furthermore exploit the symmetry properties. Assuming that any spur affects the differential inputs to the same extent $(v_{n,CM})$, for the first stage we can use the half schematic shown in Fig. 5.18 a). Fig. 5.18 b) displays its small signal equivalent circuit and also includes the small signal equivalent circuit of the second stage. The voltage at the output of the first stage, $v_{out,n1}$ is amplified by the gain of the second stage, $A_2 = g_{m5}/(g_{ds5}+g_{ds6})$. This term is superposed to the direct spur affecting the second stage.

The common-mode noise $v_{n,CM}$ comes from a possible coupling between the noisy supply lines and the input terminals. It can be represented by

$$v_{n,CM} = k_+ v^+_n + k_- v^-_n \tag{5.24}$$

where k_+ and k_- denote attenuation factors.

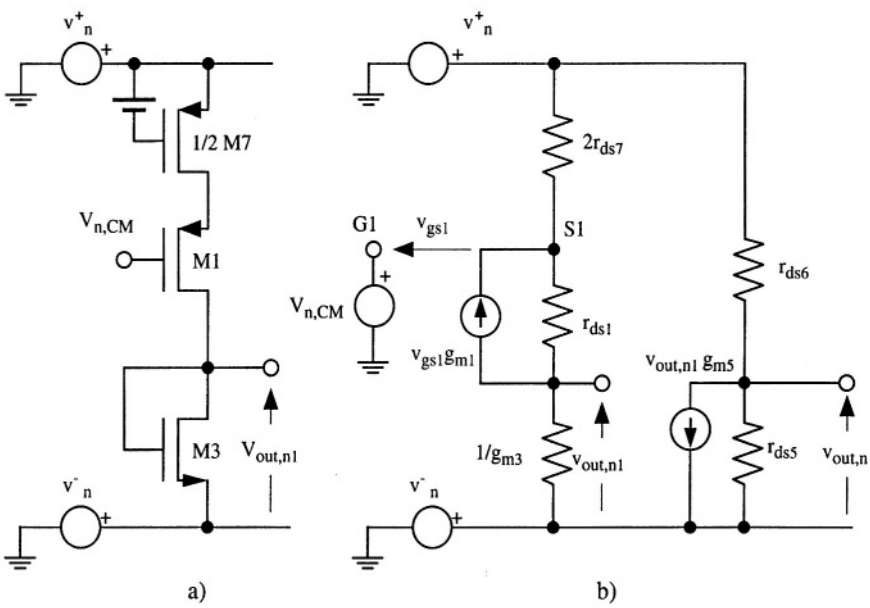


Fig. 5.18 - Circuits for calculating the effect of spur signals on the supply lines.

Therefore, we have to account for four terms: the two components that make $v_{n,CM}$, v^+_n and v^-_n. With the system being linear we can handle the problem by considering separately the contributions of the four spur generators. Then, the achieved outputs must be properly superposed. The terms produced by the same noise generator are correlated; therefore, their effects must be superposed linearly. By contrast, v^+_n and v^-_n are uncorrelated and their contributions must be superposed quadratically. The reader can do, as an exercise, the analysis using the equations that describe the small signal equivalent circuit. Here we study the problem by inspecting the circuit. Before, we shall estimate the noise currents; then, we shall calculate the effected the noise voltage.

Let us consider the first stage before. Observe that the source of M_1 follows its gate. Therefore, the spur current produced by $v_{n,CM}$ is approximated by

$$i_{n,CM} = \frac{v_{n,CM}}{2r_{ds,7}} = -\frac{k_+ v^+_n + k_- v^-_n}{2r_{ds,7}} \tag{5.25}$$

The equivalent resistance seen from the source of M_3 is the series connection of diode-connected transistor M_3 and the cascode structure $M_1-(1/2)M_7$. Since $1/g_{m3}$ is negligible with respect to the cascode resistance, we assess the noise current developed by v^-_n as

$$i^-_n = -\frac{v^-_n}{r_{ds,7}g_{m1}r_{ds,1}} \tag{5.26}$$

The equivalent resistance seen from the source of $(1/2)M_7$ is given by the drain resistance of $(1/2)M_7$ in series with the resistance seen from the source of M_1. Here, $1/g_{m1}$ is negligible with respect to $2r_{ds,7}$; therefore, the effect of v^+_n is

$$i^+_n = \frac{v^+_n}{2r_{ds,7}} \tag{5.27}$$

Note that the noise currents due to v^+_n in (5.25) and (5.27) have opposite sign.

The superposition of the four noise currents flows into $1/g_{m3}$ and develops a noise voltage given by

$$v^2_{o,n1} \cong \frac{(1-k_+)^2(v^+_n)^2 + \left(k_- + \frac{1}{g_{m1}r_{ds1}}\right)(v^-_n)^2}{(2g_{m3}r_{ds,7})^2} \tag{5.28}$$

it, amplified by $-g_{m5}/(g_{ds5} + g_{ds6})$, comes out across the drain-to-source of M_5.

The noise due to the first stage is superposed to the one given by the second stage. By inspection of the equivalent circuit in Fig. 5.18 b), one obtains

$$(v_{o,n2})^2 = \frac{g_{ds6}^2((v_n^+)^2 + (v_n^-)^2)}{(g_{ds5} + g_{ds6})^2} \tag{5.29}$$

and, superposing the effects

$$(v_{o,tot})^2 \approx \left(\frac{g_{ds6} - \frac{g_{m5}(1-k_+)}{2g_{m3}r_{ds3}}}{g_{ds5} + g_{ds6}} \right)^2 (v_n^+)^2 + \left(\frac{g_{ds6} - \frac{g_{m5}k_-}{2g_{m3}r_{ds3}}}{g_{ds5} + g_{ds6}} \right)^2 (v_n^-)^2 \tag{5.30}$$

An examination of (5.30) leads to the following observations:

- the noise from the first stage dominates the total result;
- the coupling coefficient k_+ reduces the outcome of v^+_n. It is therefore advisable to tight the common-mode input with the positive supply voltage (for p-channel input pair);
- the coupling coefficient k_- materializes a noise term. It is therefore advisable to well disjoint ground from the common-node input (for p-channel input pair);
- for n-channel input pair the two above recommendations must be switched.

High frequencies require a more simple study: the compensation capacitor establishes a connection from the output of the second stage to its input. Therefore, the contribution of the first stage vanishes. The spur from the supply lines is transferred to the output through the network made by $1/g_{m5}$ and r_{ds6}. Therefore, at high frequencies, it yields

$$(v_{o,loc})^2 = [(v_n^+)^2 + (v_n^-)^2]\left(\frac{1}{1 + g_{m5}r_{ds6}}\right)^2 \tag{5.31}$$

At high frequencies the local output rejects of the supply noise significantly because of the connection established by the compensation capacitor. However, this happens at frequencies around the gain-bandwidth product.

5.4.5 Effect of External Components on the *PSRR*

Power supply rejection additionally depends on the circuit that uses the op-amp. To discuss the issue, let us consider the scheme in Fig. 5.19. It illustrates a simple switched capacitor integrator. The input signal is injected by a capac-

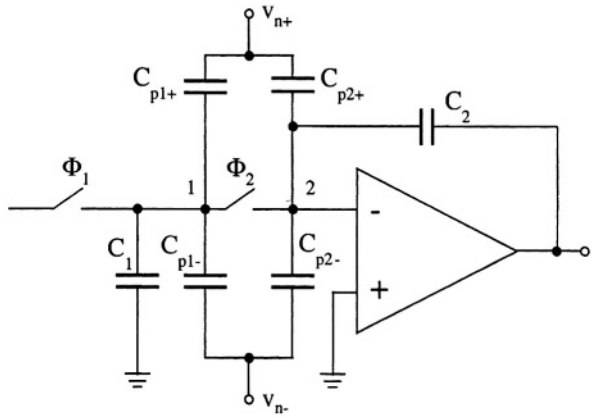


Fig. 5.19 - Spur signal injection caused by external parasitic elements.

itor C_1 and two switched driven by complementary non-overlapped phases. During phase *1* an input voltage charges capacitor C_1. During phase 2 the action of the op-amp transfer the charge on C_2. Four parasitic capacitances represent possible parasitic couplings between node *1* and *2* (the virtual ground) and the power supply lines. The parasitic capacitances are, in general, non linear elements often associated to *p-n* junctions. They include the parasitic given by the switches and the parasitic coupling created by the input transistors of the op-amp.

The spur signal at the output of Fig. 5.19 can be easily estimated. The result is different when the injecting switch is *on* (phase 2) or when it is *off (phase* 1). In the two cases we obtain, respectively

$$(v_{o,n})^2 = (v_{n+})^2 \cdot \left(\frac{C_{p1+} + C_{p2+}}{C_2}\right)^2 + (v_{n-})^2 \cdot \left(\frac{C_{p1-} + C_{p2-}}{C_2}\right)^2 \qquad (5.32)$$

$$(v_{o,n})^2 = (v_{n+})^2 \cdot \left(\frac{C_{p2+}}{C_2}\right)^2 + (v_{n-})^2 \cdot \left(\frac{C_{p2-}}{C_2}\right)^2 \qquad (5.33)$$

Therefore, depending of the phase in which we sample the output, we have a larger or smaller coupling with the spur supply lines.

All the components' parasites must be reduced to a minimum. In particular, it is essential to minimize the coupling between the connection to the virtual ground and the substrate. Long interconnections collect significant spur signals because of their substrate coupling. When long wiring to the virtual ground is necessary it is recommended to use electrically shield, achieved by a proper use of the metal layers available.

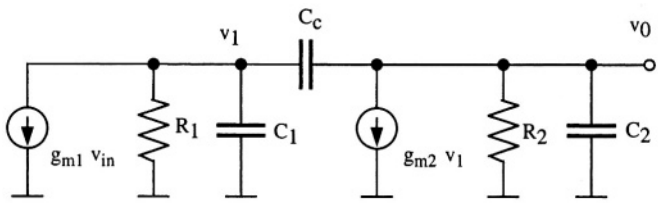


Fig. 5.20 - Small signal equivalent circuit of a two stages OTA.

5.5 FREQUENCY RESPONSE AND COMPENSATION

The small signal equivalent circuit of the two stage amplifier in Fig. 5.12 can be represented with the simplified diagram in Fig. 5.20. Each stage is represented by a transconductance generator and the parallel connection of an output resistance and a load capacitance. The *dc* gain is given by the product of the two gains $g_{m1}R_1$, $g_{m2}R_2$. The two *RC* networks contribute with two poles whose angular frequencies are

$$p'_1 = \frac{1}{\tau_1} = \frac{1}{R_1 C_1} \tag{5.34}$$

$$p'_2 = \frac{1}{\tau_2} = \frac{1}{R_2 C_2} \tag{5.35}$$

We know that the output resistances of the two stages are given by the parallel connection of two r_{ds}. The load capacitances result from the parasitic elements of the transistors used and, possibly, the capacitance loading the output. It turns out that the two time constants of the first and the second stage are not significantly different.

However, since the transfer function has two poles, we can ensure stability only if the poles are sufficiently far apart that when the module of the gain becomes *1* the second pole has not been achieved yet. The first pole is called the *dominant pole* since it dominates the frequency response behaviour in the region where the gain is larger than one.

Unfortunately, in the two stage amplifier, the poles p'_1 and p'_2 are relatively close one to the other. The circuit does not have a dominant pole and *compensation* becomes necessary.

One possible (and widely used) compensation network is achieved by the capacitor C_c connected between input and output of the second stage (see Fig. 5.20). We account for its effect in the small signal transfer function by

$$v_1(g_1 + sC_1) + (v_1 - v_0)sC_c + g_{m1}v_{in} = 0 \tag{5.36}$$

$$v_0(g_2 + sC_2) + (v_0 - v_1)sC_c + g_{m2}v_1 = 0 \tag{5.37}$$

they are the node equations at the output of the first and the second stage. Moreover, $g_1 = 1/R_1$ and $g_2 = 1/R_2$. Solving and leads to the frequency response

$$\frac{V_0}{V_{in}} = g_{m1}R_1R_2\frac{g_{m2} - sC_c}{1 + sR_1R_2g_{m2}C_c + s^2R_1R_2[C_1C_2 + (C_1 + C_2)C_c} \tag{5.38}$$

that shows one zero and two poles. The position of the poles is approximately given by

$$p_1 \cong \frac{-1}{g_{m2}R_2R_1C_c} \tag{5.39}$$

$$p_2 \cong \frac{-g_{m2}C_c}{C_1C_2 + (C_1 + C_2)C_c} \tag{5.40}$$

and the zero is located at

$$z = +\frac{g_{m2}}{C_c} \tag{5.41}$$

therefore, the action of the compensation capacitor C_c on the poles is twofold. Pole p_1 is pushed at low frequency. In fact, it is $g_{m2}R_2C_c/C_1$ lower than p'_1 in (5.34). Pole p_2 is brought to a high frequency: it is approximately $g_{m2}R_2(C_2/C_1+C_2)$ times p'_2. The double action on the poles is called *pole splitting.*

TWO STAGES OTA COMPENSATION

A capacitor in feedback around the second stage splits the poles. The pole of the first stage goes at low frequency and the one of the second stage is pushed at high frequency. The action is named pole splitting.

To better memorize the achieved result we observe that the action of the compensation capacitor can be analysed by using the Miller theorem. Since the gain of the second stage is $g_{m2}R_2$, capacitor C_c is amplified by $1+g_{m2}R_2$ when transferred across the output of the first stage and ground and remains almost unchanged when transferred to the output of the second stage. Therefore the capacitive load of the first stage becomes $(C_1+g_{m2}R_2C_c)$ and the one of the second stages slightly increased: (C_2+C_c). Moreover the compensation capacitor establishes a negative feedback around the second stage of amplification. Therefore, as we know from basic

electronic courses, the bandwidth of the second stage is enlarged by the loop gain.

The two poles of the transfer function are in the left s-plane, as result from the minus sign in (5.39) and (5.40). By contrast the zero is in the right s-plane. Therefore, the phase shift produced by the zero will be negative like a pole in the left s-plane. As a result the phase shift of the zero does not improve the phase margin but instead, worsens it. Consequently, the zero can be a problem if it is located close to the unity gain frequency. Assuming that the first pole is dominant, multiplying it by the *dc* gain obtains the expected unity gain angular frequency

$$\omega_T = 2\pi f_T = \frac{g_{m1}}{C_c} \tag{5.42}$$

by comparing (5.41) with (5.42) we see that the ratio between the zero and ω_T is equal to the ratio of the transconductance gain of the second and the first stage. Therefore, the zero is far away from the expected unity gain frequency if the transconductance of the second stage is much higher than the one of the first stage. With a *CMOS* circuit it is difficult to achieve large differences between g_{m2} and g_{m1}: we have an increase in g_m with the square root of the current and the aspect ratio. By contrast, with a bipolar technology the transconductance is proportional to the bias current. Therefore, in bipolar implementations g_{m2} can be designed suitably larger than g_{m1}, while with *CMOS* we cannot achieve the same result easily. If the zero is near the expected ω_T, its effect modifies significantly the frequency response near the *zero dB* crossing degrading the stability conditions. Thus, in practical cases, we cannot leave the circuit as it is, but we need to find some remedy to the unsatisfactory phase margin.

REMEMBER THESE EQUATIONS

Dominant pole:

$$f_1 = \frac{1}{2\pi} \frac{1}{g_{m2} R_2 R_1 C_c}$$

Unity gain frequency:

$$f_T = \frac{1}{2\pi} \frac{g_{m1}}{C_c}$$

Zero in the right s-plane:

$$f_1 = \frac{1}{2\pi} \frac{g_{m2}}{C_c}$$

The problem of the zero in the right s-plane can be solved by three techniques:

- use of a unity gain buffer
- use of the zero nulling resistor
- use of a unity gain current amplifier

If we examine equations (5.36) and (5.37) we identify the term $-sC_c v_1$ in

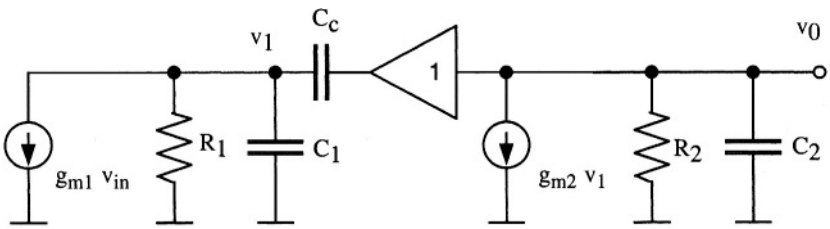


Fig. 5.21 - Use of a unity gain buffer to cancel the right-plane zero

(5.37) as the responsible that lead to the zero. If we are able to eliminate such a term the zero disappears from the solution. We achieve the result with a unity gain buffer connected to the output node and feed the compensation capacitor (Fig. 5.21). The second term in (5.37) disappears, as the current from the output node to the compensation capacitor is provided by the buffer. Actually, in addition to the term $-sC_c v_1$, we also eliminate $sC_c v_0$; therefore, we will have some changes in the pole positions. If we use instead of (5.37) the relationship

$$v_0(g_2 + sC_2) + g_{m2}v_1 = 0 \tag{5.43}$$

the solution of the new system made by (5.36) and (5.43) becomes

$$\frac{V_0}{V_{in}} \cong \frac{-g_{m1}g_{m2}R_1R_2}{1 + sR_1R_2g_{m2}C_c + s^2R_1R_2(C_1 + C_2)C_c} \tag{5.44}$$

whose denominator is only slightly different from that in (5.38). Therefore, the poles will approximately remain unchanged but the zero disappears (5.37).

Fig. 5.22 shows the possible circuit implementation of the technique discussed above. It is applied for an *n-channel* input transistors architecture. In the right side circuit a source follower achieves the required buffer. Note that, to operate properly, the buffer needs at least $V_{th}+2V_{sat}$ at its input. If the output voltage is lower than this limit the current source is pushed into the triode region and the buffer does not operate properly any more. Therefore, the use of a unity gain buffer would limit the achievable output swing of the two stages op-amps.

Another important point to observe is that proper operation of the buffer matters around the unity gain frequency of the op-amp: the compensation network must correct an undesired behaviour at high frequency. At very high frequency the buffer can suffer some frequency limitations with a consequent phase shift from input to output. The resulting effect is that the zero is not eliminated but a doublet zero-pole, each close to the other, appears in the transfer function. The doublet (in the left *s*-plane) is not particularly problematic. How-

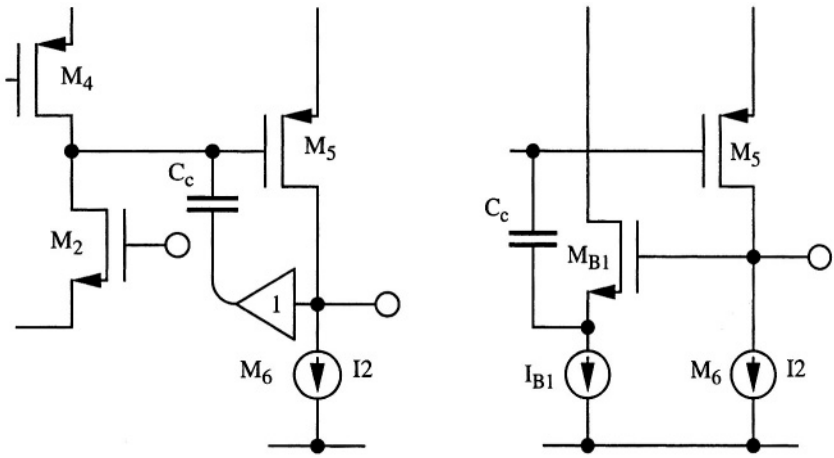


Fig. 5.22 - Eliminating the right-plane zero with a unity gain buffer Left use of a source follower.

ever, under given feedback factors it can be responsible for a worsening of the phase margin. In addition to the above limitation, we have to account for an obvious increase in power consumption and chip area.

CALL UP

The unity gain buffer possibly used in the compensation network must be very fast. It should work properly at frequencies higher than the unity gain frequency. Otherwise the right-half plane zero is not canceiled but it is replaced by a doublet pole-zero.

The second technique that handles the problem of the zero in the right *s*-plane is the zero nulling. The compensation network is made, instead of a capacitor, by an impedance: it is the series connection of a resistance and the compensation capacitance. Fig. 5.23 show the small signal schematic. If we use this kind of solution we have to replace sC_c in equations (5.36) and (5.37) with

$$sC_c \Rightarrow \frac{sC_c}{1 + sR_2C_c} \tag{5.45}$$

The solution of the modified equations (5.36) and (5.37) does not show any significant change in the denominator of the transfer function. Thus, the poles are substantially unchanged. By contrast the zero position is modified into

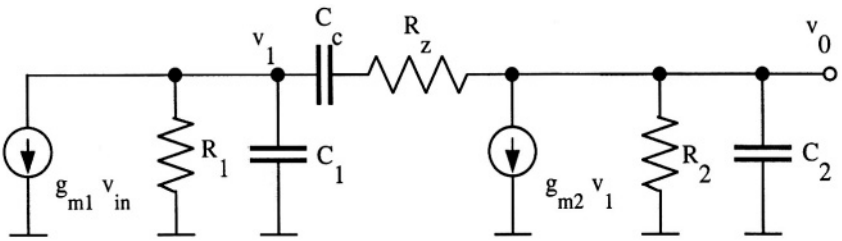


Fig. 5.23 - Use of the zero nulling resistor to move the right-plane zero.

$$z = \frac{1}{C_c\left(\frac{1}{g_{m2}} - R_z\right)} \tag{5.46}$$

We can observe that depending on the sign of $(1/g_{m2} - R_z)$ the zero is in the right or left side of the s-plane. Moreover for $(1/g_{m2} - R_z) = 0$ the zero goes to frequency ∞, or as we normally say, it is nulled. Having a zero in the left s-plane can be attractive and possibly, achieving a pole-zero cancellation we can enlarge the op-amp bandwidth. Nevertheless, it is not advisable to try this design avenue. The accuracy of integrated resistors is not particularly high; thus, a pole zero cancellation is never achieved in practical cases. Instead, it is advisable to use a value of R_z that nominally accomplishes the nulling operation. Possible variations due to technological changes will move the zero to around infinite in the positive or negative half plane but, always sufficiently far away from the critical unity gain frequency.

The zero nulling resistance should match the inverse of the transconductance gain of the second stage that, in turn, depends on the electrical and technological parameters. A possible way to track some of the parameter variations is to

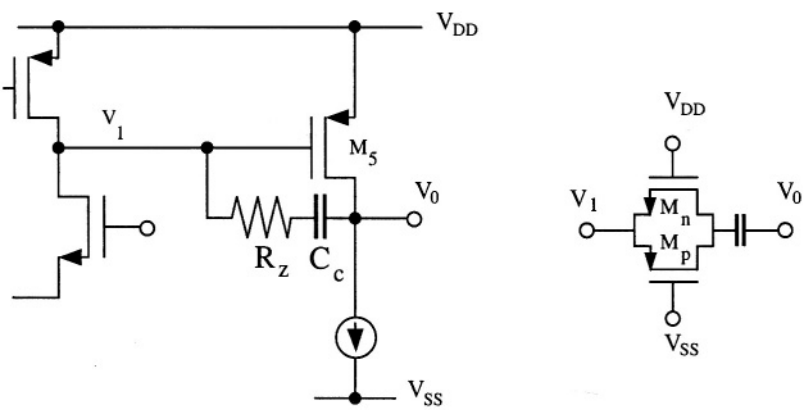


Fig. 5.24 - Zero nulling resistor achieved with complementary transistors.

use one or two transistors to realize the zero nulling resistor. Since they are in series with the compensation capacitor, the *dc* current is zero. Therefore, the transistor(s) is (are) in the triode region. Assuming we use the compensation network in right side of Fig. 5.24 the resistance R_z will be expressed by

$$R_z = \frac{R_n R_p}{R_n + R_p} \tag{5.47}$$

where

$$\frac{1}{R_n} = k'_n \left(\frac{W}{L}\right)_n [V_{DD} - V_1 - V_{Th,n}] \tag{5.48}$$

$$\frac{1}{R_p} = k'_p \left(\frac{W}{L}\right)_p [V_1 - V_{SS} - V_{Th,p}] \tag{5.49}$$

ZERO-NULLING TECHNIQUE

The resistance that cancels the zero in the right-half plane is the inverse of the transconductance gain of the second stage.

The best strategy is just to send the zero at infinite and not trying to enlarge the bandwidth by some pole-zero cancellation.

observe that the two transistors implementing the zero nulling resistor are connected toward the node V_1. From the electrical point of view it is irrelevant if we connect the zero nulling resistor on the V_1 side or on the V_0 side. However, we have a significant practical difference. Since we have amplification between V_1 and output, the voltage of V_1 will show little variation with respect to the output. Therefore, if we put the equivalent resistor on the V_1 side, the variation in R_n and R_p will be kept to a minimum. We should note that the *dc* level of V_1 is below V_{dd} by the gate to source voltage of M_5. Therefore, when we use the equivalent resistance on the V_1 side we have

$$V_{gs,n} = V_{gs,5} \tag{5.50}$$

Transistor M_5 is a *p-channel* device while M_n is a *n-channel* one. Remembering that the thresholds of the n-channel and *p-channel* can be different, and their technological variation uncorrelated, we observe that in nominal or worst case situations the voltage $V_{gs,5}$ cannot be enough to drive M_n properly. Therefore, when using the equivalent R_z on the V_1 side, many designers prefer to employ only one *p-channel* transistor (they use an *n-channel* for a complementary architecture).

If we put the equivalent resistor on the V_o side, when the voltage of the tran-

sistor channels changes, the resistance R_n moves in one direction while R_p moves in the opposite direction. Consequently, if

$$k'_n\left(\frac{W}{L}\right)_n = k'_p\left(\frac{W}{L}\right)_p \tag{5.51}$$

we can keep constant R_z by compensating the two changes.

$$\frac{1}{R_z} = k_n\left(\frac{W}{L}\right)_n [V_{DD} - V_{SS} - V_{Th,n} - V_{Th,p}] \tag{5.52}$$

However, since the condition (5.51) relies on the accuracy of technological parameters, the above benefit can be difficult to achieve in practice.

Example 5.1

Design a two stages OTA having more than 70 dB dc gain and 70 MHz GBW. The phase margin must be higher than 55° with a 2 pF output stage load. Use a zero nulling compensation network. The supply voltage is 1.8 V and the target power consumption is 0.6 μW. The desired output swing is 0.3 1.5 V. Use the transistor models given in Appendix C.

