CMOS Mixed Signal Design

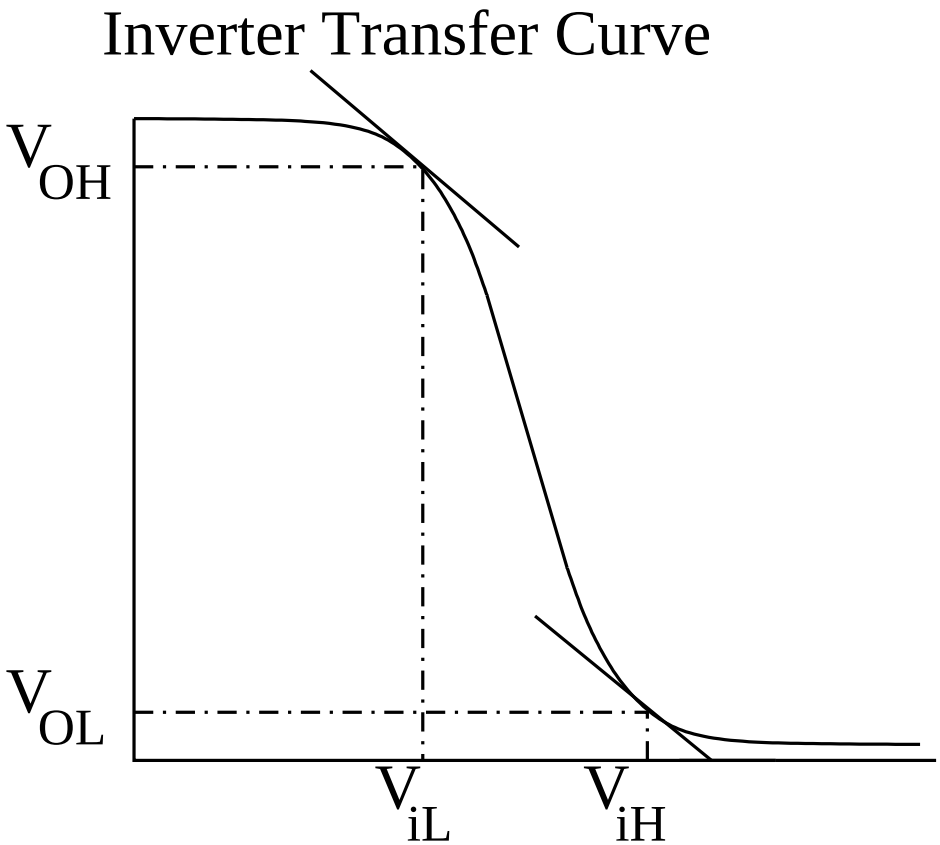
Part I: OpAmp Design

Dinesh Sharma

Microelectronics Group, EE Department
IIT Bombay, Mumbai

September 19, 2010

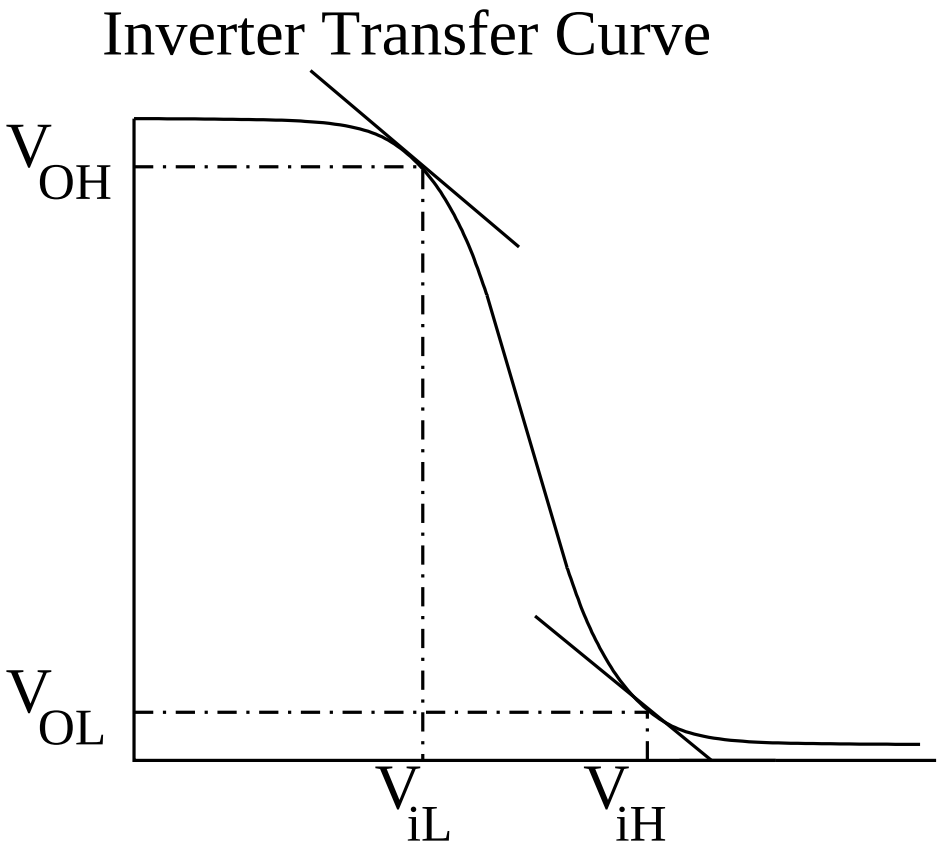
Linear Mode of Operation



- Analog circuits require the output voltage to be **sensitive** to the input voltage.