Solution: *The total bias current permitted by the specifications is around 300 μA. We assume that the power budget doesn't include the power required by the bias network. Moreover, we suppose that the current in the second stage is twice the one in the input differential pair. Accordingly, the current in M_7 (see below diagram) becomes 100 μA and the one in M_6 200 μA.*

The load of the OTA is 2 pF capacitor. This value is pretty large compared to all the parasitics affecting the output of the first and the second stage. Therefore, in the small signal equivalent circuit in

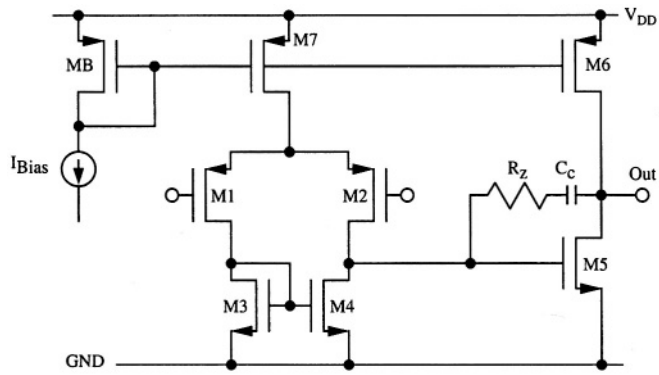


Fig. 5.23 it is allowable to use the condition $C_1 \ll C_2$. Equations (5.40) and (5.42) state

$$\omega_T = \frac{g_{m1}}{C_C} \ll \omega_2 \approx \frac{g_{m2}}{C_2}$$

where the inequality ensures a proper phase margin. Since the transconductance of the input pair is typically comparable to the one of the second stage, the compensation capacitance must be equal or, when necessary, larger than the load capacitance. Using $C_C = 2\ pF$ we have

$$g_{m1} = 2\pi f_T C_C = 6.28 \cdot 70 \cdot 2 \cdot 10^{-6} = 0.879 \cdot 10^{-3} \Omega^{-1}$$

The minimum feature of the technology is 0.25 μm. However, since the speed specification doesn't impose to go to the minimum allowed, a 0.4 μm length in input pair provides some safety margin. The Spice model parameters gives $\mu_p = 120\ cm^2/V\text{-}sec$ and $t_{ox} = 7.5\ nm$ ($C_{ox} \approx 4fF/\mu^2$). The use of the relationship

$$g_m = \sqrt{2\mu C_{ox} \frac{W}{L} I_D}$$

leads to $(W/L)_1 = 154$. However, a Spice simulation reveals that the above aspect ratio leads to a transconductance lower than expected ($0.62 \cdot 10^{-3}\ \Omega^{-1}$). This is likely due to the limits of the approximate expression of g_m used. Spice simulations show that the transistor width capable to achieve the requested g_m using 50 μA bias current is W=160 μm, (W/L)= 480. The g_m becomes equal to $0.86 \cdot 10^{-3}\ \Omega^{-1}$. The saturation voltage is 94 mV. Observe that the use of the relationship $g_m = 2I_D/V_{sat}$ produces $g_m = 1.07 \cdot 10^{-3}$ not much different from the simulated result. Thus, the second expression of g_m is more accurate than the previous one.

The length of the active loads, M_3 and M_4 should be longer than the input pair (we derived such condition for the optimization of the noise performance of the inverter with active load. It hold more in general and we use it here). We choose $L_3 = L_4 = 0.8$ μm. The width of M_3 and M_4 are designed according to the following observations: the width should lead to a reasonably low overdrive voltage; moreover, the sizing and the current in M_5 are related to the ones of M_3 and M_4. Since g_{m5} should be large, g_{m3} and g_{m4} should be reasonably large as well. In addition the overdrive of M_5 must be low enough to accommodate the required output swing (saturation 0.2 V). The above thoughts and Spice Simula-

tions lead to $W_3=W_4 = 40\ \mu m$.

Next step of the design process involves M_6 and M_7. They provides the bias current in the two stages and, in addition, M_6 works as the active load of the second stage. Possible design guide-lines are the following:

- the output resistance of M_6 must be high
- the overdrive voltages must be low (around 0.2 V)
- the length of the two transistor must be the same (to ensure matching)
- the current in M_6 is twice the one in M_7.

Using the above propositions and Spice simulations we can end on the following transistor sizing: $L_{6,7} = 0.6\ \mu m$; $W_6= 200\ \mu m$ $W_7 = 100\ \mu m$.

The aspect ratio of M_5 is defined by the condition of zero systematic offset. Therefore, using the same length of M_4 we have $L_5= 0.8\ \mu m$ and $W_5= 160\ \mu m$.

Observe that all the designed widths are round numbers. Attempts to optimize some feature can lead to widths with fraction of microns. That is not advisable because the layout would become problematic.

A .op and a . ac simulation of the circuit using the designed transistor sizing lead to the following results:

- the transconductance of the input pair is $0.85\cdot10^{-3}\Omega^{-1}$ (as expected)
- the transconductance of M_5 is as high as $3.2\cdot10^{-3}\Omega^{-1}$
- the saturation voltages of all the transistors is below 0.2 V
- the low frequency gain is 73 dB (satisfactory)
- the residual systematic offset is only 0.14 mV

Therefore, simulations confirm the validity of the design as far as the operating point and low frequency gain is concerned. The simulated output conductances are $g_{ds1}=15.31\cdot10^{-6}$; $g_{ds3}=2.28\cdot10^{-6}$; $g_{ds5}=9.81\cdot10^{-6}$; $g_{ds6}=121.55\cdot10^{-6}$. The use of these figures in (5.8) leads to a dc gain of 74 dB, just a bit more than the simulated value.

Finally we have to study the frequency compensation. The value of g_{m5} determine the zero nulling resistor: $R_Z =320\ \Omega$; with a Spice simulation we verify that a compensation capacitance of 2 pF leads to $f_T = 100$ MHz but the phase margin is only 45°. This result comes from the high value of g_{m5}. The low phase margin requires to increase the compensation capacitance. Increasing it to 3 pF leads to $f_T = 75$ MHz and phase margin 57°. We can observe that the sec-

ond pole is at 107 MHz (where the phase is margin (45°), it is about 1.5 times larger than the unity gain frequency.

The third technique used to control the right s-plane zero is again based on the observation that the zero comes from the $(-sC_c v_1)$ component of the current flowing from the output node to the node *1*. We cancel that component by using instead of a unity gain buffer a unity gain current amplifier as shown in Fig. 5.25.

The matched current generator I_{comp}, compensates the current injected into node *A* thus ensuring that the biasing condition of the first stage is not disturbed. Moreover, transistor M_c establishes a low impedance at node *A*. Therefore, the small signal current flowing in the compensation capacitance approximately results

$$i_c = v_0 s C_c \tag{5.53}$$

and, in particular is not affected by the voltage at node *1*. Because of the low impedance at its source transistor M_c delivers the signal current i_c to the node *1*, acting like a current buffer. Equations (5.36) and (5.37) are slightly modified in

$$v_1(g_1 + sC_1) + g_{m1} v_{in} - v_0 s C_c = 0 \tag{5.54}$$

$$v_0(g_2 + sC_2) + g_{m2} v_1 + v_0 s C_c = 0 \tag{5.55}$$

where we neglect the resistance seen from the source of M_c.

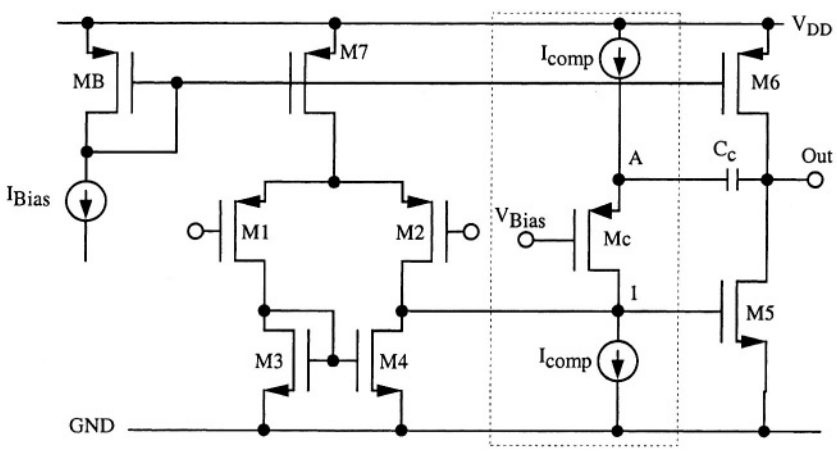


Fig. 5.25 - Cancellation of the right-plane zero by using a unity gain current amplifier.

The solution of (5.54) and (5.55) (that can be done by the reader as a useful exercise) reveals that the transfer function is no longer affected by a zero and that the two poles are approximately located as in the previously discussed solutions. In designing the bias voltage V_{bias} we have only to account for the biasing needs of transistor M_c and the current generators used. In order to avoid pushing the current generator on the upper part of the schematic in the triode region, we should use $V_{bias} \leq V_{DD} - V_{Th,p} - 2V_{sat}$.

Example 5.2

Design the unity gain current amplifier used to compensate a two stages amplifier, as shown in Fig. 5.25. Use the same transistor sizing derived in Example 5.1. Assume ideal the current generators required.

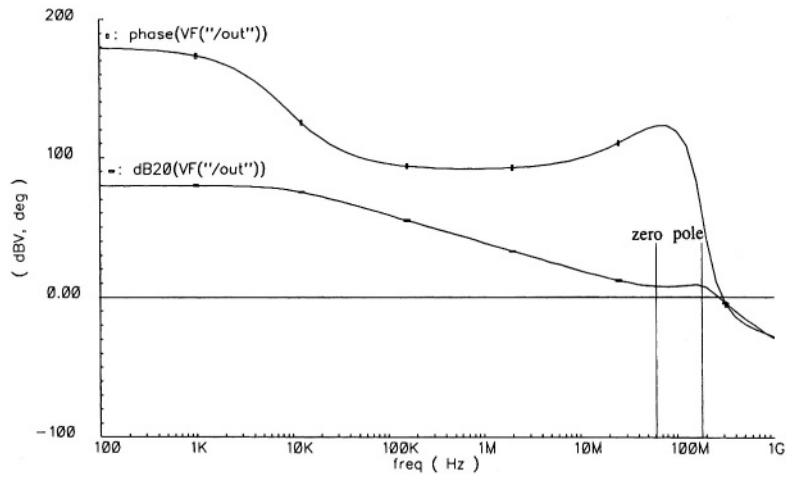
Solution: *The effectiveness of the technique rely on the use of a good unity gain current amplifier: the input conductance should be pretty high with respect to the admittance of the compensation capacitance at frequencies around the GBW. If not, the current in the compensation capacitor will exhibit a phase shift that, in turn, will lead to a zero-pole doublet. Assuming that the compensation capacitance is 2 pF, at 70 MHz the capacitive admittance is*

$$\omega C_C = 2 \cdot \pi \cdot 70 \cdot 10^6 \cdot 2 \cdot 10^{-12} = 8.79 \cdot 10^{-4} \Omega^{-1}$$

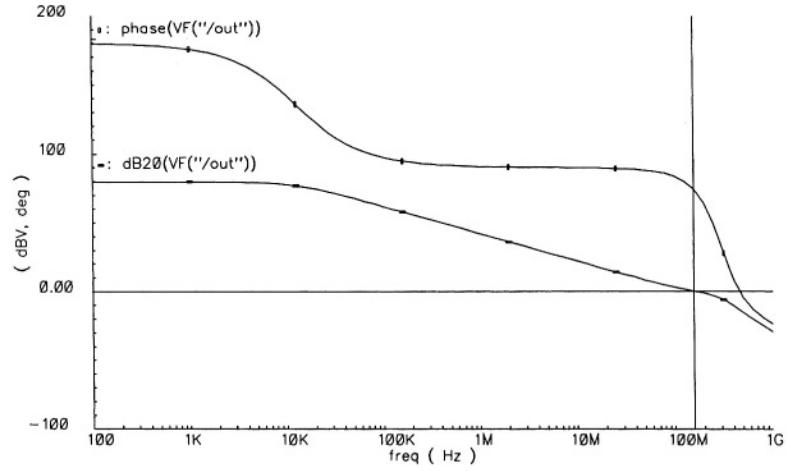
The simulation done in the Example 5.1 showed that the p-channel transistor M_6 ($W/L = 200\mu m/0.6\mu m$, I_D=200 μA) showed a transconductance $g_m = 21 \cdot 10^{-4} \Omega^{-1}$. therefore, we expect that in the current buffer it will be necessary to use such a pretty large aspect ratios and current level.

In order to investigate the effect of a limited input conductance it is helpful doing a simulation with, for instance, a p-channel transistor with aspect ratio $W/L = 200\mu m/0.4\mu m$ and a bias current of only 50 μA. The figure in the next page shows the Bode diagrams of the small signal response. Observe that the relatively low input conductance at the source of M_C produces the expected zero-pole doublet; the zero is in the left s-*plane (as shown by the phase behaviour) and it is located around 65 MHz; the pole is near 170 MHz. The effect of the doublet is not beneficial because the zero lies before the unity-gain frequency. It sustains the gain above 0 dB and when the non-dominant poles occur the frequency response displays a fast phase drop.*

In order to push the doublet at a frequency higher than the unity gain bandwidth it is necessary to increase the g_m of M_C. Bringing



I_{comp} to 250 μA, we obtain the Bode diagram shown in the below Figure. The zero-pole doublet is moved at high frequency and it doesn't affect the Bode diagram any more. The reader can acquire familiarity with the circuit and the technique with some Spice simulations. The result shown in the Figure achieves a 72° phase margin. The compensation capacitance is 2 pF and the unity gain frequency is as good as 159 MHz. The superior performances of the circuit compared to the zero nutting resistor technique is paid with a significant increase of power consumption.



5.6 SLEW RATE

A large signal applied to the input pair causes a complete umbalancement of the stage: the bias current completely flows through M_1 or M_2. Under such conditions the feedback network loose the control of the output node; the swing of the output voltage depends on the load capacitance and the compensation capacitance, that must be both charged or discharged, and the current available for this operation in the first and the second stage of the amplifier.

Let us assume the input pair unbalanced so that the current I_{M7} totally flows through M_2 (Fig. 5.26); M_1 is off. M_3 is consequently off and M_4 is off as well since it mirrors the current in M_3. As a result, the entire current I_{M7} flows through the compensation network. Therefore, current I_{M7} discharges capacitor C_c with a constant pace. The voltage across C_c decreases and, if the voltage of the drain of M_2 remains constant, the output voltage goes down with a slew rate given by

$$SR- = \left.\frac{\Delta V_-}{\Delta t}\right|_{max} = -\frac{I_{M7}}{C_c} \tag{5.56}$$

The discharge current at the other plate of C_c flows through M_5. The voltage of the drain of M_2 drops-up by a given extent so that the current furnished by M_5 increases and likely provides the current discharge of both C_c and C_L. If the increase of I_{M5} is not large enough the operation slows down. Therefore, the slew rate expressed by (5.56) is the maximum achievable.

In the other symmetrical case, when the current I_{M7} flows throw M_1, M_4 mirrors the I_{M7} current while M_2 is off. Therefore the I_{M7} current from M_4 flows comes from the compensation network. Under the assumption that the drain of M_2 jumps down by the overdrive of M_5, transistor M_5 turns off and the current

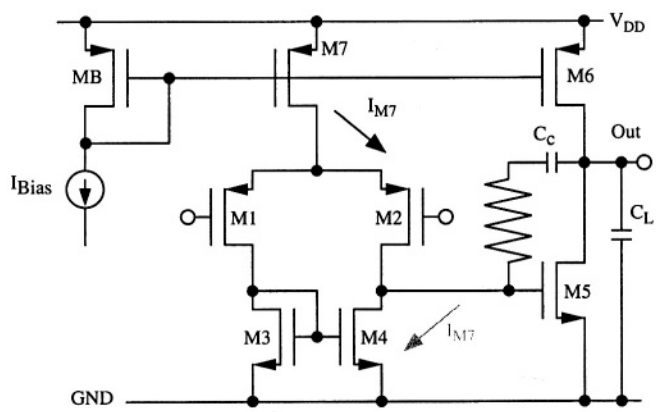


Fig. 5.26 - OTA in the positive and negative slew rate conditions.

I_{M6} is fully available to charge C_c and C_L. Therefore

$$SR_+ = \left.\frac{\Delta V_+}{\Delta t}\right|_{max} = \frac{I_{M6}}{C_c + C_L} \tag{5.57}$$

In addition, we have the current limit through C_c. It gives $SR_+ = I_{M7}/C_c$. So, the smaller one will hold. Note that the above equations are an approximation: M_5 does not switch on or off immediately and the transient of this process moderates the output node slewing up and down.Therefore, the use of equations (5.56) and (5.57) provide just guidelines in the preliminary design phase. Computer simulations achieve a more accurate estimation.

Often it is necessary to have symmetrical positive and negative slewing. This design requirement leads to the condition

$$\frac{I_{M7}}{C_c} \leq \frac{I_{M6}}{C_c + C_L} \tag{5.58}$$

that establishes a design constraint over the choice of the bias current of the first and second stage.

Since $\omega_T = \frac{1}{\tau_T} = \frac{g_{m1}}{C_c}$, the slew-rate can be expressed as

$$SR = \frac{I_{M1}}{g_{m1}}\omega_T = (V_{GS1} - V_{Th})\omega_T = \frac{V_{GS1} - V_{Th}}{\tau_T} \tag{5.59}$$

REMEMBER THAT

A symmetrical slew-rate requires a current in the second stage larger than the tail current of the input pair by the factor $(C_C+C_L)/C_C$.

therefore, the slew-rate is proportional to the overdrive voltage of the input pair and the unity gain frequency. If f_T is *40 MHz* and the overdrive is *300 mV* the slew rate becomes *75.4 V/μs*, Moreover, using the same figures, the time required for a *1 V* slewing turns out to be *3.3* times τ_T.

Example 5.3

Given the op-amp designed in Example 5.1, determine with a Spice simulation the positive and negative slew-rate. Use the OTA in the open-loop conditions and apply a large step signal at the inputs. Compare the achieved result with what expected using the approximate expressions (5.56), (5.57) and (5.59).

Solution: *Example 5.1 uses a capacitive load* $C_L = 2\ pF$ *and a compensation capacitance* $C_c = 3\ pF$. *Moreover, the bias current*

in the first stage is $I_{M7} = 100\ \mu A$ *and the current in the second stage is* $I_{M6} = 200\ \mu A$. *The use of equation (5.58) predict that the slewing in the positive and negative directions are both limited by the* $SR = I_{M7}/C_c$ *condition. In fact, using (5.56), (5.57) we obtain*

$$SR- = -\frac{I_{M7}}{C_c} = 33V/\mu sec \qquad SR_+ = \frac{I_{M6}}{C_c + C_L} = 40V/\mu sec$$

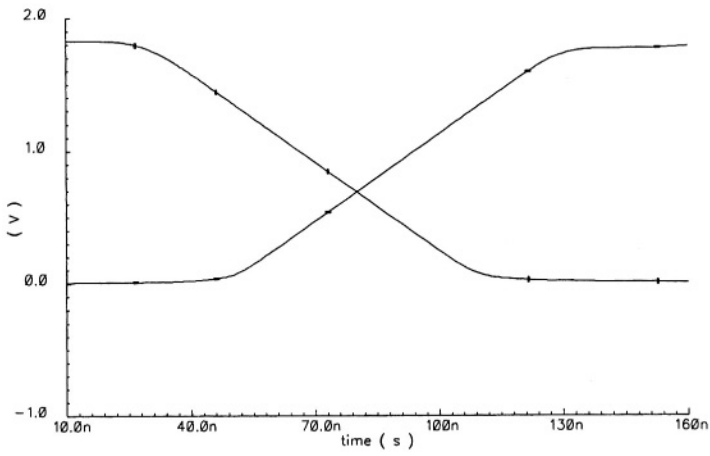
The current of the input transistors is $50\ \mu A$ *while their* g_m *is* $0.88 \cdot 10^{-3}\ \Omega^{-1}$. *Since the* f_T *achieved in Example 5.1 is 75 MHz, the use of (5.59) predicts*

$$SR = \frac{I_{M1}}{g_{m1}} 2\pi f_T = 23.56V/\mu sec$$

which is lower than the value predicted by the use of equations (5.56) and (5.57).

A time-transient simulation with a big step (± 100 mV) applied to the inputs lead to the responses shown in the below figure. Since the OTA is in the open loop conditions we observe only the slewing transients: they bring the output from the positive to the negative (ground) rails and vice-versa. The two slopes are, as expected, equal and their value is $22\ V/\mu sec$. *This value is not much different from the results of the hand calculations. However, the use of equation (5.59) leads to a more accurate forecast.*

Once again the reader should be aware that the equations derived in this book are the result of many approximations. It should not be surprisingly if the difference with simulation results are as large as



50% or more.
The transient responses in the figure mark some delay from the time at which the impulse is applied and the time the output starts slewing. Since the impulse start from 100 mV (or -100 mV) the initial condition of the output is a strong saturation at 0 or 1.8V. The transition between saturation and normal condition of operation leads to some delay. For the specific circuit studied the delay required to recover from zero is higher than the one required to exit from the V_{DD} saturation.

5.7 DESIGN OF A TWO STAGE *OTA*: GUIDELINES

The study of the features of a two-stage *OTA* determined a number of design recommendations. Some of them are conflicting one each other. So, the designer should find the best trade-off to match the design specifications. Some recommendation are statement other are expressed by an equation. The most important of them are the following:

- a reduction of the bias current increases the *dc* differential gain;
- in order to optimize the systematic offset the transistor aspect ratios should comply equation (5.14);
- the random offset depends on matching; however, for a given mismatch offset is proportional to $I_{bias}/g_{m,in}=(V_{GS1}-V_{Th})$; thus, in saturation the random offset is reduced by reducing the overdrive voltage of the input pair;
- the unity gain frequency ω_T is equal to $g_{m,in}/C_c$; additionally, in order to ensure a proper phase margin, ω_T must be smaller that $p_2 \approx g_{m2}/(C_1+C_2)$. Therefore, the compensation capacitance must be larger (by at least a factor 2) than $(C_1+C_2)g_{m,in}/g_{m2}$ [C_1 and C_2 are the capacitive load of the first and the second stage];
- since ω_T must be smaller than the non-dominant pole, p_2, the achievable gain bandwidth depends on the angular frequency of p_2. The designer achieves high speed positioning the non-dominant poles at pretty high frequencies;
- the zero nulling resistor should be equal to $1/g_{m2}$;
- symmetrical slew rates require $I_{B1}/C_c \leq I_{B2}/(C_c+C_L)$, [$I_{B1}$ and I_{B2} are the bias currents in the first and the second stage];
- with a given bandwidth we maximize the slew-rate with an high input pair overdrive [remember the relationship $SR=(V_{GS,1}-V_{Th})\omega_T$];
- optimum noise performances (topic that we will study shortly) require an

input pair transconductance higher than the one of the input stage active load. Moreover, the length of the active load should be higher than the one of the input pair.

5.8 SINGLE STAGE SCHEMES

The previous sections discussed in a detail the features of a two stages op-amp (or better an *OTA*). The gain that we can achieve with a simple stage of amplification is around *40 dB*. Thus, in order to achieve *80 dB* or so it was necessary to use the cascade of two stages. However, we have seen that two stages bring about two poles one close to the other and this requires compensation. In turn, a compensation network, beside increasing the global complexity, reduces the design flexibility.

A cascode with cascode load permits us to achieve a high gain (around *80 dB*) without the disadvantage of having two poles one close to the other. Therefore, the use of cascode based *OTA* is an interesting solution alternative to the two stages *OTA*. This section discusses the features of a "single stage" *OTA* so that the outcomes of the study will permit us to properly decide which solution is convenient for given design constraints.

5.8.1 Telescopic Cascode

The simplest version of a single stage *OTA* is the telescopic architecture, shown in Fig. 5.27: the input differential pair injects the signal currents into common gate stages. Then, the circuit achieves the differential to single ended conversion with a cascode current mirror. Note that the transistors are placed one on the top of the other to create a sort of telescopic composition (from this the name of the circuit). The small signal resistance at the output node is quite high: it is the parallel connection of two cascode configurations. Such a high resistance benefits the small signal gain without limiting the circuit functionality when we require an *OTA* function.

By inspection of the circuit one finds the low-frequency small signal differential gain

$$A_0 \cong g_{m1} \frac{r_{ds8} g_{m6} r_{ds6} \cdot r_{ds2} g_{m4} r_{ds4}}{r_{ds8} g_{m6} r_{ds6} + r_{ds2} g_{m4} r_{ds4}} \tag{5.60}$$

it is proportional to the square of the product of a transistor transconductance and an output resistance. Therefore, as expected, the telescopic cascode achieves a gain similar to the one of the two stages architecture. Moreover, by

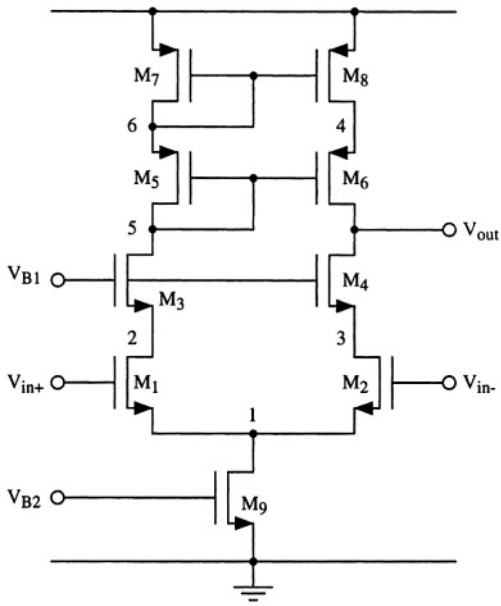


Fig. 5.27 - Telescopic cascode single stage OTA.

inspection of the circuit, one observes that all the nodes, excluded the output, shows a pretty low small signal resistance. Node *1* is an equivalent ground for differential signals; node *2, 3* and *4* are sources of a transistor, node *5* is coupled with node *6* by a diode connected transistor. Assuming the capacitance affecting the given nodes all due to parasitic contributions, the resulting time constants are all much smaller than the one associated to the output node. Therefore, the output node can easily become the dominant pole of the circuit. The other nodes should be at least A_0 apart being the non dominant ones.

Since the circuit shows one high impedance node only, it is not possible (and not necessary) to exploit the Miller effect to procure pole splitting. A possible capacitance loading the output node permits to make dominant the related pole and ensure stability.

Fig. 5.28 shows the simplified small signal equivalent circuit for a single pole approximation. The current generated by the transconductance generator $g_{m1}v_{in}$ is injected into the output resistance with in parallel the capacitance at the output node (C_L). The time constant $C_L r_{out}$ gives rise to a pole that causes a roll-off of the Bode plot with a slope of *20 dB/decade.* Therefore, the *0 dB* axis is crossed at the angular frequency $\omega_T = g_{m1}/C_L$.