- Digital logic requires the output to be **insensitive** to the exact input voltage.

Circuits need to be *biased* for operation in the linear regime.

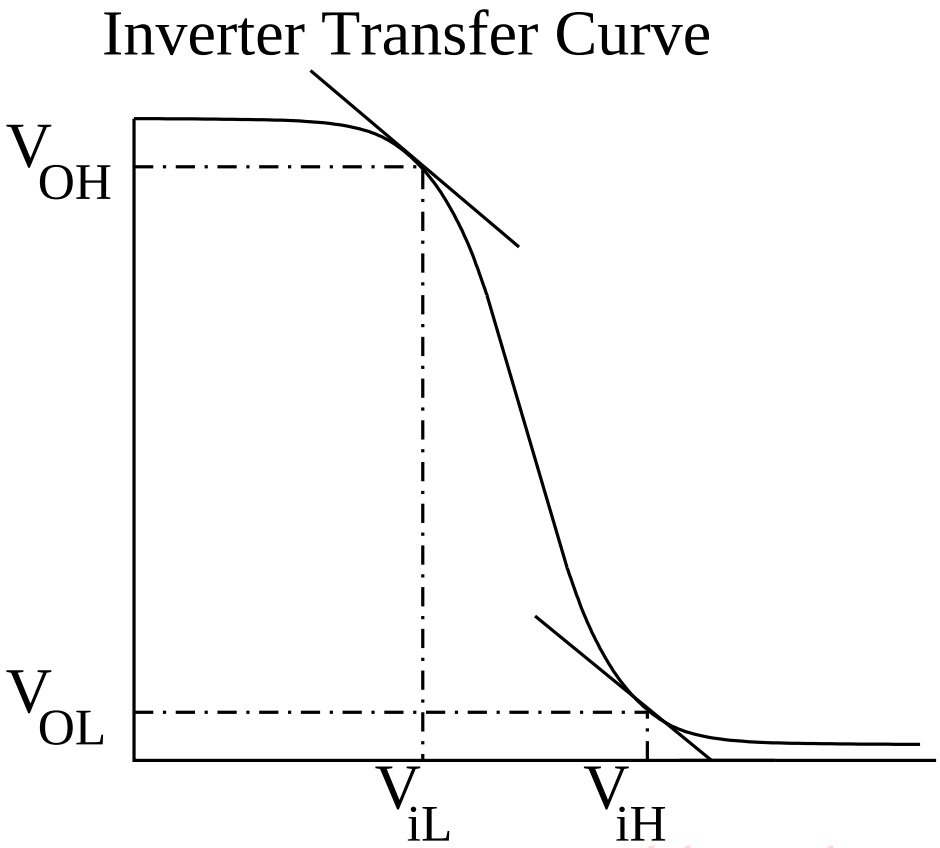
Linear Mode of Operation



- Analog circuits require the output voltage to be **sensitive** to the input voltage.
- Digital logic requires the output to be **insensitive** to the exact input voltage.

Circuits need to be *biased* for operation in the linear regime.

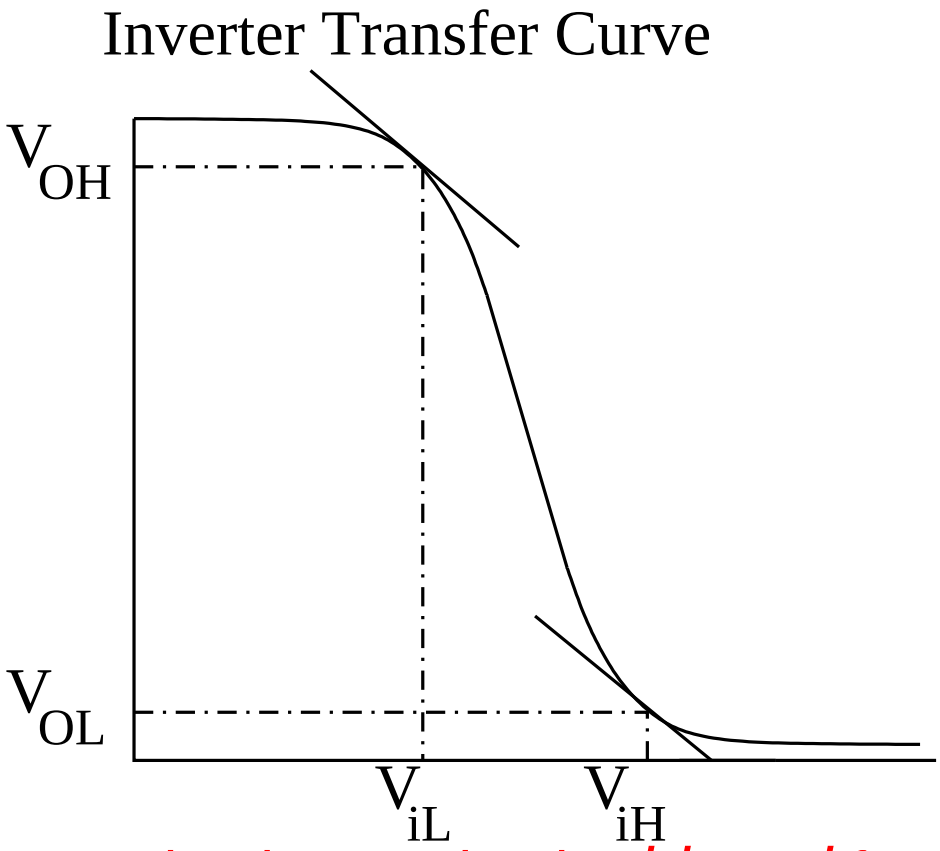
Linear Mode of Operation



- Analog circuits require the output voltage to be **sensitive** to the input voltage.
- Digital logic requires the output to be **insensitive** to the exact input voltage.

Circuits need to be *biased* for operation in the linear regime.