An accurate study of the frequency response is not done here: it should involve the analysis of a pretty complex equivalent circuit. Since we have many nodes, we expect that the transfer function has a high order polynomial

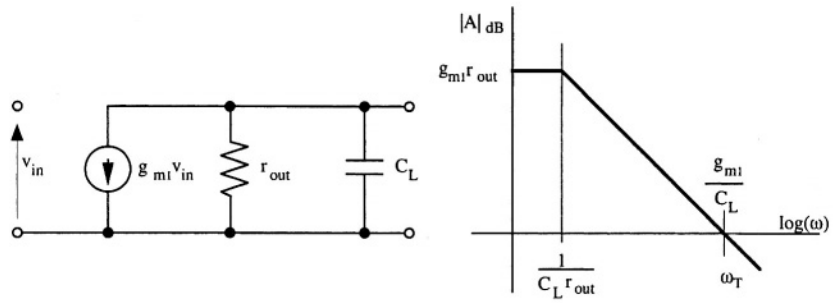


Fig. 5.28 - Simplified equivalent circuit of a single stage OTA and its Bode diagram.

denominator which requires approximations to be solved. Typically, the results achieved after approximations are not much different from the ones resulting from the intuitive estimation given above.

The telescopic configuration uses only one bias current. It flows through the differential input stage, the common base stage and the differential to single ended converter. Therefore, for a given bias voltage, the power is used at the best. By contrast, we have disadvantages: they concern the limited output dynamic range and the request to have an input common mode voltage pretty close to ground (or V_{SS}).

The triode limit of M_6 establishes the maximum allowed output voltage. By inspection of the circuit it is given by

$$V_{out,max} = V_{DD} - V_{GS7} - V_{GS5} + V_{GS6} - V_{sat,6} \cong V_{DD} - V_{Th,p} - 2V_{sat,p} \quad (5.61)$$

For typical situations it is *1 V* or more below the positive supply voltage.

The lower boundary of the output voltage depends the triode limit of M_4 that, in turn depends on V_{B1}

$$V_{out,min} = V_{B1} - V_{GS4} + V_{sat,4} = V_{B1} - V_{Th,n} \quad (5.62)$$

Normally the designer broadens the output swing by keeping low V_{B1}. However, the value of V_{B1} affects the minimum level of the input common mode voltage

$$V_{in,CM} \leq V_{B1} - V_{GS4} - V_{sat,2} + V_{GS2} \cong V_{B1} - 2V_{sat,n} \quad (5.63)$$

In turn, the input common mode voltage should allow M_9 to be in the saturation region.

$$V_{in,CM} \geq V_{sat,9} + V_{GS2} \quad (5.64)$$

Therefore, we can achieve an optimum negative swing ($3V_{sat,n}$ above ground) keeping the input common mode voltage as low as $V_{sat,9} + V_{GS2}$ (approximately equal to $V_{Th,n} + 2V_{sat}$). Assuming a symmetrical output swing around $V_{out,max}$ and $V_{out,min}$, the output common mode voltage becomes

$$V_{out,CM} \cong V_{DD} + V_{B1} - V_{Th,n} - V_{Th,p} - 2V_{sat,p} \tag{5.65}$$

KEEP IN MIND!

The use of five transistors one on the top of the other burdens the output dynamic range. A proper choice of the bias voltage and a minimum input common mode (for n-channel input pair) enlarge the output swing.The cost is that the common mode output differs from the input.

that for a typical design is quite a bit higher than V_{B1}. Therefore, the output common mode voltage is different (higher) than the input common mode voltage. This, in some applications is a limit: for instance, it is not possible to connect the telescopic cascode in the unity gain buffer configuration.

An interesting feature of the telescopic cascode is that it needs only two wires for the biasing: one for V_{B1} and the other for V_{B2}. This is more than the two stages amplifier that needs only one wire, but is less than other single stage architecture that will be studied shortly. Having a reduced wiring is important since it limits the chip area and, more important, prevents possible spur injection.

Example 5.4

Design a telescopic cascode single-stage OTA (Fig. 5.27) having about 70 dB dc gain and GBW 60 MHz. The required phase margin is 55° with a capacitive load of 3pF. The circuit must operate with a 1.8 V single supply voltage. Since it is necessary to procure the greatest the output swing, suitably modify the scheme in Fig. 5.27. Use the transistor models in Appendix C.

Solution:

It is required to procure the largest output swing with only 1.8 V supply voltage. The scheme in Fig. 5.27 has from the output to ground three drain-to-source jumps. Therefore, a proper choice of V_{B1} and an appropriate input common-mode voltage would permit a minimum output swing just at three saturations above ground. Assuming that the saturation equals 0.1 V, it results

$$V_{out,min} = 3V_{sat,n} \approx 0.3V.$$

By contrast it is not possible to have the output swing close to the

positive rail. Equation (5.61) bounds the upward swing to

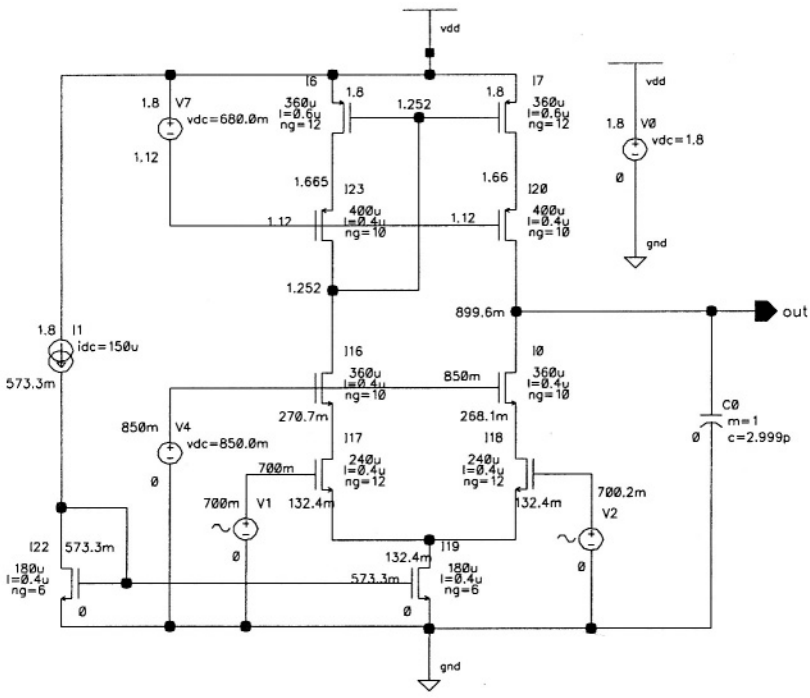
$$V_{out,max} = V_{DD} - V_{Th,p} - 2V_{sat,p} \approx (1.8 - 0.7)V = 1.1V$$

Where $V_{Th,p} \approx 0.5\ V$ *and, again, the saturation voltage is expected to be 0.1 V*

Since the major limit to the upward swing comes from the cascode current mirror, it is advisable to replace it with an equivalent counterpart that can reach a wider swing. The high compliance current mirror studied in Section 4.1.7 is a good candidate: a proper choice of the bias voltage used permits the output swing to approach the positive supply voltage by two drain-to-source drop voltages: 1.6 V, (the assumed saturation is 0.1 V). The diagram below shows the modified version of the telescopic cascode. The figure details the transistor sizes and the dc voltages.

The first step of the design consists in the estimation of the bias current. Assuming the GBW established by the dominant node, we have

$$2\pi f_T = \frac{2I_D}{(V_{GS} - V_{Th})C_L},$$



and assuming for the input pair $V_{GS} - V_{Th} = 0.1$ *V, we calculate*

$$I_D = \pi f_T (V_{GS} - V_{Th}) C_L = 56.5\mu A .$$

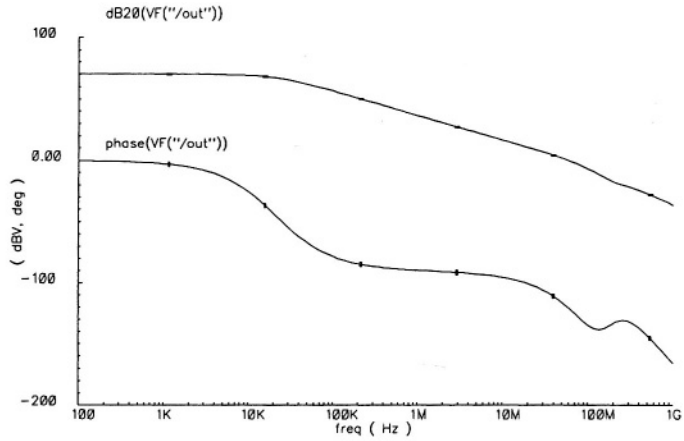
The current used is a bit higher (150 μA *tail current); it accounts for some margin and for different* V_{DS} *in the reference generator and the tail current source. The circuit employs large aspect ratios for all the transistors. This because the overdrive voltage must be pretty low (the expect value is 0.1 V).*

The drain-to-source voltages can not be pushed at the limit of saturation (0.1 V). The output resistance of transistors drops at low value affecting, in turn, the overall gain. Simulations show that drain-to-source voltages around 0.13 V permit to achieve the target of 70 dB. Moreover, in order to increase the output resistance of the p-channel cascades, we use the length of the transistors connected to V_{DD} *equal to 0.6* μm.

The voltage at the input of high compliance cascode is 1.25 V while the (quiescent) output is 0.9 V. This difference justify the small systematic offset (0.2 mV) used at the input.

The schematic shows that all the transistors are split into a given number of parts (ng, number of gates). The numbers used are suitable for a proper layout: the input pair and the p-channel elements connected to V_{DD} *are divided into 12 elements to be possibly composed in common centroid or/and an inter digitized arrangement. The cascoding transistors can not form inter digitized pairs: they are divided into 10 elements to allow some gap in the layout.*

The below figure shows the Bode diagrams of the output voltage. The achieved gain is 71 dB. The unity gain frequency is 65 MHz and the phase margin is 58°.



5.8.2 Mirrored Cascode

The mirrored cascode circuit, shown in Fig. 5.29, allows adaptability in the input common mode range. The signal currents generated by the input pair is mirrored by M_{11}-M_{10} and M_{12}-M_{13} and delivered to the source of M_3 and M_4 respectively. The operation of the rest of the circuit is identical to the telescopic version shown in Fig. 5.27.

To explain why the circuit offers a flexible input common mode range note that above the input pair drains we have two diode connected elements. Therefore, the common emitter of the input pair, node *1,* can approach the drains of M_{11} and M_{12} by just the saturation of the input pairs. Therefore, we have

$$V_{in,CM} \leq V_{DD} - V_{GS11} - V_{sat,1} + V_{GS1} \tag{5.66}$$

If the threshold of the *n-channel* transistors and the one of the *p-channel* matches V_{GS1} and V_{GS11} are comparable: the maximum value of $V_{in,CM}$ becomes close to V_{DD}. The minimum value of $V_{in,CM}$ is the one that brings M_9 in the triode region

$$V_{in,CM} \geq V_{GS,1} + V_{sat,9} \tag{5.67}$$

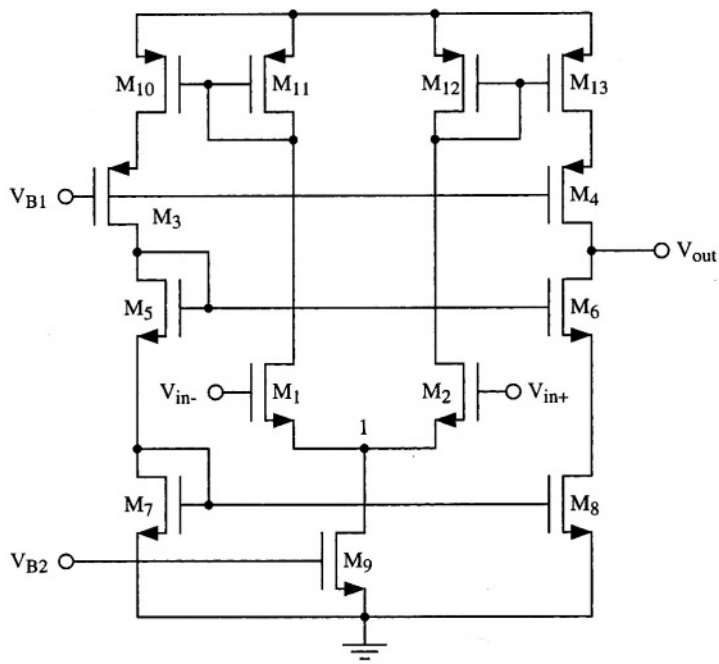


Fig. 5.29 - Mirrored cascode single stage OTA.

Therefore, the permitted range of the input common mode voltage, $V_{in,CM}$, becomes pretty large.

The output voltage range is limited by the conditions that bring M_4 and M_6 in triode. In turn, the saturation limit for M_4 depends on the V_{B1} used. By inspection of the circuit, we have

$$V_{B1,max} = V_{DD} - V_{sat} - V_{GS4} \tag{5.68}$$

this upper limit for V_{B1} leads the drain of M_{13} just one V_{sat} below V_{DD}. Therefore

$$V_{out,max} = V_{DD} - 2V_{sat} \tag{5.69}$$

that is only one saturation less than what we can achieve with a simple inverter with active load. By contrast, the cascode arrangement M_6-M_8 limits the negative swing. The voltage at the drain of M_8 is a replica of the one of M_7. Thus, we have

$$V_{out,min} = V_{GS7} + V_{sat} \tag{5.70}$$

Therefore, the output swing range is not symmetrical with respect to the positive and negative rail.

OBSERVATION

The use of current mirrors to lead the current to the output node improves the output swing, However, the additional non-dominant nodes affect the speed. As a general rule, the simpler is the circuit the faster is the response,

The mirror factors of M_6-M_7, M_{10}-M_{11} and M_{12}-M_{13} are additional design degrees of freedom. Of course, it is required to match the currents of M_{13} with the one of M_7; this determine a constraint between the mirror factors. However, the designer can use the remaining degrees of freedom to establish the currents in the branches with M_6 and M_7.

The current mirrors used to alleviate the constraints on the input common voltage limit the speed of the circuit. Compared to the telescopic cascode the mirrored version has two additional non dominant poles. They, possibly, reduce the phase margin. Therefore, in order to ensure stability, the designer has to likely increase the capacitance that loads the output, thus the speed (bandwidth and slew-rate) diminishes.

Observe that transistors M_3 and M_5 are used to improve symmetry in the circuit. They are not strictly necessary; by removing them the circuit is less symmetrical, but two non-dominant poles are removed as well.

Example 5.5

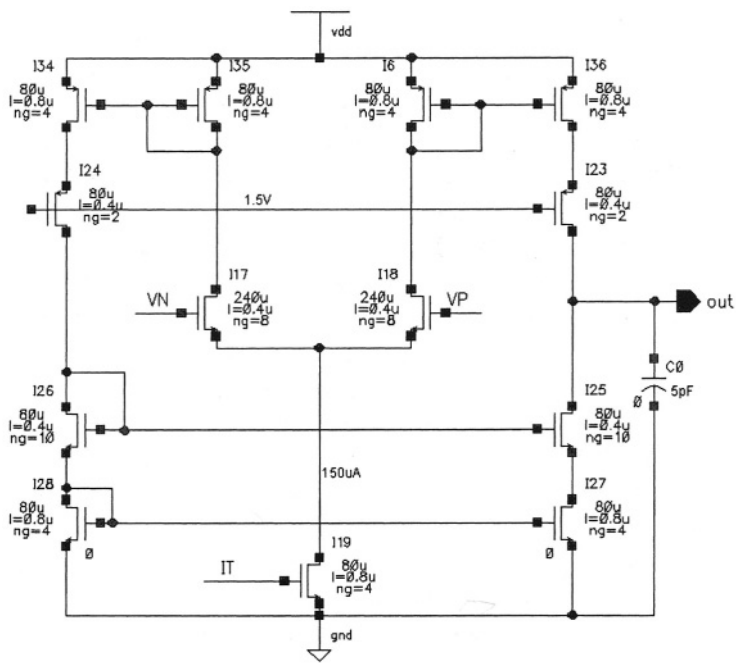
Design the single-stage mirrored cascode OTA shown in Fig. 5.29. The required dc gain and bandwidth are 80 dB and 50 MHz respectively. The phase margin must be better than 55°. Determine the load capacitance that complies with specifications. Study the behaviour of the small signal dc gain as a function of the output voltage. Use 3.3 V as single supply voltage and keep the power consumption below 1 mW. Use the transistor models of Appendix C.

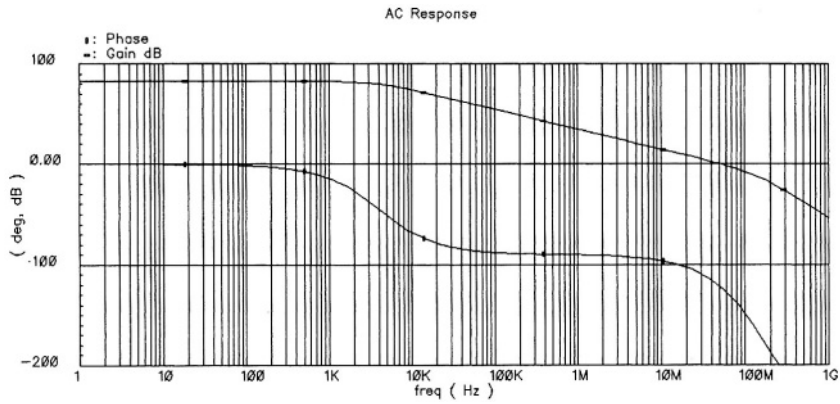
Solution:

The available bias current is $1mW/3.3V = 300\ \mu A$. *Assuming a 1:1 mirror factor for the pairs* M_{10}-M_{11} *and* M_{12}-M_{13} *the bias current is evenly distributed between the four transistors. Therefore, the current in the output branches is* $75\ \mu A$ *and the tail current generator of the differential pair,* M_9, *drains* $150\ \mu A$.

Example 5.4 used an input pair with W/L $240\mu/0.4\mu$. *With a tail current equal to* $150\ \mu A$ *the transconductance is 1.81 mA/V. This value is reasonably high: for achieving the 80 dB of gain the output resistance must be just* $5\ M\Omega$.

The output resistance of a cascode configuration is given by an r_{ds} *multiplied by the gain of a transistor. A typical value of* r_{ds} *ranges*

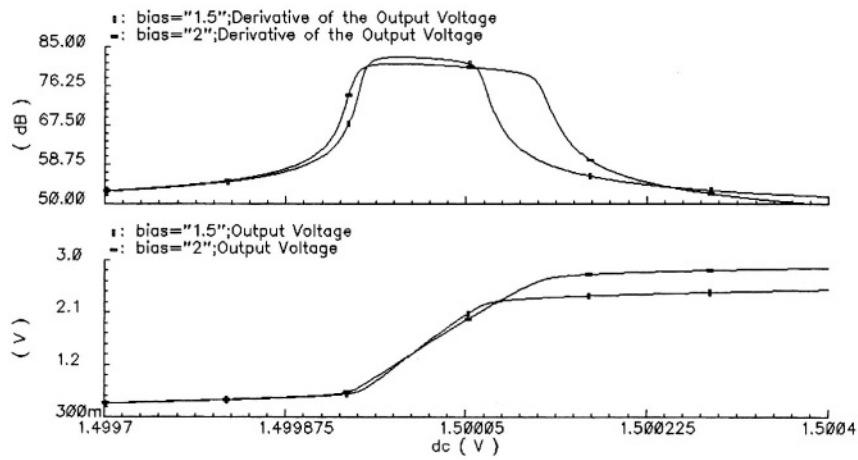




from 100 KΩ to 300 KΩ while the gain of a transistor is around 50-100. Therefore the output resistance of each cascode branch is easily larger than 5 MΩ. The transistor sizing of the schematic used leads to an equivalent resistance of 8.5 MΩ for the upper cascode and 27 MΩ for the lower cascode. The two values are unbalanced; a proper resizing of transistors would lead to more even values. Nevertheless, the used sizing make the geometry of the n-channel and the p-channel output cascodes equal. This feature can be attractive because it facilitates the layout.

The small signal analysis of the circuit leads to the results shown in the figure of the previous page. The dc gain is 83 dB, the phase margin is 59° and the gain-bandwidth is 54 MHz.

A sweep of the input voltage carried out by a .dc analysis permits us to determine the input output transfer characteristic. A simple



post-processing obtains the dc gain (its derivative). The results (see the bottom figure of the previous page) show that two different bias voltages used in the upper cascode: 1.5 V and 2 V lead to quite different performances. In the case of 1.5 V bias the gain is a bit larger but the output dynamic range is lower by approximately 0.5 V than the condition with 2 V bias.

5.8.3 Folded Cascode

The basic difference between the mirrored cascode and the telescopic one resides in the disconnection between the biasing of the input and the output branch. This results from the current mirrors that provide replica of the signal currents produced by the input stage and furnishes them to the output branch. The folded cascode, shown in Fig. 5.30 achieves the same conceptual function by a direct transfer of the small signal currents from the input to the output branch. The low impedance shown by the sources of M_3 and M_4 ensure the signal currents transfer. Moreover, the biasing of the input and output stages don't

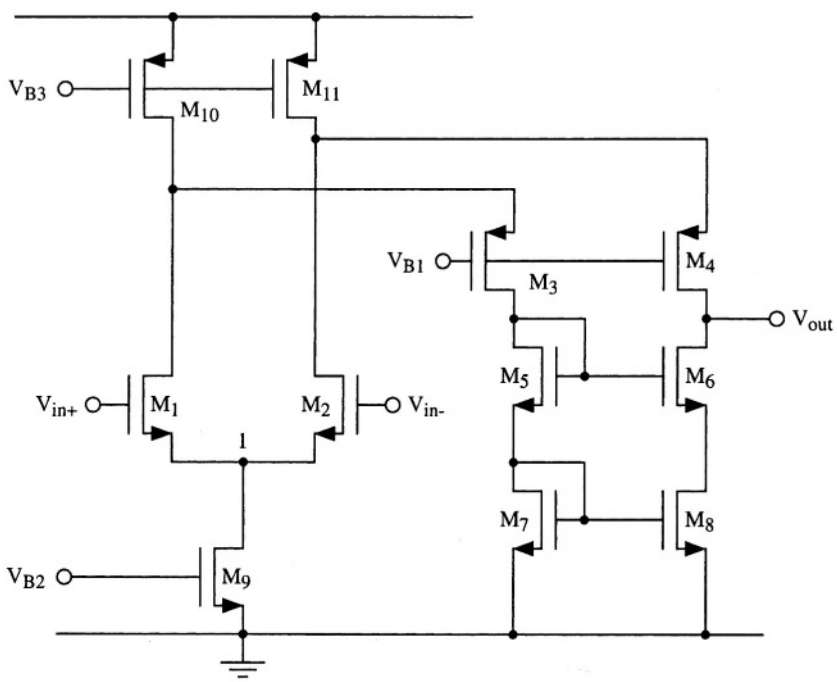


Fig. 5.30 - Folded cascode transconductance amplifier.

affect one each other. Therefore, the circuit achieves the same benefits for the common mode input and output range as for the mirrored cascode.

The folded version requires two additional current sources (M_{10} and M_{11}) to provide the current necessary for the input and the output branch. Therefore, the folded cascode demands for an additional wire connection to properly bias the two current sources.

The current source M_{11} affects the resistance of the upper cascode of the output branch. In fact, the resistance seen from the source of M_4 to ground (and node *1,* ground for differential signals) is the parallel connection of the drain resistance of M_2 and M_{11}

$$r_{eq} \cong \frac{g_{ds2} g_{ds11}}{g_{ds2} + g_{ds11}} \tag{5.71}$$

therefore, to calculate the small signal *dc* gain we should replace the pertinent term with above expression in equation (5.60).

Current sources M_9 and M_{10} (or M_{11}) set independently the currents of the input pair and the output branch. The degree of freedom available permits the designer to establish the value of the two currents freely. Assuming that the input stage is balanced we have

$$I_{M1} = \frac{I_{M9}}{2}; \quad I_{M7} = I_{M10} - \frac{I_{M9}}{2} \tag{5.72}$$

Thus, it is possible to maximise the gain by properly setting the input transconductance and the output resistance.

NOTICE

The mirrored cascode and the folded cascode need supplementary biases that, in turn, demand for extra bias generators and new bus lines.

Recall that a capacitive coupling with noisy lines can corrupt the bias voltages.

The constraints on the output dynamic range of the folded cascode are similar to the ones studied for the mirrored version. The bias voltage V_{B1} limits the upward swing. The cascode structure M_6-M_8 bounds the downward swing, and, as already discussed, the value of V_{B1} should be low enough to consent the saturation of M_{11}. Therefore, if we use for V_{B1} the maximum value expressed by equation (5.68) we obtain the same output swing constraints given by (5.69) and (5.70). However, it is possible to improve the downward swing. We should replace the plain cascode current mirror with an high swing version, like the modified cascode (see § 4.1.6) or the high compliance cascode (see § 4.1.7).

Example 5.6

Assume to come across the data-base of a folded cascode OTA. The circuit has been designed using a 0.8 μ CMOS technology which Spice models are given in Appendix A. The folded cascode operates with a 5 V supply voltage, drives 10 pF, and consumes 3 mW. The following input list describes the circuit.

```
Folded cascode
M1   2 10  1 0 MODN W=400U  L=1.5U
M2   3 11  1 0 MODN W=400U  L=1.5U

M3   5 9 2 4   MODP W=200U  L=1U
M4   8 9 3 4   MODP W=200U  L=1U
M5   5 5 6 0   MODN W=200U  L=1U
M6   8 5 7 0   MODN W=200U  L=1U
M7   6 6 0 0   MODN W=100U  L=2U
M8   7 6 0 0   MODN W=100U  L=2U
M9   1 12  0 0 MODN W=100U  L=2U
M10  2 13  4 4 MODP W=100U  L=2U
M11  3 13  4 4 MODP W=100U  L=2U

MBN 12  12  0 0  MODN W=100U  L=2U
MBP 13  13  4 4  MODP W=100U  L=2U

IBN 0 12 0.5M
IBP 13 0 0.3M
VDD 4 0 5
V10 10 0 2.5
VB1 9 0 3

CL 8 0 10P

RF1  8 99  1G
RF2  99 11  1G
C1 99 0 1
C2 98 11 1
VIN 98 0 0 AC 1
```

Identify the key points of the design and try to understand the strategy followed by the designer.

Solution:

The technology is a 0.8μ CMOS. The lentos used for the devices M_3, M_4, M_5, and M_6 are close to the minimum allowed. The possible explanation of the choice is that the transistors must provide gain to enhance the cascode output resistance. The expected value of g_m/g_{ds} must be in the range of 50 to 100. A relatively low length sustains the transconductance and likely permits us to achieve the expected gain.

The length of the input pair transistors is 1.5 μ. This figure is almost twice the minimum allowed. This choice probably answer to the need of a low output conductance. The resistance of the upper output cascode depends on $g_{sd2}+g_{ds11}$. The length of the input transistor should lead to comparable values for output conductances of M_2 and M_{11}.

The length of all the transistors connected to the supply lines is 2 μ: for all these transistors it is essential to have pretty low g_{ds}.
The Spice simulation of the circuit furnishes the results given below. Observe that the saturation voltage of most transistors is lower than 150 mV. The saturation voltage of M_9 is not relevant. What surprises us is the large saturation voltage of M_3 and M_4. Moreover, the used valuer of V_{B1} is much lower than the maximum possible. Presumably, the design aims at a symmetrical output swing around 2.5 V. Since the downward swing is bounded by the lower cascode it doesn't make sense to expand the upward dynamic range.

	m1	m2	m9	m5	m6	m7	m8
model	modn	modn	modn	modn	modn	modn	modn
id	2.49E-04	2.49E-04	4.99E-04	4.81E-05	4.81E-05	4.81E-05	4.82E-05
vgs	1.259	1.259	1.341	1.095	1.093	0.973	0.973
vds	2.632	2.636	1.241	1.095	1.525	0.973	0.975
vbs	-1.241	-1.241	0.000	-0.973	-0.975	0.000	0.000
vth	1.152	1.152	0.866	1.066	1.064	0.862	0.862
vdsat	0.119	0.119	0.362	0.060	0.060	0.117	0.117
gm	3.38E-03	3.38E-03	1.95E-03	1.28E-03	1.29E-03	6.04E-04	6.04E-04
gds	1.44E-05	1.44E-05	1.23E-05	1.01E-05	9.10E-06	2.64E-06	2.64E-06
gmb	6.59E-04	6.59E-04	5.91E-04	2.52E-04	2.53E-04	1.98E-04	1.98E-04
cbd	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
cbs	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
cgsovl	1.48E-13	1.48E-13	3.70E-14	7.40E-14	7.40E-14	3.70E-14	3.70E-14
cgdovl	1.48E-13	1.48E-13	3.70E-14	7.40E-14	7.40E-14	3.70E-14	3.70E-14
cgbovi	2.33E-16	2.33E-16	3.10E-16	1.55E-16	1.55E-16	3.10E-16	3.10E-16
cgs	8.63E-13	8.63E-13	2.88E-13	2.88E-13	2.88E-13	2.88E-13	2.88E-13
cgd	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
cgb	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00

m3	m4	m10	m11	mbn	mbp	
model	modp	modp	modp	modp	modn	modp
id	-4.81E-05	-4.81E-05	-2.98E-04	-2.98E-04	5.00E-04	-3.00E-04
vgs	-0.873	-0.876	-1.252	-1.252	1.341	-1.252
vds	-1.805	-1.376	-1.127	-1.124	1.341	-1.252
vbs	1.127	1.124	0.000	0.000	0.000	0.000
vth	-0.809	-0.811	-0.711	-0.711	0.866	-0.711
vdsat	-0.090	-0.092	-0.469	-0.469	0.363	-0.469
gm	9.53E-04	9.32E-04	1.03E-03	1.03E-03	1.96E-03	1.04E-03
gds	8.68E-06	9.44E-06	2.07E-05	2.07E-05	1.18E-05	1.94E-05
gmb	8.26E-05	8.09E-05	1.57E-04	1.57E-04	5.92E-04	1.58E-04
cbd	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
cbs	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
cgsovl	7.40E-14	7.40E-14	3.70E-14	3.70E-14	3.70E-14	3.70E-14
cgdovl	7.40E-14	7.40E-14	3.70E-14	3.70E-14	3.70E-14	3.70E-14
cgbovl	1.05E-16	1.05E-16	2.15E-16	2.15E-16	3.10E-16	2.15E-16
cgs	2.74E-13	2.74E-13	2.81E-13	2.81E-13	2.88E-13	2.81E-13
cgd	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
cgb	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00

The width of all transistors is multiple of 100μ. This choice permits the designer to split all the transistors into fingers 25μ wide and to achieve a modular and compact layout.
The ,ac simulation leads to 80 dB dc gain, unity gain frequency equal to 42 MHz and phase margin equal to 63°. The procured results are quite conventional for a general purpose OTA found in a $0.8\ \mu$ CMOS cell library.

5.8.4 Single Stages with Enhanced dc Gain

We studied in Chapter 4 that it is possible to increase the output resistance of a cascode by using a multiple cascode architecture or a local feedback. The use of the same approaches permits us to enhance the gain of a single-stage *OTA*: the gain comes from the product of the input pair transconductance and the output resistance. Therefore, for a given transconductance we increase the gain if we enhance the output resistance. Nevertheless, we have to remember that the supplementary circuits used bring about additional nodes that, in turn, lead to new poles that worsen the phase margin.

Fig. 5.31 shows a mirrored *OTA* configuration with a double-cascode load. The new circuit has four additional transistors (M_{3a}, M_{4a}, M_{5a} and M_{6a}) for achieving four common-gate stages. The output resistance increases by approximately a $g_m r_{ds}$ factor and the *dc* gain increases by the same factor.

The circuit has four additional nodes (the sources of the four added transistors). Therefore, the compensation of the stage can result critical. The small signal resistance of the added nodes is in the order of $1/g_m$; the associated capacitance comes from parasitics; therefore, the corresponding poles will be at

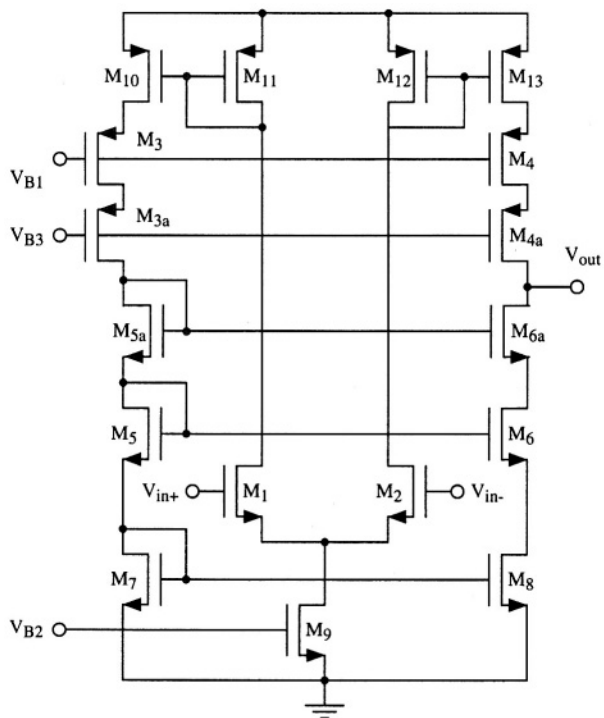


Fig. 5.31 - Mirrored double-cascode single stage OTA

a relatively high frequency. Nevertheless, the additional poles are one close the other and the phase margin is worsened because of the combined action of the cluster of non-dominant poles. Consequently, we have to bring the dominant pole at a lower frequency to ensure a more solid non-dominant position of the additional non-dominant poles.

The circuit necessitates an additional bias voltage, V_{B3}. The value of value V_{B3} determines the upward output swing constraint. Moreover, V_{B3} must be designed in combination with V_{B1}. Assuming that the gate-source voltages of M_3 and M_{3a} and the one of M_4 and M_{4a} matches, the drain-to-source voltage of M_3 and M_4 are equal to $V_{B1} - V_{B3}$. Therefore, if transistors should be kept at the limit of the triode condition, the two bias voltages must differ by just a saturation voltage and, more important, their difference should not change with temperature and process variation.

Fig. 5.32 shows the use of local feedback to increase the *dc* gain in a cascode architecture. The circuit includes two gain stages. One is used to enhance the output resistance of the upper cascode, the other to boost the output resist-

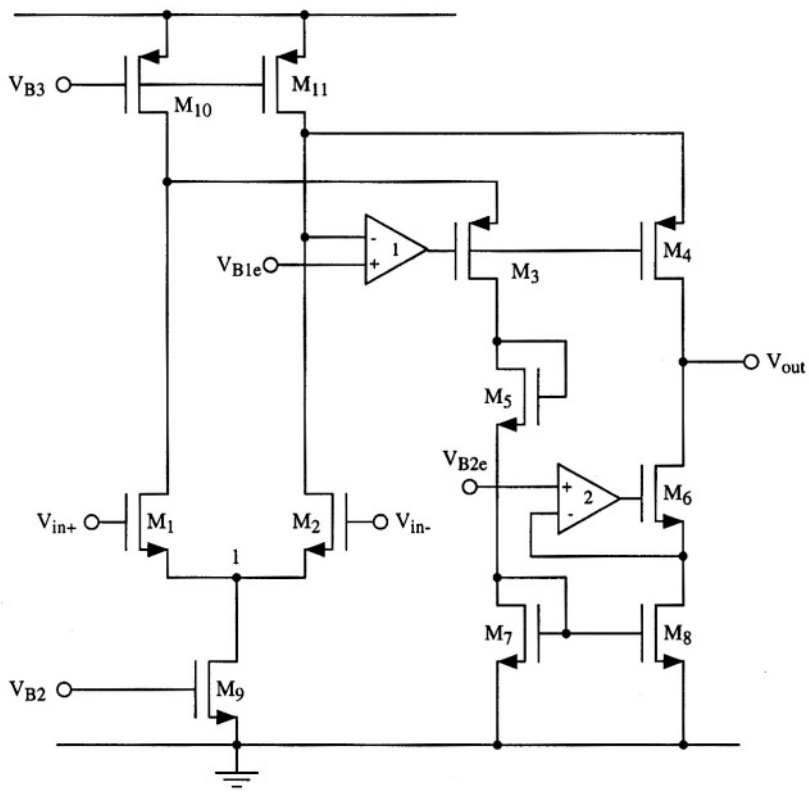


Fig. 5.32 - Use of local feedback to enhance the *dc* gain of a folded cascode.

ance of the lower one. Both gain stages control the source of the common gate elements (M_4 and M_6) to be constant. The bias voltages V_{B1e} and V_{B2e} correspond to the desired source voltages. In order to maximize the output swing these bias voltages should just be one saturation apart from the positive and negative rails.

Transistor M_5 in the normal cascode scheme biases the gate of M_6; in Fig. 5.32 it provide a drop voltage to reduce the V_{DS} of M_3 and a possibly attenuate a systematic difference between the currents in M_5 and M_4.

The frequency behaviour of the circuit depends on the used gain stages and on their frequency performance. We observe that the feedback factor in the local loops is equal to one. Therefore, the bandwidth of the feedback loop equals the unity gain frequency of the gain stage used. In turn, to ensure stability, the unity gain frequencies of the gain stage must be higher than the unity gain frequency of the cascode *OTA*.

The common-mode input of the upper gain stage is close to V_{DD} while the one of the lower gain stage is close to ground. The design of a gain stage which common mode input is close to the supply rails is not an easy task. Typically, the upper gain stage uses an *n-channel* input pair and the lower gain stage uses a *p-channel* input pair.

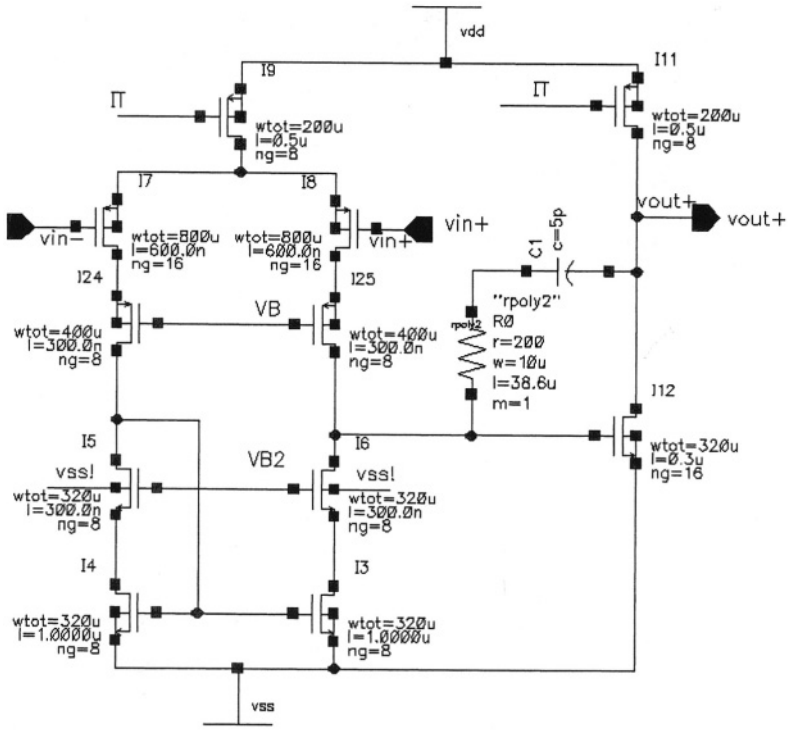
Example 5.7

Design an OTA able to drive a capacitive load of 2 *pF with 100 dB dc gain and* $f_T = 70$ *MHz (phase margin* $>$ *50°). The supply voltage is 1.8 V. The power consumption must be less than 0.8 mW. In order to achieve high gain and low power consumption use a two stages configuration with a telescopic cascode as the first stage and an inverter with active load as second stage. Use the Spice models of Appendix C.*

Solution:

The low value of supply voltage doesn't provide room for a conventional telescopic cascode. It is necessary to use a high compliance current mirror that allows room for the input pair and its current tail generator. Moreover, all the overdrive voltages must be kept at a minimum level to procure an acceptable dynamic swing. This will lead to transistors with a pretty large aspect ratio.

The above reflections lead to the schematic shown in the figure. The tail current in the first and the bias of the second stage is 150 μA. *The circuit uses a p-channel input pair which large W/L (800μ/0.6μ) achieves a transconductance as large as 2.42 mA/V. The value of input transconductance leads, with a compensation capacitance equal to the capacitive load equal to 2 pF, to a gain band-*



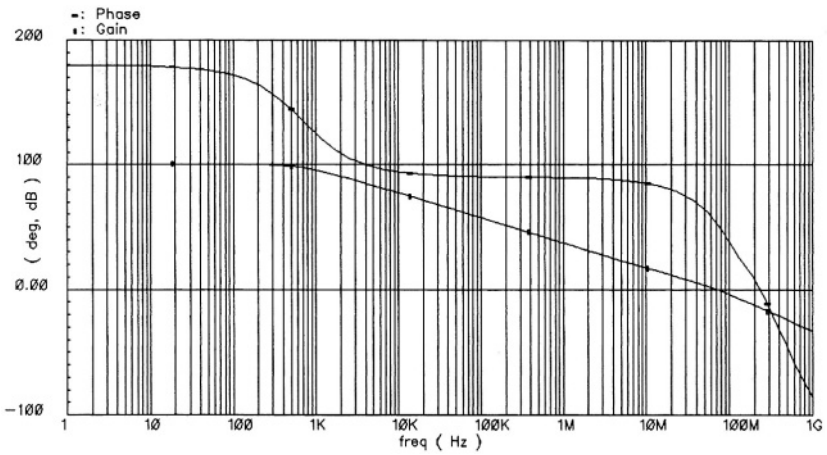
width equal to

$$f_T = \frac{1}{2\pi}\frac{g_m}{C_C} = \frac{1}{6.28}\frac{2.42 \cdot 10^{-3}}{2 \cdot 10^{-12}} = 192 \text{ MHz}$$

that is about three times the specification. This seems a good margin. However, as shown by simulations, the non-dominant poles require a pretty large compensation capacitance. To have an acceptable phase margin it is required to use $C_C = 5$ *pF. The expected* f_T *drops by a factor 2.5 and becomes 76.8 MHz.*

The large input transconductance permits to achieve a good first stage gain. The high compliance cascode current mirror keeps the transistors close to the saturation region. This prevents to obtain a very large output resistance. However, the used design leads to a first stage output resistance a bit smaller than $1\ M\Omega$ *providing a first stage gain equal to 68 dB.*

The second stage should provide at least a dc gain of 32 dB. The target is not particularly difficult. The sizing of input transistor equal to $(W/L) = 320\mu/0.3\mu$ *and an active load with* $(W/L) = 200\mu/0.5\mu$ *fulfil the equest.*



As already mentioned five low-impedance nodes in the input stage burden the phase margin. Computer simulations reveal that in order to procure a phase margin better than 50° it is necessary to use $C_C = 5\ pF$.
The above figure shows the Bode diagram of the output voltage. Other relevant results are summarized in the Table.

Parameter	Value	Unit
dc gain of the first stage	67	dB
dc gain of the second stage	34	dB
f_T	75.3	MHz
phase margin	52°	-
input referred offset	7.33	μV

5.9 CLASS AB AMPLIFIERS

Usually *AB* class amplifiers drive small resistive loads. Since in integrated *CMOS* applications only capacitors constitute the load, the objective of the *IC* designer is different than when addressing discrete component solutions. For an off-the-shelf op-amp it is necessary to ensure a low output resistance for any possible feedback and load. The aim here is to be able to drive large capaci-

tance's quickly and with a minimum harmonic distortion. For high speed or large capacitor, the target is difficult to achieve, especially when using circuits like the *A* class *OTA's* studies so far. The *AB* class amplifier better addresses the problem since it permits to deliver currents larger than the quiescent value.

5.9.1 Two Stages Scheme

The two-stages amplifier of Fig. 5.34 achieves the class *AB* operation in the second stage. The figure shows how this function results from a modification of the already studied two stages amplifier. The output of the first stage drives both the *n-channel* and the *p-channel* transistor of the second stage. Driving of transistor M_5 equals the one in the *p-channel* input counterpart of Fig. 5.12. Transistor M_6 is no more a current source but the control of its gate is a shifted down replica of the input-stage output. Transistor M_8 and the bias current M_9, possibly controlled by the same voltage used for M_7 achieve the level shift. Both transistors M_5 and M_6 contribute to the transconductance gain of the second stage. Therefore, being the gain of the second stage

$$A_2 = \frac{g_{m5} + g_{m6}}{g_{ds5} + g_{ds6}} \tag{5.73}$$

as a side benefit, the low frequency gain will result somewhat increased.

The network C_{c1}-R_z ensures the compensation of the two stages amplifier. As already discussed, the zero nulling resistor acts to remove the zero in the right s-plane. However, at high frequency, the *0°* phase shift provided the

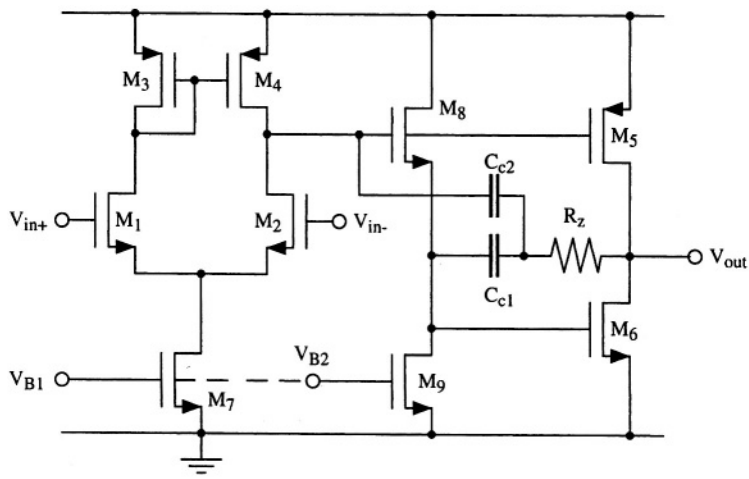


Fig. 5.33 - The two stages push-pull OTA.

buffer M_8-M_9 can deteriorate the phase response. The use of capacitor C_{c2} attenuates this possible problem.

A critical point of the design involves the control of the quiescent current, I_6 of the second stage. By inspection of the circuit we can find the following relationship between the gate-to-source voltages of three transistors

$$V_{DD} = V_{GS6} + V_{GS8} + V_{GS5} \tag{5.74}$$

two of the transistors are *n-channel* and one is *p-channel,* therefore

$$V_{DD} = V_{Th,p} + 2V_{Th,n} + \sqrt{\frac{2}{k'_n}\left(\frac{L}{W}\right)_8 I_8} + \left(\sqrt{\frac{2}{k'_p}\left(\frac{L}{W}\right)_5} + \sqrt{\frac{2}{k'_n}\left(\frac{L}{W}\right)_6}\right)\sqrt{I_6} \tag{5.75}$$

$$\sqrt{I_6} = \frac{V_{DD} - V_{Th,p} - 2V_{Th,n} - \sqrt{\frac{2}{k'_n}\left(\frac{L}{W}\right)_8 I_8}}{\sqrt{\frac{2}{k'_p}\left(\frac{L}{W}\right)_5} + \sqrt{\frac{2}{k'_n}\left(\frac{L}{W}\right)_6}} \tag{5.76}$$

A possible variation of the supply voltage (in a commercial power supply generator it can be ± *10%*) and the technological alterations of threshold voltages make pretty low the accuracy of the first three terms in the numerator of (5.76). Thus, when using an uncontrolled current mirror (as it is done if in Fig. 5.33 V_{B1}=V_{B2}) the current in M_6, may vary significantly. A possible way to fix the problem is to control the overdrive voltage of M_8 and compensate with it the variation of the first three terms in the numerator of (5.76).

Fig. 5.34 shows a possible bias network that we can use to drive the current source M_9. The part within the dashed lines matches half of the input stage. It replicates the voltage of the gate of M_8 and use it to drive M_{8b} (a replica of M_8) The current mirror M_{5b1}-M_{5b} copies the current of M_{6b} and compares it with the reference current. Any difference between the two currents affects V_{B2}. This voltage drives M_{9b} so that the connection closes the loop. Voltage V_{b2} is then used in the scheme in Fig. 5.33.

TAKE HEED!

The mail limitation of the two stages AB class OTA in Fig. 5.33 derives from the uncertainty of the quiescent output current. Its control with a suitable feedback is essential.

Observe that we have three gain stages around the feedback loop. The load of stages M_{6b}-M_{5b} and M_{9b}-M_{8b} is given by $1/g_{m,5b}$ and $1/g_{m,8b}$ respectively. Consequently, their gain is limited. In practice, the entire gain loop is ensured by M_{5b1} and the output resistance of the current reference generator, I_{ref}.

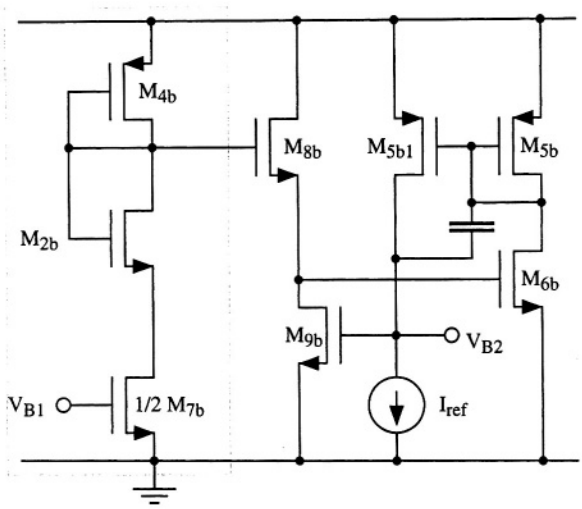


Fig. 5.34 - Possible bias network for driving transistor M_9 in Fig. 5.33

Because of the three stages around the loop, the circuit needs compensation. The capacitor across the input and output of M_{5b1} takes benefit of the Miller amplification. That is therefore the best position where to place the compensation capacitor.

5.9.2 Unfolded Differential Pair

Basically, the limit to slewing in a single stage amplifier occurs because of the current source used in the input differential pair. Even with a large differential signal the current that flows through one of the output nodes (see Fig. 5.35 a) can not be larger than the bias current, I_{bias}. In turn, we have a limit in the current that the circuit can deliver to the output nodes. It is possible to remove the above mentioned limitation by "unfolding" the input pair as shown in Fig. 5.35 b). The result is a stack of two complementary transistors that provide output currents from the two drains. The driving of the pair requires a *dc* difference between the control of the *n-channel* gate and the *p-channel* gate, $V_{AB,0}$. This difference controls the quiescent current, I_Q

$$V_{AB,0} = V_{th,n} + V_{th,p} + \sqrt{\frac{2I_Q}{C_{ox}}}\left(\sqrt{\frac{L_n}{\mu_n W_n}} + \sqrt{\frac{L_p}{\mu_p W_p}}\right) \qquad (5.77)$$

moreover, the circuit allows an extensive increase ΔI of the current if V_{AB} augments to demand slewing. Equation (5.77) permits us to calculate the current-

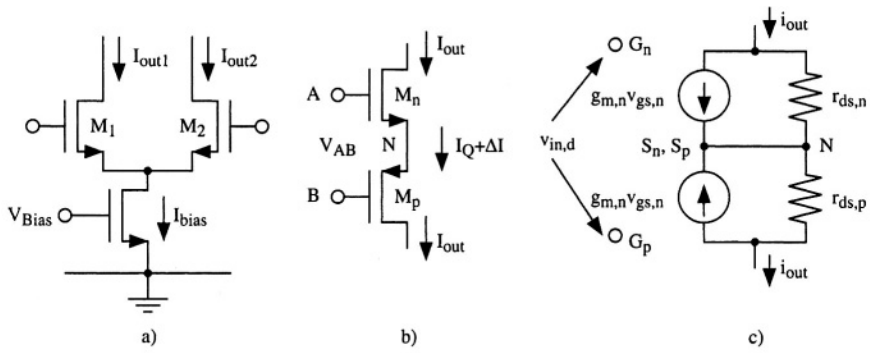


Fig. 5.35 - The "unfolded" differential pair: it is used in a class AB single stage amplifier

voltage relationship for any possible value of V_{AB}. When V_{AB} becomes lower than $V_{th,n} + V_{th,p}$ the current goes to zero. For voltages higher than such a condition we have

$$\frac{2I_{out}}{C_{ox}\mu_n\left(1+\frac{\mu_p W_p L_n}{\mu_n W_n L_p}\right)} = \left(\sqrt{\frac{2I_Q}{C_{ox}\mu_n\left(1+\frac{\mu_p W_p L_n}{\mu_n W_n L_p}\right)}} + \Delta V_{AB}\right)^2 \qquad (5.78)$$

that is a non-linear relationship between the input (signal) voltage and the output current.

Observe that any increase of V_{AB} is shared between the overdrive of the *n-channel* and the *p-channel* transistor. We can write

$$\Delta V_{AB} = \Delta V_{ov,n} + \Delta V_{ov,p} \quad ; \quad \frac{\Delta V_{ov,n}}{\Delta V_{ov,p}} = \sqrt{\frac{\mu_p W_p L_n}{\mu_n W_n L_p}} \qquad (5.79)$$

therefore, if we want to have equally split the variation of V_{AB} we have to use transistor aspect ratios inversely proportional to the mobility ratio. This will keep constant the voltage of the intermediate node *N* for a symmetrical change of node *A* and *B* voltages.

Fig. 5.35 *c)* shows the low frequency small-signal equivalent circuit of the unfolded pair. Neglecting the drain resistances permits to achieve an approximated (but effective) result. Without the drain resistances the small signal output current becomes equal to the one of two current generators

$$i_{out} = g_{m,n} v_{gs,n} = -g_{m,p} v_{gs,p} \qquad (5.80)$$

moreover, $v_{gs,n}$ and $v_{gs,p}$ are linked to the input differential voltage by

$$v_{in,d} = v_{gs,n} - v_{gs,p} \tag{5.81}$$

using (5.80) and (5.81) one obtains

$$i_{out} = \frac{g_{m,n} \cdot g_{m,p}}{g_{m,n} + g_{m,p}} v_{in,d} = g_{eq} v_{in,d} \tag{5.82}$$

that defines the equivalent transconductance gain g_{eq} of the unfolded pair. Observe that the equivalent transconductance equals half of the transconductance of each single transistor if $g_{m,n} = g_{m,p}$; moreover, the signal voltage is the entire differential input. Therefore, if $g_{m,n} = g_{m,p}$ the unfolded pair operates exactly as the normal (folded) counterpart.

COMMENT

For small signals, an unfolded differential pair operates like the folded counterpart. The equivalent transconductance is given by

$$g_{eq} = \frac{g_{m,n} \cdot g_{m,p}}{g_{m,n} + g_{m,p}}$$

A more precise analysis that includes the effect of drain resistances is not done here. The result will show some slight difference with respect to the simple result given above. The reader can do such an analysis as an exercise.

In addition to the equivalent transconductance estimation we should determine the equivalent output resistances of the unfolded pair. That is done by inspection of the circuit in Fig. 5.35 c). Nulling the input signal, we immediately estimate the total equivalent resistance: it is the series connection of $r_{ds,n}$ and $r_{ds,p}$. For input differential signal half of that series value depicts the resistance between each output node and analog ground.

5.9.3 Single Stage AB-class OTA

The circuit in Fig. 5.36 uses two unfolded pairs. Their connections are crossed so that the current of the pair M_1-M_2 is accessible on the right-bottom side and the one of the pair M_3-M_4 is at hand on the right-top side. These currents are mirrored by Wilson schemes and summed up at the output node. The same input voltage controls transistors M_1 and M_4. By contrast, two level shifts, ΔV, (represented in the figure with two batteries) disjoint the positive input from the gates of M_3 and M_2. Alternatively, we can use level shifts to disjoint the voltages of the gate of M_1-M_3 from the differential input applied throughout M_2 and M_4. However, the latter option is not optimum: parasitic capacitances will surely affect the level shifts; it is better to use them on the positive input since it is typically connected to the analog ground.

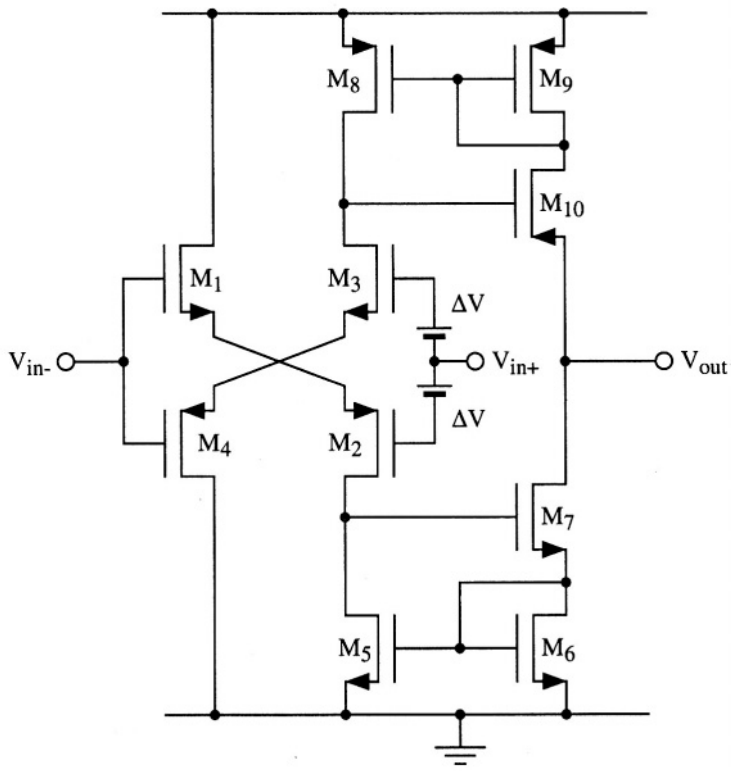


Fig. 5.36 - Single stage AB class op-amp

If, for instance, the input umbalancement brings the positive input higher than the negative one, the current in the pair M_1-M_2 diminishes (and, possibly goes to zero) while the current in the pair M_3-M_4 goes up. Therefore, thanks to the mirroring, the circuit accomplishes a push pull operation.

In the quiescent conditions the level shifts ΔV controls the current in the output stage. In fact, if the current from the output node is zero, the following relationships hold

$$I_7 = k_{56} I_1 = I_{10} = k_{89} I_3 \tag{5.83}$$

$$k_{56} = \frac{(W/L)_6}{(W/L)_5}; \quad k_{89} = \frac{(W/L)_9}{(W/L)_8} \tag{5.84}$$

where k_{56} and k_{89} are the current mirroring factors, assumed equal. The above equations assume that all the transistors are in the saturation region. Moreover, matched transistors and matched mirroring factors avoid a systematic offset in the stage. In addition we have

$$\Delta V = V_{th,n} + V_{th,p} + \sqrt{\frac{2I_1}{C_{ox}}}\left(\sqrt{\frac{L_1}{\mu_n W_1}} + \sqrt{\frac{L_2}{\mu_p W_2}}\right) \tag{5.85}$$

Therefore, a given ΔV leads to a well defined quiescent current.

A possible output current results from the difference

$$I_{out} = k_{89}I_3 - k_{56}I_1 = k_{56}(I_3 - I_1) \tag{5.86}$$

remembering that

$$V_B + V_{in} = V_{GS2} + V_{GS4} = \tag{5.87}$$

$$= V_{Th,n} + V_{Th,p} + \left(\sqrt{\frac{2}{k'_n}\left(\frac{W}{L}\right)_2} + \sqrt{\frac{2}{k'_p}\left(\frac{W}{L}\right)_4}\right)\sqrt{I_2}$$

$$V_B - V_{in} = V_{GS1} + V_{GS3} = \tag{5.88}$$

$$= V_{Th,n} + V_{Th,p} + \left(\sqrt{\frac{2}{k'_n}\left(\frac{W}{L}\right)_3} + \sqrt{\frac{2}{k'_p}\left(\frac{W}{L}\right)_1}\right)\sqrt{I_1}$$

it results

$$I_{out} = k_{56}(I_1 - I_2) = \alpha k_{56} V_B V_{in} \tag{5.89}$$

CALL UP

The time allowed by a sampled data system to charge or discharge a capacitance is partaken between the slewing phase and the settling of the output voltage. The longer is the slewing the shorter is the time for settling.

where α is a proper coefficient that results from the above equations. It depends on the technology and the transistor sizing. Observe that the compensation of quadratic terms makes the output current linearly proportional to the large signal input voltage. Equation (5.89) holds until some current flows on both the unfolded pairs. When the current in one of them goes to zero the output current increases in a non-linear way as described by equation (5.78). Fig. 5.37 sketches the relationship between output current and input differential voltage. In a given interval, as predicted by (5.89) the relationship is linear. Outside the linear region the current increases faster than in the linear region. Such a feature is effective for enhancing the slew-rate, but, on the other hand, leads to harmonic distortion.

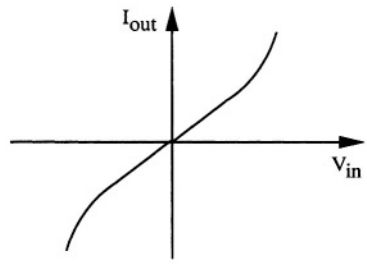


Fig. 5.37 - Output current as a function of the input voltage in the class AB stage of Fig. 5.36

Designer can derive the small signal performances of the circuit in Fig. 5.36 by inspection of the circuit. The unfolded stages generate small signal currents that, after being multiplied by the mirror factor k_{56} (= k_{89}) are summed up at the output node. The combined signal current flows into the output resistance and leads to the voltage gain

$$A_v = 2k_{56}g_{eq}r_{out} \tag{5.90}$$

The output resistance is the parallel connection of the resistances of the two Wilson current mirror. Remembering the study done in the previous chapter, we have

$$r_{out,p} = \frac{g_{m8}}{g_{m9}} \cdot \frac{r_{ds10}g_{m10}}{g_{ds3} + g_{ds8}} = \frac{1}{k_{89}} \cdot \frac{r_{ds10}g_{m10}}{r_{ds3} + r_{ds8}} \tag{5.91}$$

$$r_{out,n} = \frac{g_{m5}}{g_{m6}} \cdot \frac{r_{ds7}g_{m7}}{g_{ds2} + g_{ds5}} = \frac{1}{k_{67}} \cdot \frac{r_{ds7}g_{m7}}{g_{ds2} + g_{ds5}} \tag{5.92}$$

Equation (5.82) establishes the equivalent transconductance of the unfolded pair. Thus, the voltage gain is

$$A_v = 2k_{56}\frac{g_{m1}g_{m2}}{g_{m1} + g_{m2}}\left[\frac{1}{k_{56}}\left(\frac{r_{ds7}g_{m7}}{g_{ds2} + g_{ds5}}\right) /\!/ \frac{1}{k_{89}}\left(\frac{r_{ds10}g_{m10}}{g_{ds3} + g_{ds8}}\right)\right] \tag{5.93}$$

that, using the condition $k_{56} = k_{89}$, becomes

$$A_v = 2\frac{g_{m1}g_{m2}}{g_{m1} + g_{m2}}\left[\frac{r_{ds7}g_{m7}}{g_{ds2} + g_{ds5}} /\!/ \frac{r_{ds10}g_{m10}}{g_{ds3} + g_{ds8}}\right] \approx \gamma(g_m r_{ds})^2 \tag{5.94}$$

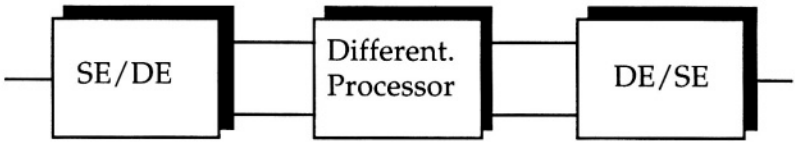


Fig. 5.38 - Fully differential signal processing chain.

Therefore, the single stage *AB* class *OTA* achieves a *dc* gain equivalent to the one of a two stage or a conventional single stage *OTA* studied so far. Observe that a possible benefit achieved with a gain in the current mirror vanishes because of an equivalent reduction in the output resistance. Thus, an amplifying factor in the current mirrors doesn't produce any benefits as far as the dc gain is concerned. Instead, higher quiescent currents in the output branch (that, keep note, produces higher power consumption) hand over higher output current for the (capacitive) load.

5.10 FULLY DIFFERENTIAL OP-AMPS

Mixed signal circuits extensively use fully-differential signal processing. We have seen that a fully-differential solution is beneficial for the clock feedthrough cancellation. We acquire similar advantages in the rejection of spur affecting the power supplies. In general, a fully differential solution rejects all the common mode components that we have at the input. Therefore, assuming that disturbing signals affect in the same extent the differential path, when considering the differential effect they are cancelled out.

In many processing systems a single-ended input and not a differential one is usually available. Moreover, numerous applications require a single ended output. Therefore, in addition to the differential processor we need a single-ended to differential and a differential to single-ended converter, as shown on Fig. 5.38. The additional blocks consume silicon area and power and, more importantly, add noise. Therefore, at the system level, it is essential to carefully estimate the global costs and the benefits of a fully-differential implementation.

5.10.1 Circuit Schematics

All the single ended schemes studied so far can be easily transformed into a fully-differential version. Often, it is just required to remove the differential to single ended converter and, possibly, use two second gain stages.

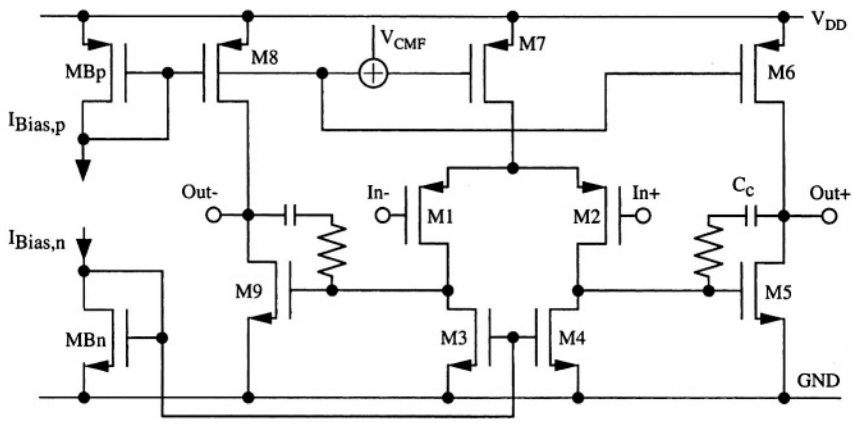


Fig. 5.39 - Fully differential two stages OTA.

Fig. 5.39 shows the fully-differential version of the two stages *OTA*. The differential to single ended conversion, achieved by connecting M_4 in the diode configuration has been removed. M_3 and M_4 both operate as active load and the dc gain of the first stage becomes

$$A_1 = \frac{1}{2}\frac{g_{m1}}{g_{ds1} + g_{ds4}}. \tag{5.95}$$

Two second stages with zero nulling compensation networks complete the scheme. Observe that the circuit in Fig. 5.39 requires one additional bias line to control the gate of M_3 and M_4. Therefore, the designer must generate that bias voltage and, more importantly, bring around another connection line.

An important function required by a fully differential scheme is the common mode feedback. We discuss this necessity by considering the circuit in Fig. 5.39. A proper operation of the network requires to balance the quiescent current of a *p-channel* transistor with the current of the equivalent *n-channel* transistor while both remain in saturation. The scheme in Fig. 5.39 has three of such a necessities: M_6 and M_8 need to match the current of M_5 and M_9 respectively; the current in M_8 demands to be equal to the sum of the currents in M_3 and M_4. Unavoidable mismatches between current references and mirror factors prevent to achieve the fittings. It is therefore necessary to use a specific circuit that takes care of the current matching issue: this circuit is called the *common mode feedback.* It detects the average value of the output voltages, compares it with a desired level and produces a controls signal that, in feedback, regulates one of the currents that we want to match.

Fig. 5.39 shows a possible input of the common mode control, V_{CMF}. Assume that the currents drained by the p-channel transistors of the second

stages are less than what the n-channel counterparts sink to ground. The common mode output voltage would drop down and the circuit doesn't operate properly. A positive V_{CMF} control solves the problem. It diminishes the overdrive of M_7 thus reducing the bias current in the input stage. Therefore, the quiescent currents of M_5 and M_9 diminish until they match the current of M_6 and M_8. A specific circuit capable to generate V_{CMF} is not discussed here. It is achieved by an inverting amplification of the common mode output voltage.

Fig. 5.40 and Fig. 5.41 show the fully differential version of the mirrored cascode and the folded cascode *OTAs*. The two schemes follow directly from the single ended versions: the connections of the cascode current mirrors have been taken away to provide the fully differential outputs. In both circuits we need to properly bias the transistor pairs M_3-M_4, M_5-M_6 and M_7-M_8. Therefore, with respect to the single ended version, the circuit requires additional bias voltages. The designer can possibly limit the design burden by using the same voltage used to control M_9 for the biasing of M_7-M_8.

The schematics in Fig. 5.40 and Fig. 5.41 highlight the points where the designer can introduce the common-mode feedback. We can achieve the nec-

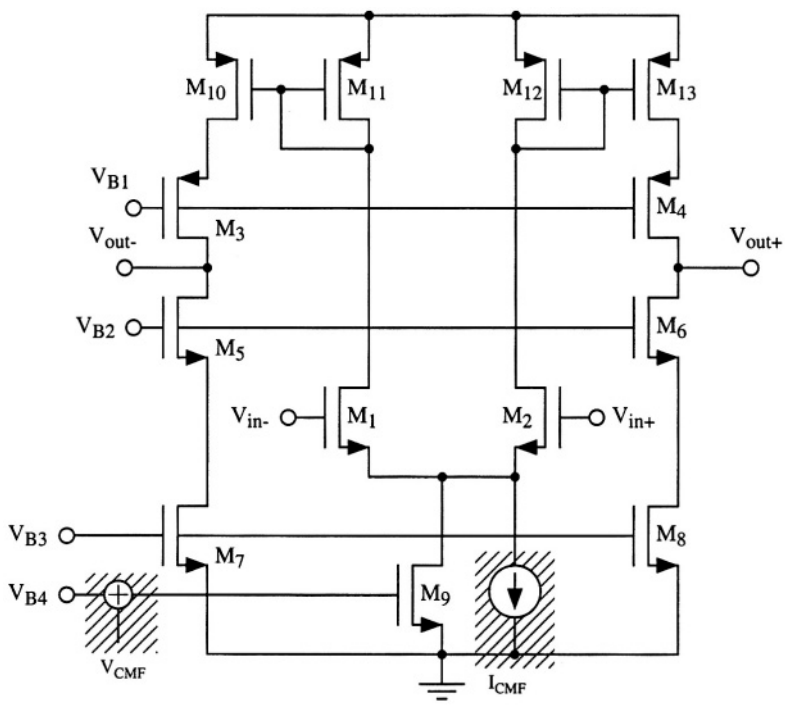


Fig. 5.40 - Fully differential version of the mirrored cascode OTA. The circuit shows two options for the common-mode feedback input: one for a voltage-mode the other for a current-mode implementation.

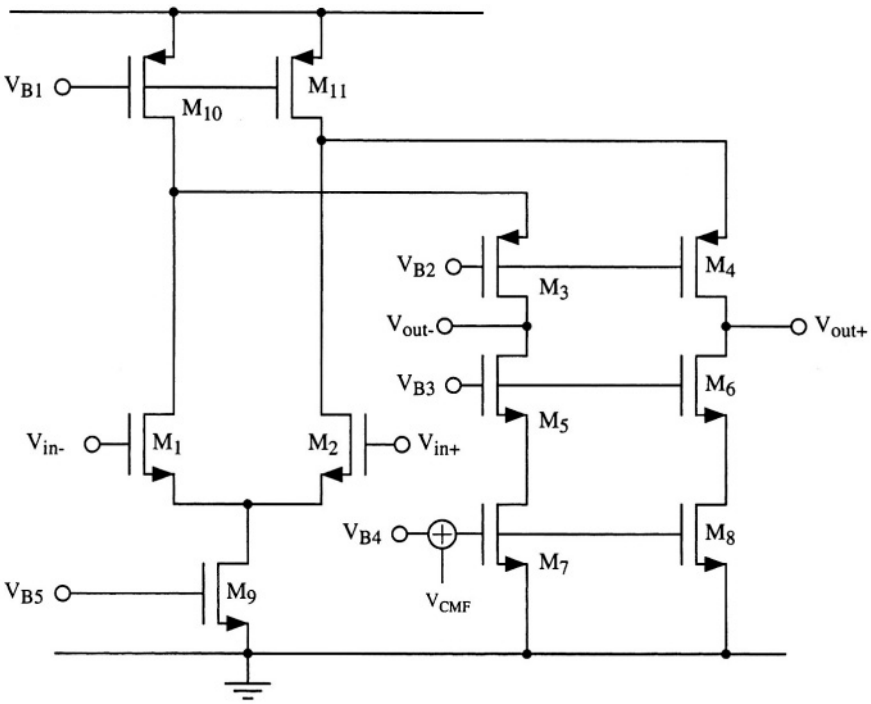


Fig. 5.41 - Fully differential version of the folded cascode OTA.

essary balancement of currents by adjusting in feedback the gate voltage of M_9 (Fig. 5.40) or the one of M_7-M_8 (Fig. 5.41). Moreover, we balance currents by a straight injection of a common-mode feedback current.

The output nodes, similarly to the single ended version, are the only high impedance nodes. All the other nodes are connected to the source of a transistor or to a diode connected element. Therefore, the pole splitting compensation or other compensation techniques are not necessary. The stability depends on the position of the dominant pole. Its location can be controlled by a possible increase of the capacitive loads at the two differential outputs.

5.10.2 Common Mode Feedback

Before studying circuit implementations of common mode feedback let us discuss again, from a different perspective, why a fully differential op-amp needs the common mode feedback. Fig. 5.42 shows a single ended application of an op-amp (or an *OTA)* and the corresponding fully differential version. In the single-ended circuit the non-inverting input is connected to the analog ground while the feedback network links the inverting terminal to the output. If

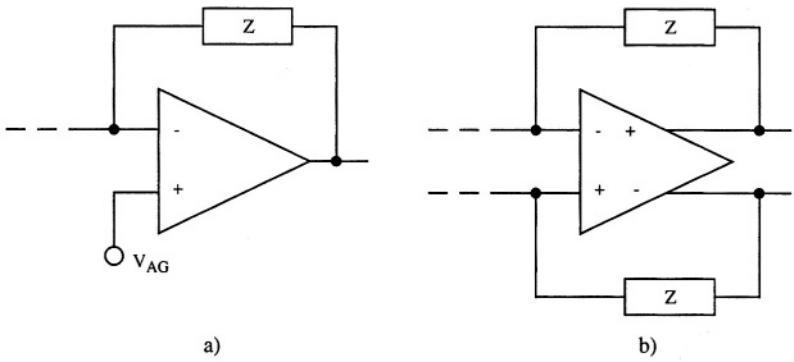


Fig. 5.42 - a) single ended use of an op-amp (or an OTA); b) fully differential version.

the voltage gain is large the differential input is low, at the limit zero. Therefore, as we normally say, the inverting terminal is virtually connected to the analog ground. In turn, since the feedback binds output and inverting terminal, even the common-mode output equals the analog ground. Therefore, the biasing of the non-inverting terminal fixes either the common-mode input and the common-mode output.

OBSERVATION

All the high impedance nodes need some control to secure their dc voltage. When the differential feedback loose its effectiveness an additional specific control loop is mandatory.

By contrast, the differential version utilizes the two input terminals to establish the feedback in both paths. Consequently, the circuit doesn't have any bind to analog ground. The *dc* gain ensures that the differential input is small or, at the limit, zero but there is no condition that fixes the quiescent voltage of input and output terminals.

Fig. 5.43 shows the basic functions of the

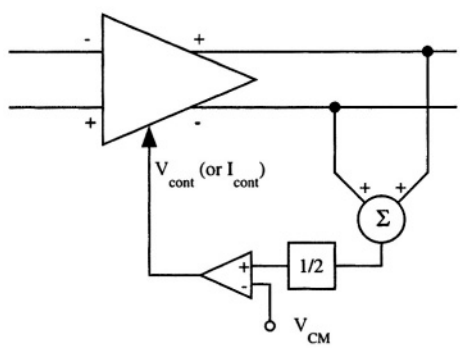


Fig. 5.43 - Basic functions performed by the common-mode feedback

common mode feedback. The adder and the 1/2 amplifier determine the common mode output; the other block compares the result with the desired common mode level; the possible error (if necessary amplified) controls in feedback a suitable node of the fully-differential architecture. We will see in the next subsections that either a continuous-time or a sampled data approach achieve the above basic functions.

5.10.3 Continuous-time Common-mode Feedback

Fig. 5.44 a) shows a simple circuit sensitive to the average value of differential outputs. It is made by a matched pair of source-coupled transistors. The output is a shifted-down replica (by $V_{GS1} = V_{GS2}$) of equal inputs; small differential changes of the inputs doesn't modify the output. The result drives the current generator, M_1 to produce a current signal. Fig. 5.44 a) doesn't incorporate the block that compares the common-mode output with a desired value. Nevertheless, the comparison results implicitly from the circuit

$$V_{CM} = V_{GS,1} + V_{GS1} \tag{5.96}$$

If the differential swing is larger than the overdrive of M_1-M_2 one of the transistors turns-off and the output follows the higher input voltage. It turns out that the circuit in Fig. 5.44 a) operates properly only for limited differential swings. Nevertheless, the solution is valuable in non-demanding cases because of its simplicity and the relative low power consumption. The circuit in Fig. 5.44 b) uses two diode connected transistors M_4 and M_5 to degenerate the pair M_1 and M_2. The range of operation of the common-mode feedback is extended at the expense of an increased V_{CM}.

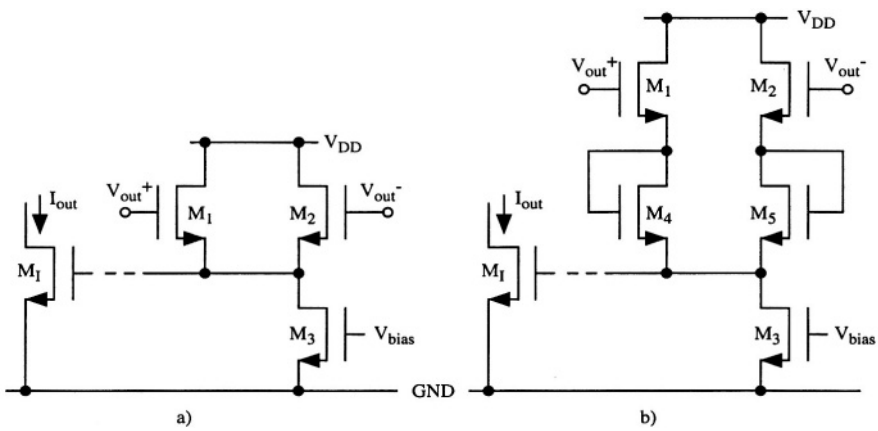


Fig. 5.44 - a) Simple source coupled common-mode feedback. b) Source coupled pair with degeneration (resistors can, possibly, replace the diode connected transistors).

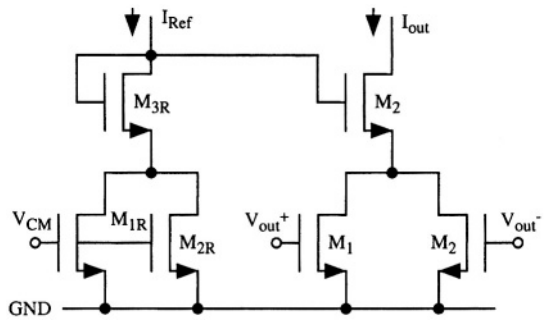


Fig. 5.45 - Use of a current mirror with adjustable mirror factor in the common-mode feedback.

Fig. 5.45 shows another solution that achieves a simple but effective common mode feedback. It is a modified version of the current mirror with adjustable mirror factor studied in Chapter 4. Two transistors, instead of one, degenerate M_{3R} and M_3.The gate of M_1 and M_2 are controlled by the output differential signals which common mode voltage must be kept equal to V_{CM}.

The transistors M_1 and M_2 are equal. Assuming them in the triode condition, we have

$$I_1 = \frac{1}{2}k\left(\frac{W}{L}\right)_1[2(V_{out+} - V_{Th})V_{DS} - V_{DS}^2] \tag{5.97}$$

$$I_1 = \frac{1}{2}k\left(\frac{W}{L}\right)_1[2(V_{out-} - V_{Th})V_{DS} - V_{DS}^2] \tag{5.98}$$

similarly, for the reference branch, being M_{1R} and M_{2R} equal and matched with M_1

$$I_{1R} = I_{2R} = \frac{1}{2}k\left(\frac{W}{L}\right)_1[2(V_{CM} - V_{Th})V_{DS,R} - V_{DS,R}^2] \tag{5.99}$$

Transistors M_3 and M_{3R} are matched; then, if the currents I_{Ref} and I_{out} are not much different the V_{DS} of the pair M_{1R}-M_{2R} equals the one of M_1-M_2. Thus, using $V_{DS,R}=V_{DS}$ in (5.99), from (5.97) and (5.98) it results

$$I_{out} = I_1 + I_2 = I_{Ref} + 2k\left(\frac{W}{L}\right)_1 V_{DS}\left(\frac{V_{out+} + V_{out-}}{2} - V_{cm}\right) \tag{5.100}$$

that corresponds to the required control. The achieved transconductance gain is $2k'(W/L)_1 \cdot V_{DS}$.

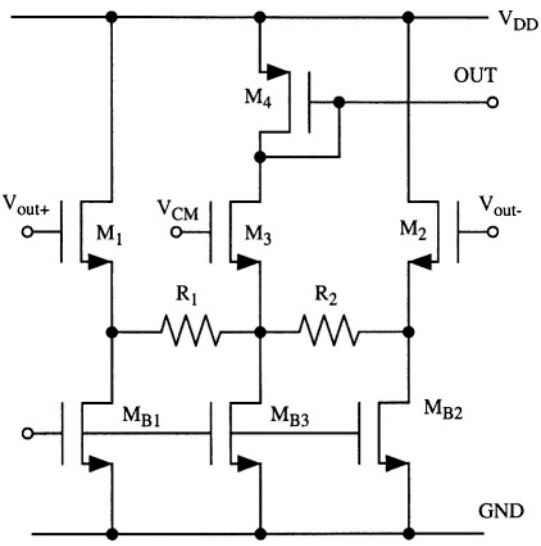


Fig. 5.46 - Common-mode feedback based on source followers.

Similarly to the solution in Fig. 5.44, the circuit in Fig. 5.45 doesn't accomplish a direct comparison of common-mode output with a desired value. The operation results implicitly because of the symmetrical action of V_{out+}/V_{out-} and V_{CM}.

Fig. 5.46 shows a more complex implementation of the common mode feedback. It uses two source follower to generate a shifted down replica of the differential outputs. Moreover, a third source follower shifts down the desired common-mode voltage, V_{CM}. The resistors R_1 and R_2, nominally equal connect the sources of transistors M_1, M_2 and M_3. If the voltages at the three gates are equal no current flows on the resistors. A common mode change of V_{out+} and V_{out-}, for example in the positive direction, causes two equal currents flowing out of the sources of M_1 and M_2. These currents will diminish by a corresponding extent the current in M_3 (and M_4) being M_{B3} a current source. Assuming one the gain of the source follower and neglecting the effect of the finite resistance of the follower, the current variation in the diode connected transistor M_4 is simply given by

$$\Delta I_{M4} = \frac{2(V_{out+} + V_{out-} - 2V_{CM})}{R_1} \tag{5.101}$$

By contrast a differential signal induces a balanced swing at the sources of M_1 and M_2. Therefore, a current proportional to the differential signal flows from the source of M_1 into the source of M_2 without affecting the branch in the middle.

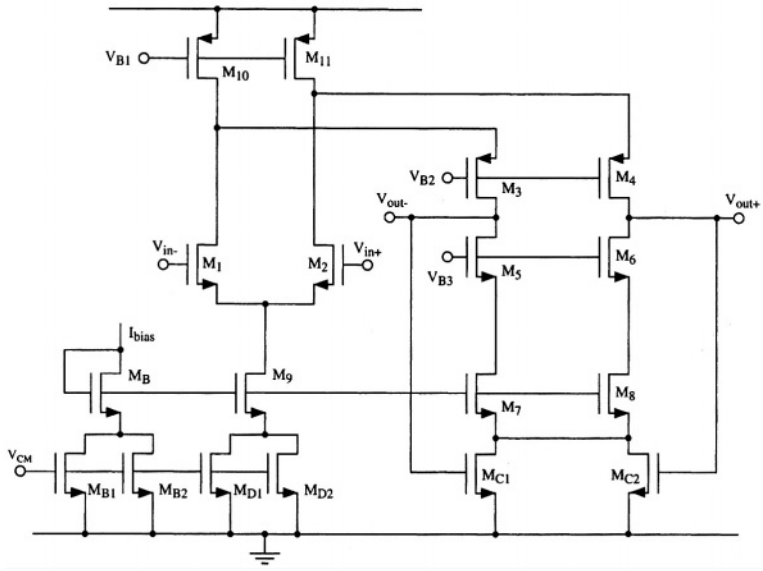


Fig. 5.47 - Fully differential folded cascode with the common-mode feedback of Fig. 5.45.

Fig. 5.47 incorporates the common mode feedback in Fig. 5.45 in the fully differential folded cascode configuration. The circuit includes the two additional transistors M_{C1} and M_{C2} in the output branch and the matching transistors M_{B1}-M_{B2} and M_{D1}-M_{D2} to degenerate the reference current and the source of the input stage. Moreover, he circuit uses the same control for both current transistors M_6 and M_7.

The circuit operates properly until M_{C1} or M_{C2} are in the triode region. Large output differential signals can bring one of them in the saturation or even in the off state. The entire current flows in the other transistor and the

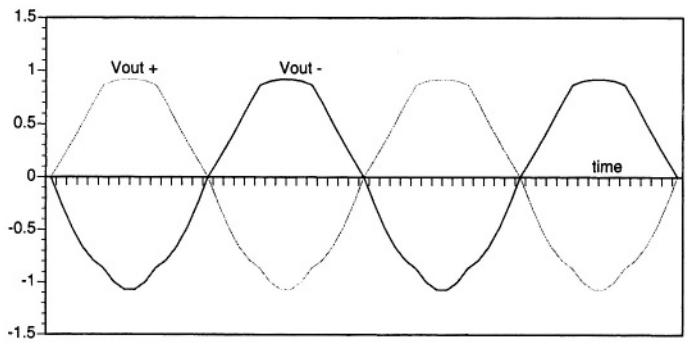


Fig. 5.48 - Typical distortion of the output differential voltages due to a poor common-mode control.

common mode control ceases working properly. The negative swing of the output doesn't compensate the positive one any more and the common mode output drops down. Fig. 5.48 shows a typical plot of the output voltage for large output differential signals. The result looks poor; however, what is important to us is the difference and not the single components shown in the figure. Actually, after the differential to single-ended conversion the signal becomes an accurate sine-wave. The limit that conies out from the response in Fig. 5.48 anyway regards the dynamic range. The output signal must remain in the region where the gain is large enough. Since a poor common mode control broadens downward the necessary dynamic range, the maximum output swing diminishes accordingly.

5.10.4 Sampled-data Common-mode Feedback

Fig. 5.49 shows how capacitors and switches can implement the basic functions required by the common mode feedback. Two complementary non-overlapped phases control the circuits. The network in Fig. 5.49 a) determines the average of the output voltages. It works as follow: during phase *1* the voltages V_{out+} and V_{out-} pre-charge the two nominally equal capacitors *C*. Then, during phase *2*, the capacitors are connected in parallel. They share their charges and generate the voltage

$$V_{parall} = \frac{Q_T}{2C} = \frac{CV_{out+} + CV_{out-}}{2C} = \frac{1}{2}(V_{out+} + V_{out-}) \qquad (5.102)$$

The circuit in Fig. 5.49 b) subtracts the two voltages V_2 and V_1. During phase *1* V_2 pre-charges the capacitor. During the next phase *2* the top plate is floating and V_1 biases the bottom plate. Therefore, the top plate pops to V_2-V_1.

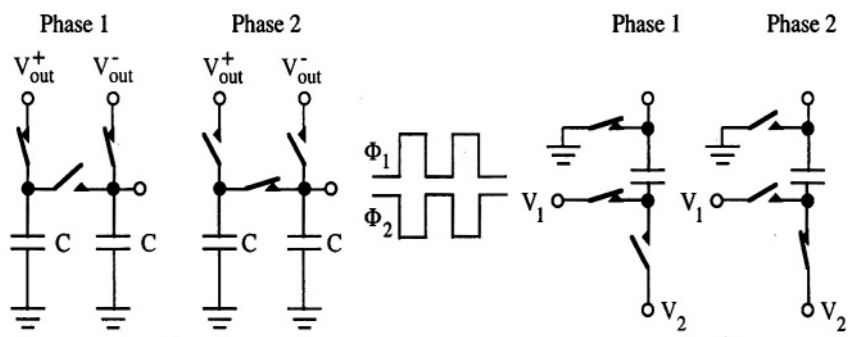


Fig. 5.49 - a) Switched capacitor network for the calculation of the average of two voltages. b) Circuit useful for the subtraction of two voltages,

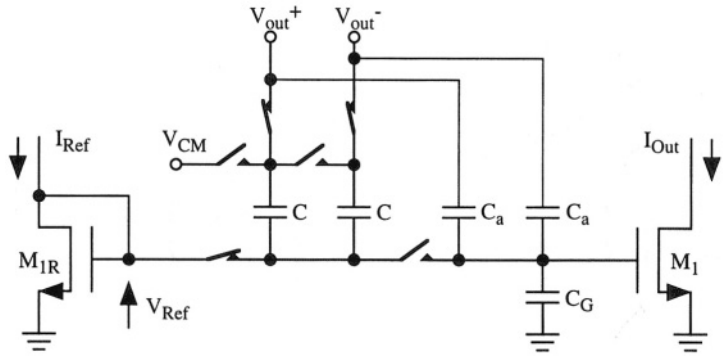


Fig. 5.50 - Sampled-data common-mode feedback. A common mode output larger than V_{CM} leads to an output current lower than the reference.

The circuit in Fig. 5.50 incorporates the two basic operation discussed above thus implementing a sampled-data common-mode feedback operation.

The reference current I_{Ref} injected in the diode connected transistor M_{1R} generates a gate reference voltage V_{Ref}. The output current would be equal to I_{Ref} if the voltage V_{Ref} were straight applied to the gate of M_1. The switched capacitor network in between the gate of M_{1R} and M_1 possibly modify this condition. During phase *1* V_{out+} and V_{out-} minus V_{Ref} charge the two equal capacitors *C*. During phase *2* (complementary to the one illustrated in the figure), a switch connects the capacitors in parallel while the top plate is connected to V_{CM}. If the average of V_{out+} and V_{out-} equals V_{CM} the voltage developed at the bottom plate becomes V_{Ref}. Now, if the average of V_{out+} and V_{out-} is larger than V_{CM} the voltage at the bottom plate becomes lower than V_{Ref}, and vice versa. Therefore, a common mode output voltage larger than what was expected diminishes the output current.

Actually during ϕ_2 the two capacitors *C* are placed in parallel with C_G. Capacitance C_G represents the gate capacitance of M_1 and an actual capacitor that is likely used to sustain the gate voltage of M_1 during the phase *1*. Capacitor C_G also smooths the changes induced every clock cycle. (The circuit acts like a switched-capacitor network whose function is equivalent to an *RC* low-pass filter).

The discussed common mode feedback measures the output voltage during one phase and controls the gate voltage during the successive phase, therefore, it doesn't react to instantaneous changes of the output voltages. Capacitors C_a, through a capacitive couplings between V_{out+} and V_{out-} and the gate of M_1, ensure a high speed path for the common mode feedback circuit.

The circuit in Fig. 5.50 diminishes the output current as response to a com-

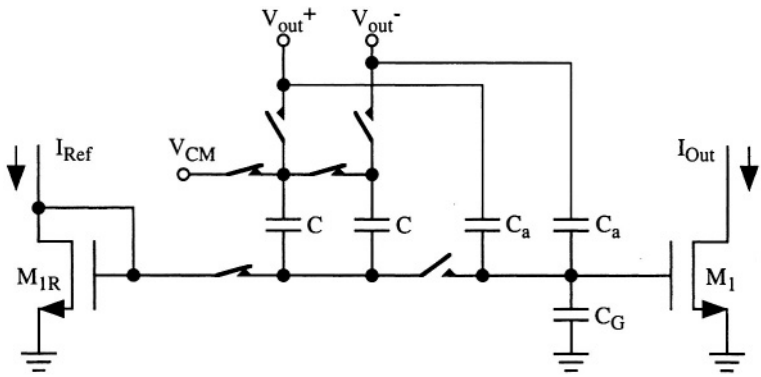


Fig. 5.51 - Sampled-data common-mode feedback with effect opposite to the one of the circuit in Fig. 5.50. A common mode output larger than V_{CM} leads to an output current higher than the reference.

mon mode higher than V_{CM}. In some cases the fully differential circuit demands for an opposite behaviour. Inverting the phases of the switches controlling the upper plate of capacitors *C* increases of the current. Fig. 5.51 shows the circuit configuration.

5.11 MICRO-POWER *OTA*'S

Many applications require very low power consumption. Portable apparatuses are powered by small batteries with a limited power capacity. Thus, to ensure the largest autonomy it is required to minimize the current (consequently, the power) in the basic blocks used. When the bias current in a *MOS* transistor becomes pretty low, the region of operation is no more the saturation but transistor enters in the sub-threshold region. Here as we have learned in Chapter 1, the current-voltage relationship is exponential and the transconductance becomes

$$g_m = \frac{I_D}{nV_T} \tag{5.103}$$

Moreover, the gain of a simple inverter with active load reaches a maximum value

$$A_v = -1/\left[n\frac{kT}{q}(\lambda_n + \lambda p)\right] \tag{5.104}$$

At room temperature, it can be around *60 dB*. Therefore, by using low bias current it is easy to obtain pretty high *dc* gains. Thus, a single stage or a two-stage amplifier are more than enough for satisfying usual gain requirements.

The key design problem results from the very small current available (typically in the ten of micro-ampere range or less). Small currents determine limited bandwidths and this can be accepted for low frequency signal processing. More important, small currents lead to small slew-rates; and this is, normally, the most severe design limitation. Therefore, when designing a micro-power *OTA* it is necessary to use specific techniques to entrance the slew-rate. The designer can use two methods:

- dynamic biasing of the current tail
- dynamic voltage biasing in push-pull stages

Below we will study both methods.

5.11.1 Dynamic-biasing of the Tail Current

The basic concept behind dynamic biasing is quite simple; to provide more current than the quiescent level when slewing needs it. To achieve the result it is necessary to use a slew-rate detector and to have a current bias boost.

In the slew-rate region the current in one branch of the input differential stage equals the tail current while the current in the other branch goes to zero. Thus, we detect the slew-rate condition by measuring the full current umbalancement in the input differential stage.

The circuit inside the dashed lines in Fig. 5.52 takes the difference between I_2 and I_2 that, if positive, flows into the sensing transistor $M_{8,1}$. The similar circuit (made by $M_{3,2}$-$M_{8,2}$) permits us to sense (I_1-I_2) (if positive). The com-

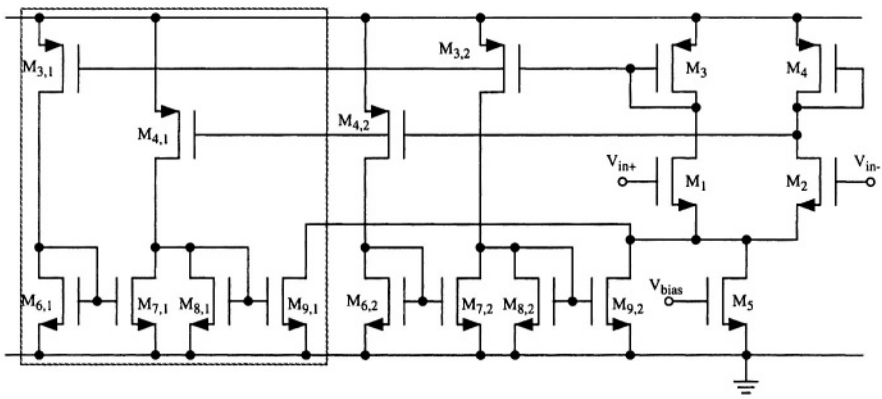


Fig. 5.52 - Dynamic biasing used in a simple gain stage.

bination of the two results, formed by the mirroring elements $M_{9,1}$ and $M_{9,2}$, provide $k|I_1\text{-}I_2|$. Where k comes from the mirror factors used.

The right part of the circuit is the input stage of the *OTA*. Possibly, a mirrored output stage completes the scheme. Assuming M_1 and M_2 in the subthreshold region the differential current is given by

$$|I_1 - I_2| = (I_D + I_{B1} + I_{B2}) atan \frac{|v_{in}|}{nV_T} \tag{5.105}$$

but

$$I_{B1} + I_{B2} = k|I_1 - I_2| \tag{5.106}$$

combining (5.105) and (5.106) it results

$$|I_1 - I_2| = \frac{I_D atan \frac{|v_{in}|}{nV_T}}{1 - k atan \frac{|v_{in}|}{nV_T}} \tag{5.107}$$

that denotes, as results from inspecting the circuit, a positive feedback that brakes into a regenerative grown when

$$k \cdot atan \frac{|v_{in}|}{nV_T} = 1 \tag{5.108}$$

Therefore, in order to have a significant increase of the current, *k* must be larger than *1* when the input differential voltage is comparable to nV_T. Note that the external feedback network used in the circuit controls the input differential voltage. At a given time, after slewing, the input umbalancement goes down and it becomes low enough to extinguish the regenerative feedback. Then, the current bias boosting turns off.

5.11.2 Dynamic Voltage Biasing in Push-pull Stages

A push-pull scheme is fairly adequate for low-power applications. It allows the use of very low currents and works with relatively low supply voltages.The *AB* class operation also provides large current when the op-amp is required to sustain the slewing phase.

Fig. 5.53 shows a fully-differential scheme based on the same approach used in Fig. 5.36. Two cross-coupled unfolded differential stages provide signal cur-

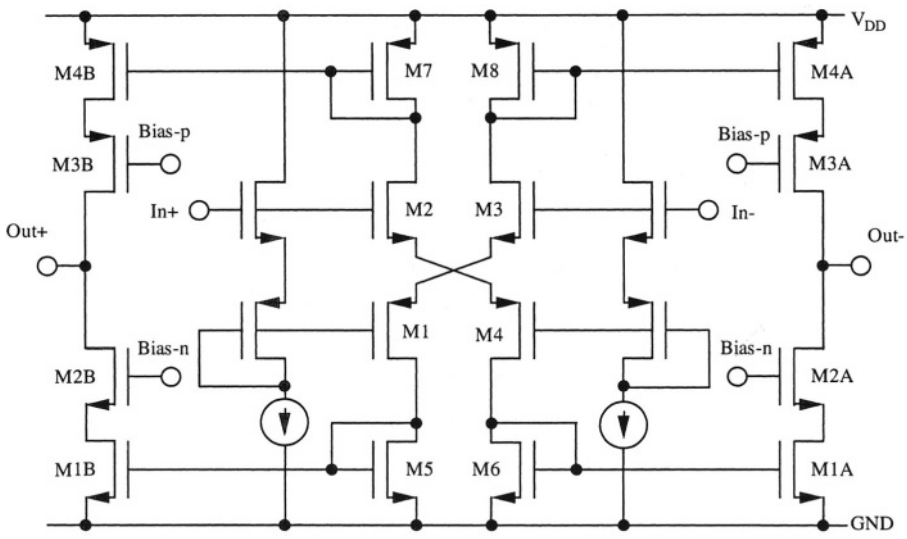


Fig. 5.53 - Fully-differential class AB

rents that, after mirroring, control the cascode output stages. The input section determines the minimum supply voltage. Assuming that the V_{DS} of the transistors of the unfolded pairs is kept a its minimum, we have

$$V_{DD,min} = V_{GS,n} + V_{GS,p} + V_{sat,n} + V_{sat,p} \tag{5.109}$$

that, for typical technologies, can be as low as 2 V or so.

Observe that, in order to ensure low voltage, the circuit uses a simple diode connected transistor to detect the signal current in the input stage. Moreover, suitable bias voltages control the gates of the output cascodes. The input differential signal controls the n-channel input directly. Two level shifters shift down the input voltages by $V_{GS,n}+V_{GS,p}$ to provide the control of the p-channel inputs. For an optimum operation the crossing point in the input stage must be at $V_{DD}/2$. This means that the common mode input must be V_{GS2} higher than $V_{DD}/2$. By contrast the optimum common mode output is at $V_{DD}/2$. The above discrepancy is a minor problem and typically is solved with proper level shifters at the system level.

The differential gain is given by the transconductance of the input pair multiplied by the mirror factor and the output resistance

$$A_v = \frac{(W/L)_{4A}}{(W/L)_{8A}} \cdot \frac{g_{m1}g_{m3}}{g_{m1}+g_{m3}} \cdot \frac{g_{m3A}}{g_{ds3A}\cdot g_{ds4A}} // \frac{g_{m2A}}{g_{ds2A}\cdot g_{ds1A}} \tag{5.110}$$

that is pretty large especially with micro-currents.

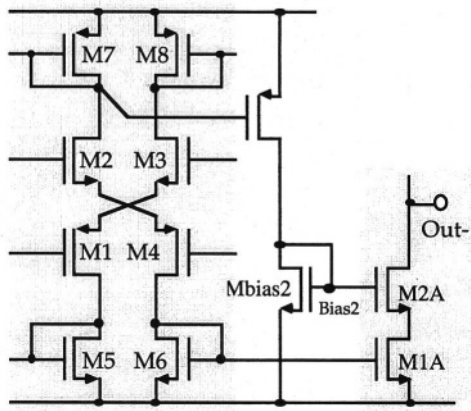


Fig. 5.54 - Dynamic control of the bias voltages in micro-power push-pull OTA.

The main design problem concerns the bias of the cascoding transistors. In order to maximize the output swing the bias voltages $V_{bias\text{-}p}$ and $V_{bias\text{-}n}$ must approach as close as possible to the rail voltages. However, during the slewing conditions the boost of the overdrive of transistors in the output stage doesn't lead to a substantial increase of the slew current. The restriction to V_{DS} caused by the cascoding elements leads the boosted transistors into the triode region. A solution to the problem comes from the use of a dynamic control of the bias voltages. During the slewing phase the dynamic control must increase the V_{DS} voltages and keep the boosted transistors into saturation.

Fig. 5.54 shows a possible implementation of the dynamic control of bias voltages. It applies to the cascode M_{1A}-M_{2A}. The other three cascodes will use similar circuits. The bias voltage V_{Bias2} results from a diode connected transistor whose current is a replica of the current in M_7. A suitable choice of the aspect ratio of M_{bias2} will determine the right V_{Bias2}. During slewing the current in M_6 increases. Accordingly the current in M_7 growths. Therefore, the augmented bias current in M_{bias2} pulls up V_{Bias2}. This happens until the circuit remains in the slewing phase. In the normal conditions of operation I_{M7} returns to its quiescent level and V_{Bias2} goes back to the nominal value.

5.12 NOISE ANALYSIS

The noise performances of the OTAs studied in the previous sections are controlled by the noise of the single transistors used in the circuit. As already discussed the noise analysis assumes uncorrelated the noise generators of dif-

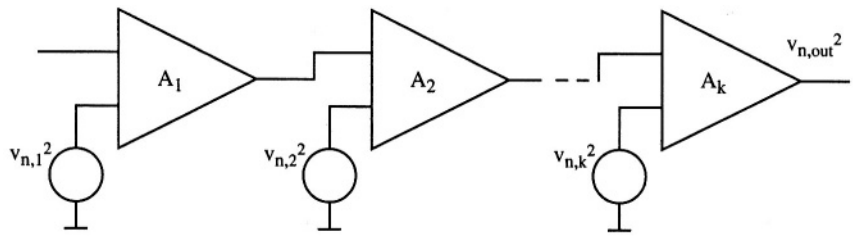


Fig. 5.55 - Noise representation in a cascade of gain stages.

ferent transistors. Therefore, their effect is superposed quadratically.

For the cascade of gain stages the most important contribution comes from the first stage. Assuming to represent the noise of each stage with an input referred generator as shown in Fig. 5.55, we have

$$v_{n,out}^2 = (A_1 A_2 \ldots A_k)^2 v_{n,1}^2 + (A_2 \ldots A_k)^2 v_{n,2}^2 + \ldots + A_k^2 v_{n,k}^2 \qquad (5.111)$$

referred to the input of the chain, it results in

$$v_{n,in}^2 = v_{n,1}^2 + \frac{v_{n,2}^2}{A_1^2} + \ldots + \frac{v_{n,k}^2}{(A_1 A_2 \ldots A_{k-1})^2} \qquad (5.112)$$

that proves the above statement under the assumption that the gain in the first stage is large enough and expecting comparable input referred noise generators.

Fig. 5.56 shows the first stage of a two-stage amplifiers. It outlines the relevant input referred noise generators. The noise associated to M_5 is not included since it yields a minor effect on the output voltage. In fact, the noise generator of M_5 produces a noise current that, thanks to the differential input stage, flows in the two branches evenly. The pair M_3-M_4 mirrors half of the noise current and injects it into the output node. As a result the subtraction of two fully correlated terms following in the output node produces a null contribution. Therefore, the noise of M_5 can be neglected up to frequencies at which the current mirror operates properly.

The output noise caused by the generators in Fig. 5.56 is estimated by

$$v_{n,out}^2 = [(v_{n1}^2 + v_{n2}^2) g_{m1}^2 + (v_{n3}^2 + v_{n4}^2) g_{m3}^2] \left(\frac{1}{g_{ds2} + g_{ds4}} \right)^2 \qquad (5.113)$$

The input referred noise is derived by dividing $v^2_{n,out}$ by the square of the gain of the stage

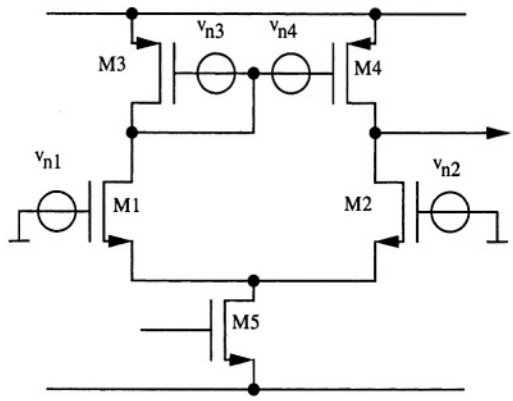


Fig. 5.56 - Schematic for the noise analysis of the first stage of a two-stages OTA.

$$v_{n,in}^2 = \frac{v_{n,out}^2}{g_{m1}^2}(g_{ds2} + g_{ds4})^2 = 2\left[v_{n1}^2 + \left(\frac{g_{m3}}{g_{m1}}\right)^2 v_{n3}^2\right] \tag{5.114}$$

Assuming the transistors in saturation, we can use for the transconductance the following relationship

$$g_m = \sqrt{2\mu C_{ox}\frac{W}{L}I} \tag{5.115}$$

Moreover, according to equations (1.69) and (1.70), the white and the *1/f* term of the spectrum of the noise generator are given by

$$v_n^2 = \left(\frac{8kT}{3}\frac{}{g_m} + \frac{K_F}{2\mu C_{ox}}\frac{1}{WL}\cdot\frac{1}{f}\right)\Delta f \tag{5.116}$$

using the white component of the noise and (5.115) in equation (5.114) it results

$$v_{n,in,white}^2 = 2v_{n1}^2\left(1 + \frac{g_{m3}}{g_{m1}}\right) = 2v_{n1}^2\left(1 + \sqrt{\frac{\mu_3(W/L)_3}{\mu_1(W/L)_1}}\right) \tag{5.117}$$

indicating that the white component of the input referred noise, $v^2_{n,in,white}$, equals the noise of the input pair multiplied by a proper factor. Often it is required to have the input referred noise dominated by the input pair. Therefore, the multiplying factor should be kept at the minimum. By inspection of (5.117) we obtain the following design conditions:

- the mobility in the input pair should be larger than the one in the active load

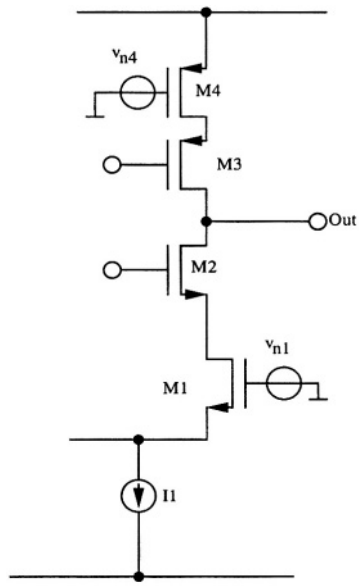


Fig. 5.57 - Schematic for the noise analysis of a telescopic cascode OTA

(that implies the use of *n-channel* transistors at the input);

- the aspect ratio of the input pair should be larger than the one of the active load (that satisfies the request to have a large transconductance gain).

The use of the expression of the *1/f* component of the noise leads to

$$v_{n,in,1/f}^2 = 2\frac{K_{F1}}{\mu_1 C_{ox} W_1 L_1}\cdot\frac{1}{f}\left(1+\frac{K_{F3}L_1^2}{K_{F1}L_3^2}\right) \qquad (5.118)$$

Again, the input referred noise is given by the noise of the input transistor enlarged by a multiplier factor. The designer keeps it at the minimum using a length of the active load larger than the one of the input pair.

Fig. 5.57 shows a schematic useful to estimate the noise performances of a telescopic cascode. The figure depicts half of the circuit only (the other half brings in the same amount of noise). Moreover, the figure shows the noise generators of M_1 and M_4 only. The noise of

NOISE OPTIMIZATION TIPS

- Use of an *n-channel* input pair to minimize the white noise.
- Employ a (W/L) of the input pair larger than one of the active load.
- Use an active load longer than the input pair (to minimize the 1/f noise).

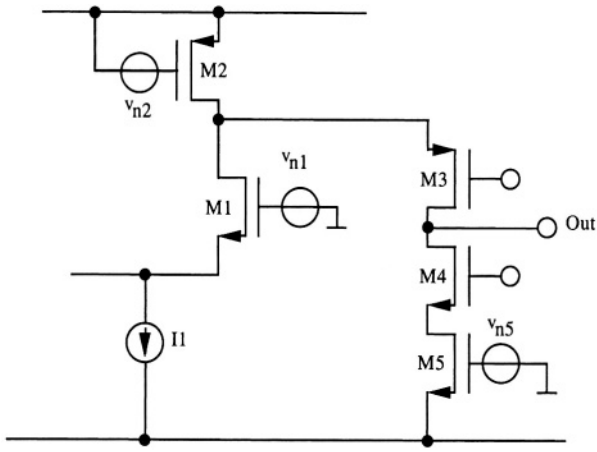


Fig. 5.58 - Schematic for the noise analysis of a folded cascode OTA

the current generator I_1 brings about a negligible effect for the same reasons discussed in the simple differential amplifier. The noise generators of M_2 and M_3 are not considered because the gain from their gates to the output is considerably lower than the gain from the gates of M_1 and M_4. In fact, by inspection of the circuit, one obtains

$$A_1 = g_{m1}R_{out}; \qquad A_4 = g_{m4}R_{out} \tag{5.119}$$

$$A_2 = g_{ds1}R_{out}; \qquad A_3 = g_{ds4}R_{out}. \tag{5.120}$$

The input referred noise becomes

$$v_{n,in}^2 = 2\left[v_{n1}^2 + \left(\frac{g_{m4}}{g_{m1}}\right)^2 v_{n4}^2\right] \tag{5.121}$$

that equals the equation that we obtained for the simple differential stage. Therefore, the designer should follow for input pair and the transistors connected to the positive supply voltage the same recommendation given for the input pair and the active load of Fig. 5.56.

In the folded cascode scheme (see Fig. 5.58) we have an additional transistor that contributes to the noise. It is the transistor used to achieve the current generator M_2. The analysis of the circuit leads to

$$v_{n,in}^2 = 2\left[v_{n1}^2 + \left(\frac{g_{m2}}{g_{m1}}\right)^2 v_{n2}^2 + \left(\frac{g_{m5}}{g_{m1}}\right)^2 v_{n5}^2\right] \tag{5.122}$$

therefore, it is necessary to follow consistent design recommendations for both

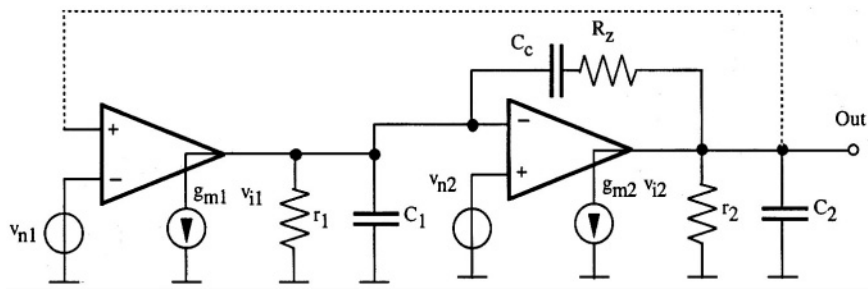


Fig. 5.59 - Equivalent circuit for estimations the noise transfer functions in a two-stages OTA.

transistors M_2 and M_5 in Fig. 5.58.

The previous analysis didn't accounted for the frequency behaviour of the gain stages and the effect of the feedback network. Fig. 5.59 shows the equivalent circuit of a two stages amplifier whose external feedback establishes (for noise calculation purposes) a unity gain configuration. The analysis of the network in Fig. 5.59 or a Spice simulation lead to the transfer functions for the two noise generators. Fig. 5.60 shows the results of a Spice simulation. The overall gain is 80 dB, the gain-bandwidth is 8 MHz and the phase margin is 60°. The transfer function from the first stage shows that the noise is conveyed at the output with a unity gain until the f_T of the OTA. Two poles, f_T and f_2 determine a roll-off of the transfer function. Since f_T and f_2 are pretty close one to the other the roll-off of the Bode diagram rapidly becomes *-40 dB* per

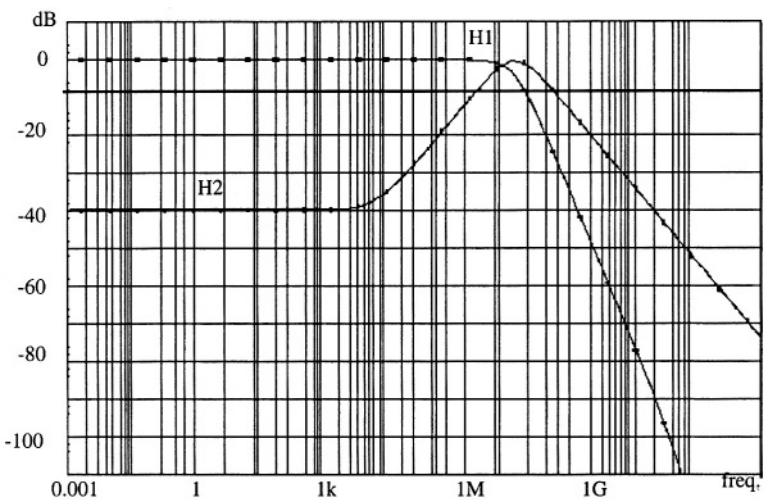


Fig. 5.60 - Noise transfer functions of a typical two-stages OTA.

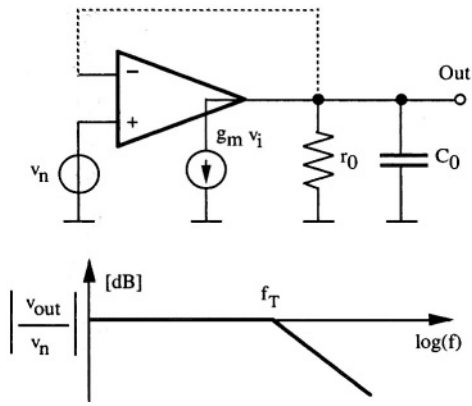


Fig. 5.61 - Equivalent circuit for estimating the noise transfer functions in a single-stages OTA.

decade. The noise transfer function form the input of the second stage has a different behaviour. As expected at low frequency the gain is $1/A_2$ (- *40 dB*). The Bode diagram shows a zero at a given frequency and two poles at f_T and f_2. The zero stays approximately at f_T/A_2. The result confirms that the noise performances are dominated by the first stage. However, at frequencies around the gain-bandwidth the noise provided by the second stage can become comparable with the one of the first stage.

Fig. 5.61 shows the equivalent circuit for estimating the noise transfer function in a single stage OTA. The model used outlines a single pole operation. Therefore, the input referred noise generator influences the output up to f_T. After f_T the noise transfer function rolls down by *20 dB* per decade.

Fig. 5.61 show a low pass filtering action of the op-amp (or OTA) over the input referred noise spectrum. Therefore, independent on the bandwidth of the load, the noise brings about a finite power. The maximum noise power results from the integral over the infinite interval of the output noise spectrum. Assuming the white term of the input spectrum represented by $v^2_{in} = 2(1+\alpha)(8/3)kT/g_m$ one obtains

$$\overline{V_{n0}^2} = \int_0^\infty v_n^2 \frac{df}{\left|1 + s/(2\pi f_T)\right|^2} = \tag{5.123}$$

$$= 2(1+\alpha)\frac{8}{3}kT\int_0^\infty \frac{1}{g_{m1}}\frac{df}{1 + (2\pi f C_0/g_{m1})^2} =$$

$$= 2(1+\alpha)\frac{8}{3}kT\frac{1}{2\pi C_0}\int_0^\infty \frac{dx}{1+x^2} = \frac{4}{3}2(1+\alpha)\frac{kT}{C_0}$$

where we assume $f_T = 2\pi C_0/g_{m1}$.

Equation (5.123) shows that the noise power depends on the load capacitor of the op-amp and the factor α, denoting the noise exceeding the input pair contribution. The fact that the result doesn't depend on the transconductance of the input pair can be confusing. To justify the result, observe that an increase of the transconductance leads to a lower noise but, at the same time, the bandwidth of the noise transfer function increases.

5.13 LAYOUT

The previous chapters examined the techniques for laying out a single transistor or a basic building block. This section considers the more comprehensive issue of laying-out a complete op-amp or an OTA. The performances of circuits studied so far critically depends on the electrical design. Additionally, any design presumes matching between transistors and symmetry between sections. Moreover, it is expected that parasitics and spur couplings produce negligible effects. Therefore, a properly executed layout is the key to comply the design targets.

In general, the layout should conform with the following guidelines:

- to procure the same symmetry that the circuit has at the electrical level;
- to ensure that transistors supposed to be equal or assumed to have a given aspect ratio match at the geometrical and technological level;
- to bring at minimum and/or to match the parasitic drops voltage across interconnections;
- to obtain balanced paths in signal interconnections;
- to avoid (or to make minimum when unavoidable) capacitive couplings especially the one with high impedance nodes.

Symmetry in circuits are important to minimize the random offset and to reject common-mode unwanted signals. Since the random offset comes from the mismatch of paired transistors, inter digitized or, even better, common centroid arrangements of critical components are necessary.

5.13.1 Parasitic Effects

Spurs affect circuits in different ways. They appear because of the interference between interconnection lines. They come from the substrate or can be the consequence of opening a switches. Therefore, the protection from spur signals involves both transistors and interconnections. The use of symmetrical

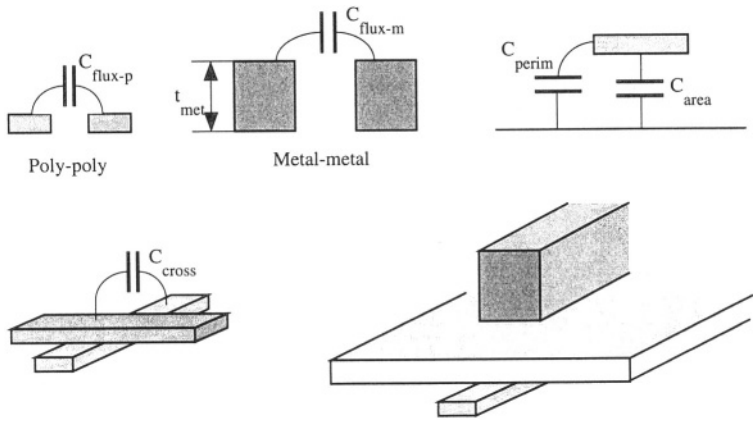


Fig. 5.62 - Possible parasitic coupling affecting metal and poly interconnections.

layouts made by two mirrored half-cells helps in obtaining balanced paths and symmetrical features. However, a mirrored layout ensures solely a symmetry axis; a gradient in direction orthogonal to that axis causes mismatch. Therefore, it is probably better to use small mirrored half-cells wired with balanced interconnections.

Fig. 5.62 shows the capacitive couplings that can occur in a typical layout. A flux capacitance affects two lines running in parallel. For polysilicon strips at $0.5\ \mu$ distance the value of the specific capacitance $C_{flux\text{-}p}$ is typically $50\ aF/\mu$. Therefore, two lines running in parallel at the minimum separation for $200\ \mu$ can be affected by a *10 fF* link. Moreover, the thickness of metal lines, $t_{met} \approx 0.5\text{-}1.5\ \mu$, is larger than the height of poly strips by a factor *2* to *3*. Even if the minimum distance between metal is larger than the separation of poly the flux capacitance $C_{flux\text{-}m}$ can be larger than $C_{flux\text{-}p}$.

Since the metal-metal flux capacitance is large, some technologies exploit

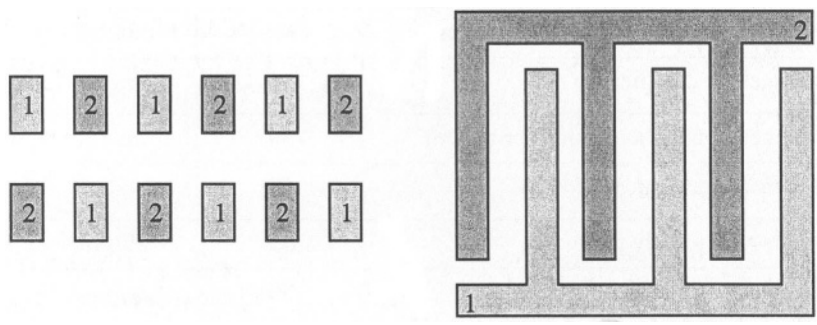


Fig. 5.63 - Cross section and layout of a flux capacitor using two metal layers.

the feature to obtain capacitors. The structures used connect in parallel the various metal layers available so that the flux capacitance enlarges. Fig. 5.63 shows the cross section and the layout of a possible flux capacitor. The structure uses two metal layers. An interleave in the horizontal and vertical directions of the two metal achieves plate land plate 2.

Table 5.2 provides the parasitic parameters for a typical $0.25\,\mu$ CMOS technology. Observe that the parasitic in the vertical direction between substrate and connection lines or between different conductive layers depends on the area and, because or the fringing effects, on the perimeter of the considered geometry. Therefore, the estimation of the coupling due to the intersection of connecting wires must account for both crossing area and perimeter. Observe that a relatively narrow line (below $1\,\mu$) brings about fringing parasitic larger than the direct coupling. For example, a $0.5\,\mu$ *metal-1* line running over the field oxide for $100\,\mu$ leads to *2 fF* for the overlap and *6fF* for the fringing coupling. Moreover, the fringing term dominates the capacitive coupling in the crossing of a $1\,\mu$ *metal-1* and a $1\,\mu$ *metal-2*. The vertical coupling, *30 aF,* is a small fraction of the total, *230 aF.*

TABLE 5.2 - *Parasitic Capacitances*

Parameter	Value	Unit
Poly-poly flux	50	aF/μm^2
Metal1-metal2 flux	70	aF/μm^2
Metal2-metal2 flux	80	aF/μm^2
poly over field oxide area	40	aF/μm^2
poly over field oxide perimeter	30	aF/μm
metal1 over field oxide area	30	aF/μm^2
metal1 over field oxide perimeter	40	aF/μm
metal 2 over field oxide	15	aF/μm^2
metal2 over field oxide perimeter	30	aF/μm
metal1 to poly area	20	aF/μm^2
metal1 to poly perimeter	35	aF/μm
metal2 to metal1 area	30	aF/μm^2
metal2 to metal1 perimeter	50	aF/μm

Albeit the estimated values of parasitic capacitance are quite low, their effect is significant for analog applications. Example 5.8 shows how an improper layout induces significant spur on sensitive analog lines.

Example 5.8

A metal-2 line overlap for 50 μ a metal-1 line. The width of both lines is 0.5 μ. The first line carries a digital clock that switches from 0 to 3.3 V. The second line connects to the gate of an analog transistor which input capacitance is 0.2 pF. Calculate the disturbance caused by the parasitic coupling. Use the parasitic parameters given in Table 5.1.

***Solution*:**

The coupling comes from the overlap area and the fringing contribution

$$C_p = (30 \cdot 25 + 50 \cdot 100)\text{aF} = 2.25\text{fF}$$

The digital clock, attenuated by the capacitive divider C_p-C_{in}, affects to the gate of the analog transistor by

$$V_{gate} = V_{clock}\frac{C_p}{C_{in}+C_p} = 36.7mV$$

which is a large number for any analog applications.

In any schematic an interconnection is intended to be a wire with zero resistance. Unfortunately, in real circuits that is not true. Metal lines achieve interconnections with a given series resistance and, relevant for high frequency operation, a parasitic capacitance. The specific resistance of metal depends on the technology: interconnections may use aluminium or copper; the thickness of metal lines can be a fraction or more than one micron. As a result, the specific resistance of metals varies over a pretty large range. A typical value is around *0.1 Ω/square*.

The width of the interconnections must account for the carried current. To avoid electromigration (that affects the reliability) it is necessary to ensure a given width per unity current. Table 5.3 furnishes typical parasitic parameters for metal interconnections. Observe that in addition to the metal connections it is necessary to account for the resistance of via when interconnections use different metal layers.

A wire of few ten of squares produces some Ω of parasitic resistance. When a relatively large currents flows, the drop voltage across the wire can be compa-

TABLE 5.3 - *Parasitic Resistances and Current Densities*

Parameter	Value	Unit
Poly-1 specific resistance	20	Ω/□
Metal-1 specific resistance	100	mΩ/□
Metal-2 specific resistance	70	mΩ/□
Copper Metal specific resistance	20	mΩ/□
Metal-1 poly contact resistance	10	Ω/contact
Metal-1 metal-2 via	1	Ω/via
Poly current density	0.4	mA/μm
Metal-1 current density	1	mA/μm
Metal-2 current density	1.5	mA/μm
Via current	1	mA/via

rable to the offset caused by the threshold mismatch in adjacent transistors. Widening metal lines reduces the series resistance. However, the parasitic coupling increases possibly limiting the frequency performances or augmenting the spur injection. Therefore, it is necessary to find the best trade-off between the parasitic resistance and the capacitance limitation.

Fig. 5.64 compares a layout that leads to mismatched interconnections and a more suitable implementation. The current toward the sources of M_1 and M_2 flows from the left side of the layout. Therefore, the voltage at the source of M_1 is $V_S - R_1(I_1+I_2)$. Moreover, the difference between the voltage of the sources of M_1 and M_2 is R_2I_2. Assuming M_1 and M_2 matched and in saturation

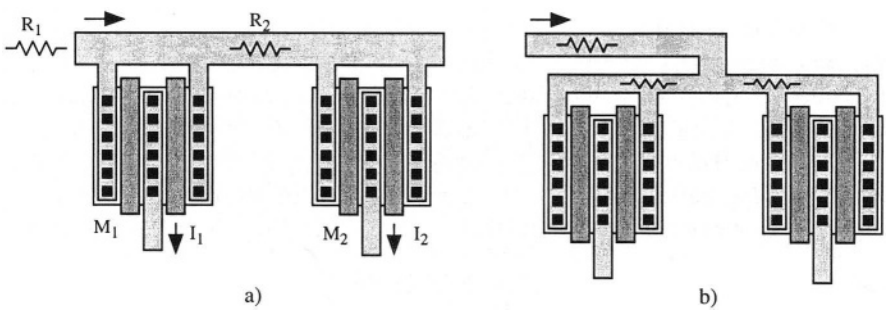


Fig. 5.64 - a) Interconnection of matched transistors causing an offset error.
b) Interconnection matching the drop voltage across the metal lines.

the current I_2 becomes

$$I_2 = k\left(\frac{W}{L}\right)(V_{GS1} - R_2 I_2 - V_{Th})^2 = I_1\left(1 - \frac{2R_2 I_2}{V_{GS1} - V_{Th}}\right) \qquad (5.124)$$

If, for example, the drop voltage across R_2 is 1 *mV* and the overdrive voltage is *200 mV* the current mismatch is *1%*. Actually, what degrades the circuit performances is not the drop voltag $R_1(I_1+I_2)$ but $R_2 I_2$. Therefore, instead of using wider metals it can be more effective to match the parasitic drop voltages. Fig. 5.64 b) shows the modified the layout. The current flows through the common wire until a junction point; from that point the connections to the sources of M_1 and M_2 are symmetrical. The use of two metals leads to a same result.

5.13.2 Stacked Layout

Splitting transistors in a number of fingers favours a stack arrangement and improves the layout matching. For example Fig. 4.3 presented the layout of a simple current mirror. The four finger of each transistor were interleaved so that the centroid of the two transistors are one close to the other. The arrangement of the stack was *AABBAABB* (where *A* and *B* represent the fingers of M_1 and M_2 respectively). An alternative organization was *ABBAABBA* that lead to an identical common centroid. However, the boundary conditions are not symmetrical: two fingers of M_1 establish the two boundaries while M_2 has all the fingers inside the array.

The above layout strategy can be generalized for more complex cells. However, it is important to work on a design that favours the stacked approach. The transistors' finger that should be laid-out on the same stack must have the same width. This is often possible: all the circuits include transistors which sizes are not particularly critical. A small change of the widths don't modify the performances but permits a better layout.

We will discuss the stacked layout technique with examples. Let us consider the two stages OTA of Fig. 5.65. The scheme specifies the number of fingers and the (W/L) of each transistor. However, the numbers used are not real dimensions. A scaling factor : *2* is used. Therefore, the fingers of the sizes of input pair are $W/L = 10\mu/0.4\mu$. The input pair sums up *12* fingers. The transistor M_{N2} and M_{N3} have *12* fingers; the same is valid for the trio M_{P0}, M_{P1}, and M_{P2}.The above observation suggests to arrange the layout using three stacks

CCDDEEEEEEEE
XABBAABBAABBAX
FFFFGGGGGGG

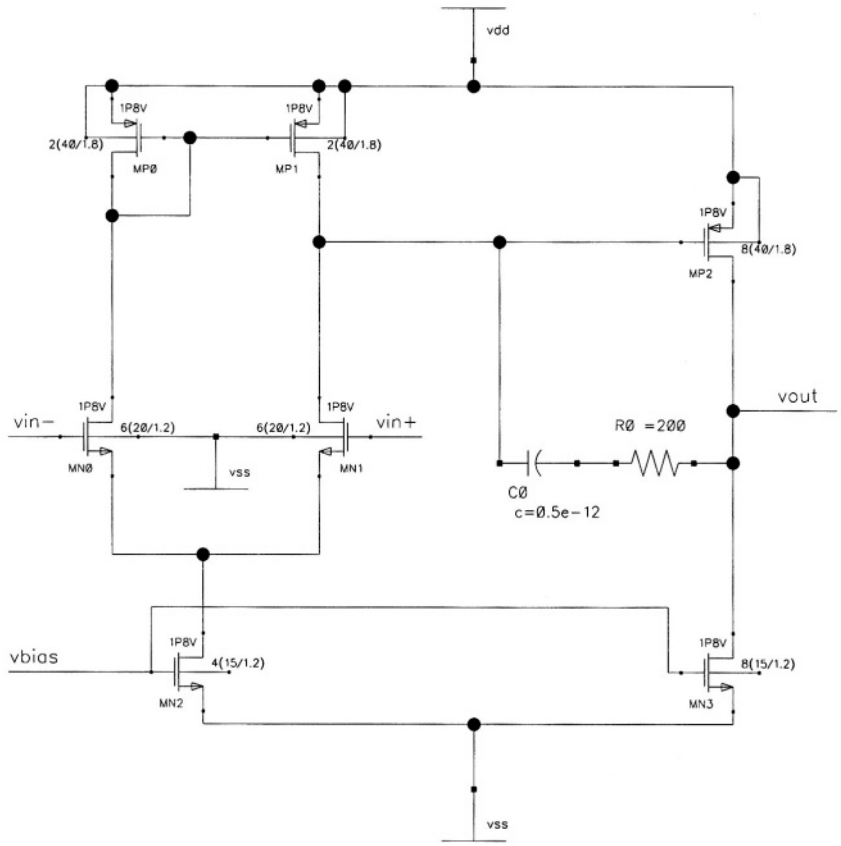


Fig. 5.65 - Circuit schematic of a two-stages OTA. The transistor sizes are convenient for a stacked layout. Use a shrink factor of 2 to have the real dimensions.

The top one arranges the fingers of the p-channel transistors. The letters *C*, *D,* and *E* represent the fingers of M_{P0}, M_{P1}, and M_{P2} respectively. The middle stack achieves the input pairs with the arrangement *ABBAABBAABBA* and two dummy fingers, *X,* at the endings. The third stack includes the tail current generator and the n-channel transistor of the second stage.

Note that the ending node of every pair of fingers is the same. Therefore, if the left terminal of the stack *CCDDEEEEEEEE* is V_{DD}, a connection to V_{DD} every second finger is required. Therefore, it is possible to interchange pair of fingers and obtain different arrangements. For example, the first stack can become *EEEECCDDEEEE* or any other convenient combination. Different finger sequences lead to a corresponding interconnection. Some routing are more problematic than other. This makes the choice of a stack arrangement a trade-off between transistor matching and interconnection balancement. For

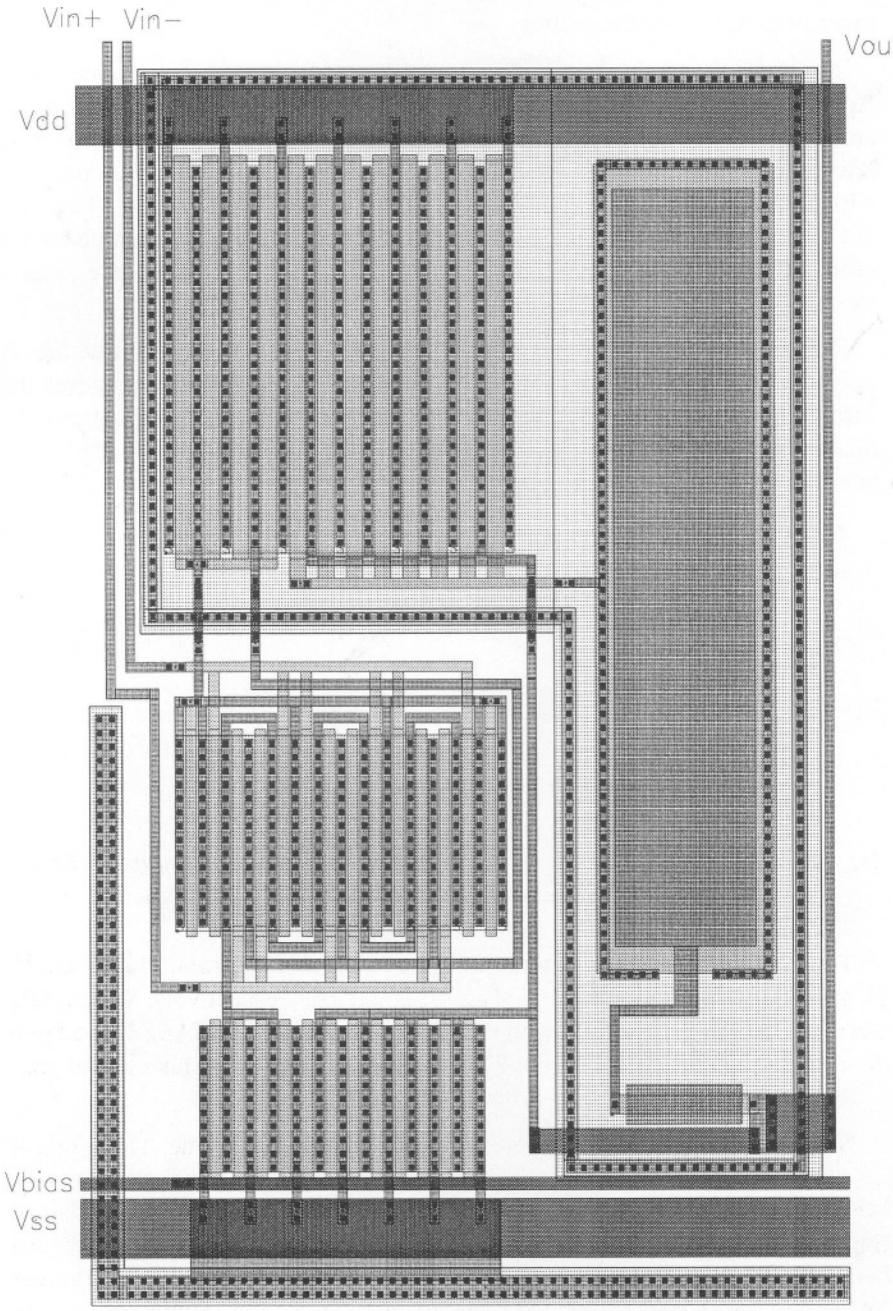


Fig. 5.66 - Layout of the two-stages OTA which schematic is shown in Fig. 5.65

low-power, medium-frequency circuits the matching between transistors is more important than the routing.

Fig. 5.66 shows the layout. Two metal layers favour the interconnections. Actually, metal-2 is used for V_{DD}, V_{SS} and V_{bias}. In addition, only few metal crossings use metal-2. Thus, metal-2 is largely available for the system level interconnections. Moreover, the use of metal-2 for V_{DD} enables the crossings with the vertical metal-1 lines leading out the inputs and the output. This lay-out organization favours the placement of various op-amps one side to the other with biasing running horizontally and external components (like capacitors and switches) placed on the top side.

The compensation network is on the right edge of the stacked structure. A poly1-poly2 arrangement obtains the capacitor while a poly strip achieves the zero nulling resistor. An n-well, biased all around the periphery, sits under the p-channel transistors and the compensation network. An *L*-shaped substrate bias closes the protection ring around the whole structure.

Fig. 5.67 shows the schematic of a folded cascode. The circuit is more

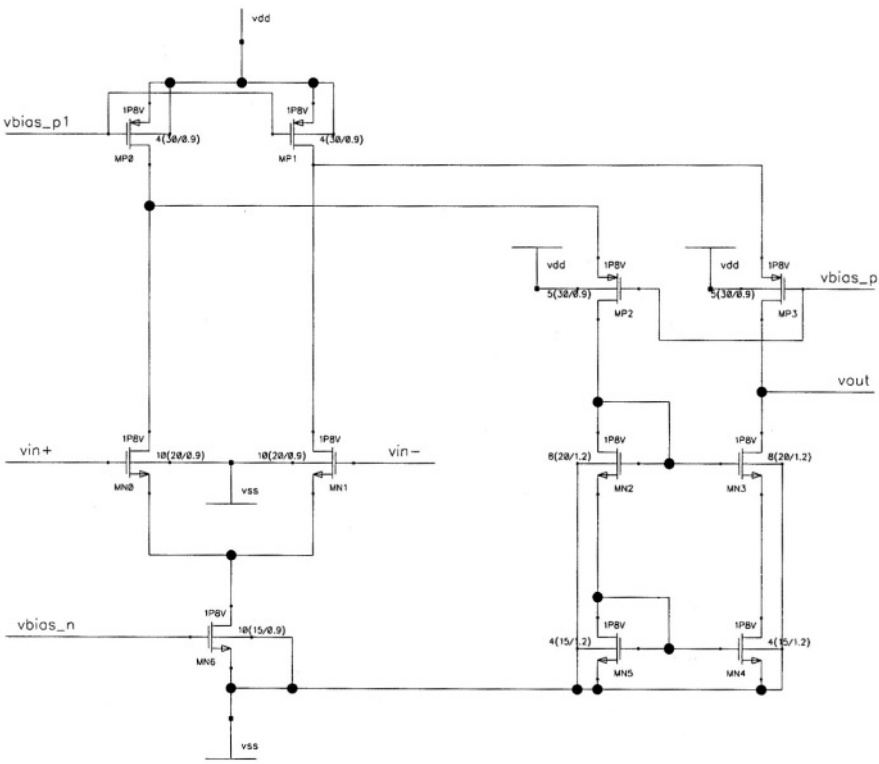


Fig. 5.67 - Circuit schematic of a a folded cascode. The transistor sizes are appropriate for a stacked layout. Use a shrink factor of 2 to have the real sizes.

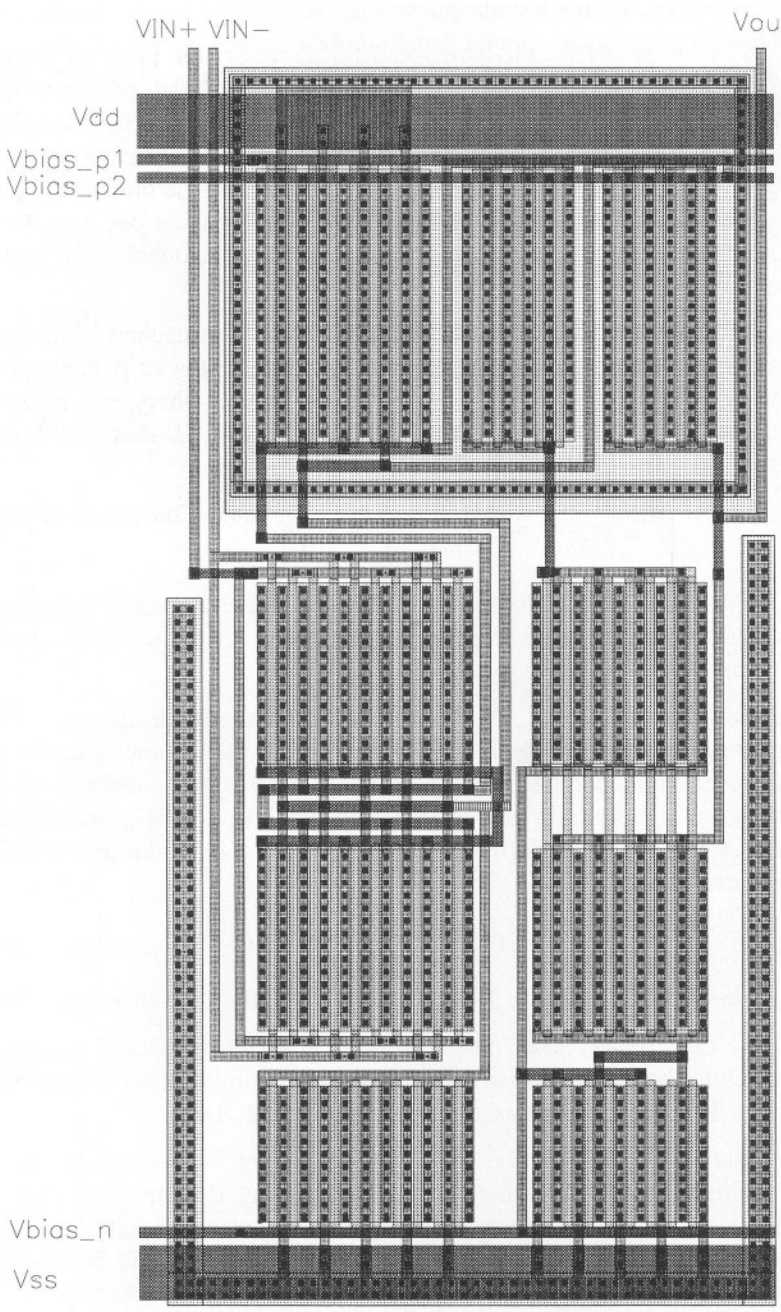


Fig. 5.68 - Layout of the folded cascode which circuit schematic is shown in Fig. 5.67

complex than the simple two-stages discussed above. However it is possible to identify a proper scheme for the stacked layout. The transistors of the input pair is made by 10 fingers. Such a large number advises the use of a common centroid layout. The below scheme shows a possible floor planning. It uses four rows of stacks divided in two vertical sections

CDDCCDDC FFFFF GGGGG
BAABBAABBA KKKKKKKK
ABBAABBAAB NNNNNNN
EEEEEEEEEE PPQQPPQQ

The used letters denotes transistor fingers corresponding to

$A \rightarrow M_{N0}$ $B \rightarrow M_{N1}$ $C \rightarrow M_{P0}$ $D \rightarrow M_{P1}$
$E \rightarrow M_{N6}$ $F \rightarrow M_{P2}$ $G \rightarrow M_{P3}$ $K \rightarrow M_{N2}$
$N \rightarrow M_{N3}$ $P \rightarrow M_{N5}$ $Q \rightarrow M_{N4}$

Fig. 5.68 shows the obtained layout. The rectangular shape and the vertical metals leading out inputs and output resembles, once again, the approach used for a standard cell layout.

The transistors M_{P2}, M_{P3}, M_{N2} and M_{N3} are non critical elements: they enhance the output resistance of the OTA or improve the symmetry of the circuit. However a possible inaccuracy of transistors' sizes don't affect much the performances. Therefore, in the layout M_{P2}, M_{P3}, M_{N2} and M_{N3} are separate elements. Actually, the width of M_{P0}, M_{P1}, M_{P2} and M_{P3} would permit a single stack arrangement. It is

FFFFFCCDDCCDDGGGGG

Instead, the sizes of M_{P2}, M_{P3}, M_{N4} and M_{N5} don't offer the same possibility.

Observe that the left stack on the top row is *CDDCCDDC*. It ensures a good matching between transistors but it would require dummy elements at the ending. This option is not used in the layout of Fig. 5.68.

The input pair has a common centroid arrangement. The use of two metal layers facilitates the required interconnection crossing. The input pair doesn't use dummy elements. The reason is that the common centroid yields an *A* and *B* element at both endings. Metal wiring joining the gates at the top and the bottom of the common centroid structure minimize the gate resistance.

A well bias surrounds the well where the *p-channel* transistors sit. Additionally, a *U* shaped substrate bias encloses the *n-channel* region.

5.14 REFERENCES

- D. A. Johns and K. Martin, *Analog Integrated Circuit Design,* J. Wiley & Sons, New York, NY, 1997.
- P. R. Gray and R, Meyer, *Analysis and Design of Analog Integrated Circuits,* J. Wiley & Sons, New York, NY, 1993.
- P.R. Gray and R.G. Meyer. *MOS operational amplifier design-a tutorial overview,* IEEE Journal of Solid-State Circuits (1982) SC-17.6 (Dec. 1982, pp. 969-982.
- B.K. Ahuja, *An improved frequency compensation technique for CMOS operational amplifiers,* IEEE Journal of Solid-State Circuits, Vol. 28, 1993, pp. 629-633.
- K. Gulati and Hae-Seung Lee, *A High-Swing CMOS Telescopic Operational Amplifier,* IEEE J. Solid-State Circuits, vol. 33, 1998, pp. 2010–2019.
- J. N. Babanezad, *A low-output-impedance fully differential op-amp with large output swing and continuous-time common-mode feedback,* IEEE J. Solid-State Circuits, vol. 26, 1991, pp. 1825-1833.
- K. Bult, G, Geelen, *A fast-settling CMOS op-amp for SC circuits with 90-dB dc Gain,* IEEE Journal of Solid-State Circuits, Vol. 25, 1990, pp. 1379-1384.
- E. Sackinger and W. Guggenbuhl, *A high-swing, high-impedance MOS cascode circuits,* IEEE Journal of Solid-State Circuits vol. 25, 1990, pp. 289-298.
- R. Castello, P. R. Gray, *A high performance Micropower Switched-capacitor Filter,* IEEE Journal of Solid-State Circuits, Vol. 20, 1985, pp. 14122-1132.
- M. Banu, J. M. Khoury, and Y. Tsividis, *Fully-differential Operational amplifiers with Accurate Output Balancing,* IEEE Journal of Solid-State Circuits, Vol. 23, 1988, pp. 1410-1414.
- S. Mallya and J.H. Nevi, *Design procedures for a fully differential cascode CMOS operational amplifier,* IEEE Journal of Solid-State Circuits vol., 24, 1989, pp.1737-1740.
- L.G.A. Callewaert and W.M Sansen, *Class AB CMOS amplifiers with high efficiency,* IEEE Journal of Solid-State Circuits, vol. 25, 1990, pp. 684-691.
- R. Kline, B. J. Hosticka, and H. J. Pfleiderer, *A very-high slew rate CMOS operational amplifier,* IEEE J. Solid-State Circuits, vol. 24, 1989, pp. 744–746.
- J. F. Duque-Carrillo, *Control of the Common-mode Components in CMOS Continuous-time Fully-differential Signal Processing,* International Journal of Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers, 1993
- M. Degrauwe, J. Rijmenants, E. A. Vittoz, and D. Man, *Adaptive biasing*

CMOS amplifiers, IEEE J. Solid-State Circuits, vol. 17, 1982, pp. 522-528.

- R. Harjani, R. Heineke and Feng Wang, *An integrated low-voltage class AB CMOS OTA,* IEEE Journal of Solid-State Circuits, vol.34, 1999, pp. 134-142.

5.15 PROBLEMS

5.1 Calculate the response of the circuit in Fig. 5.2 assuming that Z_1 and Z_2 are capacitors, C_1 and C_2 respectively, and $Z_3 = Z_4 = 0$. Assume finite the gain and the bandwidth of the op-amp.

5.2 Simulate, using Spice, the time response of the circuit in Fig. 5.3. Use the following parameters $C_1 = 0.5\ pF$; $C=2\ pF$; $C_0 = 3\ pF$; $g_m = 1\ mA/V$; $r_0 = 1\ M\Omega$. Compare the result with the case: $C_1 = 0$ and a charge injection by $1\ pCoul$ at the time $t = 0$.

5.3 Using Spice, determine the spur signal produced on the positive supply bias of the op-amp of Fig. 5.6. The bonding pad causes $8\ \mu H$. The resistance on the analog wire is $1\ \Omega$ and the current I_A is $2\ mA$. The current I_D is made by bipolar glitches with $1\ \mu sec$ periodicity and amplitude $20\ mA$. The duration of the glitches is $40\ psec$.

5.4 Simulate, using Spice and the models of Appendix B the circuit in Fig. 5.12. Use the following design parameters: $(W/L)_1 = (W/L)_6 = 100/1.5$; $(W/L)_5 = 200/2$. All the sizes are in μ. $C_c = 2pF$. Determine the differential gain, the common mode gain, the offset and the power supply rejections.

5.5 Repeat Problem 5.4 using the models of Appendix C. Shrink the geometry of all the transistors by a factor 2. Use a suitable resistance in series with C_c and estimate the value that permits the circuit to ensure a $60°$ phase margin with a capacitive load of $2pF$.

5.6 Calculate the systematic offset of the two stages amplifier described in Problem 5.4 for the following values of $(W/L)_5$: $190/2$; $210/2$; $100:1$. Moreover, perform a simulation where two equal transistors connected in parallel $(W/L = 50/1.5)$ replace M_6 and four parallel equal transistors $(W/L = 50/2)$ replace M_5. Compare the obtained offset with the initial design and comment the result. All the sizes are in μ.

5.7 Determine the random offset produced by a 5% mismatch in the W/L ratio of M_1-M_2 and M_3-M_4 of the two stages amplifier discussed in

Problem 5.5 (shrink geometers). Compare the obtained random offset with the one that comes out from a 5% inaccuracy of the *W/L* of M_6.

5.8 Design a two stages *OTA* with an *n-channel* input pair able to fulfil the following specifications: *dc* gain *80 dB;* $f_T = 50MHz$; phase margin *60°*; $C_L = 3pF$. Use the transistor models of Appendix C.

5.9 Determine the power supply rejection as a function of the frequency of the *OTA* designed in Problem 5.8. Estimate the effect of a capacitive coupling between the supply voltages and the inputs. The coupling results from a capacitor *C* joining inputs and V_{DD} and another capacitor *2C* linking inputs and ground. Study the effect of a spur voltage superposed to V_{DD} or to ground and the effect of a noisy current affecting I_{ref}.

5.10 Simulate the effect of the parasitic coupling in the circuit of Fig. 5.19. Assume $C_{p1+} = C_{p2+} = 20fF$; $C_{p1-} = C_{p2-} = 40fF$; $C_1 = 1pF$; $C_2 = 2pF$. Use for the op-amp the small signal equivalent circuit of Fig. 5.20. The gain of each stage is *40 dB*. Moreover, $C_1 = 100fF$; $C_2 = 2pF$, $C_c = 3pF$ and $g_{m1} = g_{m2} = 1mA/V$.

5.11 Design a two stages OTA using the following design conditions: Input pair *(W/L)=200/1;* tail current *300* μA; current in the second stage *600* μA; $C_L = 5pF$. Use the Spice models of Appendix B and $V_{DD} = 3.3V$. Determine the transistor sizes that lead to $A_1 = 35$ *dB*; $A_2 = 40dB$. Design the zero nulling compensation network that provides a phase margin better than 60°.

5.12 Design a compensation network with unity gain buffer to be used with for the *OTA* designed in Problem 5.11. Identify the possible zero-pole doublet that comes out because of the limited transconductance of the buffer. Use for the buffer a current lower than 50 μA.

5.13 With Spice simulations, determine the equivalent resistance of the transistor pair M_n–M_p of Fig. 5.24. $(W/L)_n = 10/1$; $(W/L)_p = 10/1$; $V_{DD} = +1.65\ V$; $V_{SS} = -1.65V$; $V_1 = 0$. Plot the result for V_1 ranging from *+1 V* to *-IV*. Compare results given by the transistor models of Appendix B and C. All the sizes are in μ.

5.14 Repeat Example 5.1 but use a supply voltage of *3.3 V*. The key design target is to limit the power consumption below *0.2* μW. The input pair should operate at the limit of the saturation/sub-threshold regions. Determine gain and bandwidth ($C_L = 2pF$, phase margin *40°*).

5.15 Repeat Example 5.2. Use the design conditions given in Problem 5.14.

and employ real current generators.

5.16 Given the *OTA* designed in Example 5.1, determine with Spice simulations, the positive and negative slew-rate. Connect the op-amp in the unity gain arrangement and utilize an input step signal jumping between *0.5* and *1.3V* (and vice-versa).

5.17 Find the small-signal voltage gain of the telescopic cascode of Fig. 5.27. Account for the substrate transconductance of transistors M_3 and M_4 Determine the gain from the gate of M_3-M_4 to the output.

5.18 Repeat Example 5.4 but assume that the capacitive load is 5 *pF*. What is the bias current that permits to achieve the same gain bandwidth product of Example 5.4. What is the *dc* gain?

5.19 Using the Spice simulation results of Example 5.4 find the position of the non-dominant poles. Estimate the effect of the source-substrate and drain-substrate parasitic capacitances on the location of non-dominant poles.

5.20 Consider the mirrored cascode of Fig 5.29. The transistor sizing is the following: $(W/L)_1 = (W/L)_2 = 150/0.4$; $(W/L)_{10} = (W/L)_{11} = (W/L)_{12} = (W/L)_{13} = 50/0.8$; $(W/L)_7 = (W/L)_8 = (50/0.6)$; $(W/L)_9 = 25/1$; $V_{DD} = 3.3V$. The current in M_g is $250\ \mu A$. Use the transition models of Appendix C. Find the value of V_{B1} and suitably modify the *n-channel* current mirror in order to obtain the widest and symmetrical output swing. Use for all the cascading transistors *(W/L)= 200/0.6.* All the sizes are in μ.

5.21 Use the simulation results of Problem 5.20 to estimate the location of the non-dominant poles. Find the load capacitance that establishes a phase margin of *60°* and calculate the ratio between the non-dominant poles and the unity gain frequency.

5.22 Repeat Example 5.6 but use a scaled version of transistors. Use a scaling factor :*2* and the Spice models of Appendix C. The supply voltage is *2.2 V,* and the capacitive load is *2 pF.*

5.23 Consider the folded cascode of Example 5.6, where the transistors are shrink by a factor :*2.5* and the supply voltage is *2.2 V.* The reference currents are reduced by the factor *1.5.* Use the local feedback technique to boost the *dc* gain. For the two amplifiers required for the gain boosting use a differential single gain stage.

5.24 Repeat Example 5.7 and try to increase the dc gain to *120 dB.* Any transistor sizes and current level are accepted.

5.25 Consider the bias network of Fig 5.34. Use the following design conditions: $(W/L)_{7b} = 50/0.6$; $(W/L)_{2b} = 200/0.4$; $(W/L)_{4b} = 50/0.8$; $I_{ref} = 200\ \mu A$; $I_{M7b} = 200\ \mu A$; $V_{DD} = 3.3V$. Design the rest of the circuit using reasonable criteria. Find sensitivity of the current I_{M6b} for a $\pm 10\%$ variation of V_{DD}. All the sizes are in μ. Use the models of Appendix C.

5.26 Design the single stage *AB* class op-amp of Fig. 5.36. The required specifications are: $A_0 = 60\ dB$; $I_{tot} = 0.1\ \mu A$. Estimate the slew rate of the output for *4 pF* load and open loop conditions Use the technology described in Appendix C.

5.27 Estimate, with Spice simulations and $(W/L)_1 = (W/L)_2 = 50/2$; $(W/L)_3 = 20/0/6$; $I_{M3} = 100\ \mu A$, the range of operation of the common-mode feedback of Fig 5.44. Find the possible range extension achieved by two degeneration resistances of $2\ K\Omega$ connected in series with the source of M_1 and M_2. Use the technology of Appendix C.

5.28 Design and simulate the current mirror with adjustable mirror factor of Fig. 5.45. $V_{S3R} = V_{S2} = 0.1\ V$. The common mode output signal is *2.5 V*. The nominal current is *0.2 mA* and must increase to *0.3 mA* when the control voltages rises to *3 V*. Plot the variation of I_{out} versus the amplitude of the differential control.

5.29 Sketch the layout of the two stages op-amp of Fig. 5.65 but with input stage made by *8* $(20\mu/1.2\mu)$ fingers and the active load by *4* $(20\mu/1.8\mu)$ fingers. It is recommended a common centroid layout for the input stage and an interdigitized active loads. The currents in the wire connections are significant. Find a proper wiring (using two metal layers) that matches the parasitic drop voltages.

5.30 Sketch the layout of the folded cascode of Fig. 5.67. The width of the fingers of M_{N4}-M_{N5} is 20μ. Transistors M_{N2}, M_{N5} and M_{N6} are made by *5, 4, 18* fingers respectively. Find the arrangement that uses *4* stacks of transistors. They must include M_{p0}-M_{p1}-M_{p2}-M_{p3}; M_{N0}-M_{N1}; M_{N6}; and M_{N2}-M_{N3}-M_{N4}-M_{N5} respectively.

5.31 Sketch the layout of the two stages amplifier designed in Example 5.1. Use a common centroid layout for the input pair. Divide the transistors in fa suitable number of fingers and propose possible modifications of the sizes in order to achieve a more efficient layout.

5.32 Layout the telescopic cascode of Example 5.4. Change the transistor sizes or the number of finger to have the p-channel transistors in the same stack. The current tail generator and its control must have an

interdigitized layout. Try to obtain a rectangular shape and ensure that the inputs and output leads are on the top side of the rectangle.

5.33 Identify two stacked arrangement for the mirrored cascode of Example 5.5. The first solution accommodates the transistors in 4 stacks, two for the p-channel transistors and two for the n-channel. The second arrangement uses three stacks, one for the input pair, one for the p-channel transistors and the third one for the remaining n-channel elements.