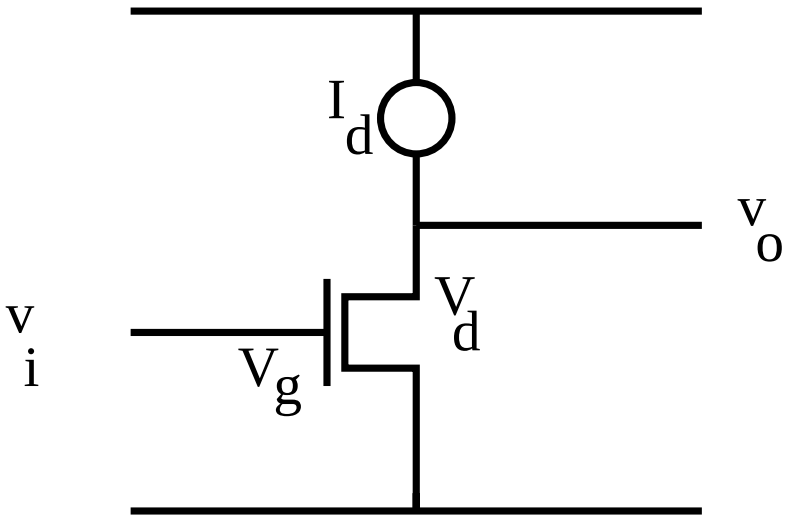
Linear Mode of Operation



- Analog circuits require the output voltage to be **sensitive** to the input voltage.
- Digital logic requires the output to be **insensitive** to the exact input voltage.

Circuits need to be *biased* for operation in the linear regime.

A Single Transistor Amplifier



$$dI_d = \frac{\partial I_d}{\partial V_g} dV_g + \frac{\partial I_d}{\partial V_d} dV_d$$

$$\frac{\partial I_d}{\partial V_g} = g_m \text{ (Transconductance)}$$

$$\frac{\partial I_d}{\partial V_d} = g_o \text{ (O/P conductance)}$$

The current source load keeps the drain current constant. So

$$dI_d = 0 = g_m v_i + g_o v_o$$

Hence, the voltage gain (A_o) is

$$A_o = \frac{v_o}{v_i} = -\frac{g_m}{g_o} = -g_m r_o$$

Transistor Characteristics

g_m and g_o depend on the transistor characteristics.
In saturation,

$$I_d \simeq \frac{K}{2}(V_{gs} - V_T)^2$$

where, K is the conductivity factor given by:

$$K = K' \left(\frac{W}{L}\right) \equiv \mu C_{ox} \left(\frac{W}{L}\right)$$

V_T is the threshold voltage
W and L are transistor width and length respectively.
μ is the mobility
and $C_{ox}$is the gate oxide capacitance per unit area.

Transconductance

$$\text{Let } V_{GT} \equiv (V_{gs} - V_T)$$

$$\text{Then } I_d = \frac{KV_{GT}^2}{2} \qquad \text{and} \qquad V_{GT} = \sqrt{\frac{2I_d}{K}}$$

$$g_m = \frac{\partial I_d}{\partial V_g} = KV_{GT} = K'\left(\frac{W}{L}\right)V_{GT}$$

$$\text{Also } g_m = KV_{GT} = K\sqrt{\frac{2I_d}{K}} = \sqrt{2KI_d} = \sqrt{2K'\left(\frac{W}{L}\right)I_d}$$

$$\text{Similarly, } K = \frac{2I_d}{{V_{GT}}^2}\text{; Therefore } g_m = \frac{2I_d}{{V_{GT}}^2}V_{GT} = \frac{2I_d}{V_{GT}}$$

Which formula?

$$g_m = K' \left(\frac{W}{L}\right) V_{GT}$$

$$g_m = \sqrt{2K' \left(\frac{W}{L}\right) I_d}$$

$$g_m = \frac{2I_d}{V_{GT}}$$

To increase g_m
should we increase V_{GT}?
or decrease it?
Is g_m linearly dependent on transistor size?
dependent on its square root?
or is it independent of transistor size?

In fact, which formula should be applied depends on how the transistor is biased and sized. If size and VGT are known, the first formula applies. If the drain current and size are known, the second one does. If gate voltage and drain current are given and the transistor is accordingly sized, the third formula should be used.

Which formula?

$$g_m = K' \left(\frac{W}{L}\right) V_{GT}$$

$$g_m = \sqrt{2K' \left(\frac{W}{L}\right) I_d}$$

$$g_m = \frac{2I_d}{V_{GT}}$$

To increase g_m
should we increase V_{GT}?
or decrease it?
Is g_m linearly dependent on transistor size?
dependent on its square root?
or is it independent of transistor size?

In fact, which formula should be applied depends on how the transistor is biased and sized. If size and VGT are known, the first formula applies. If the drain current and size are known, the second one does. If gate voltage and drain current are given and the transistor is accordingly sized, the third formula should be used.

Which formula?

$$g_m = K' \left(\frac{W}{L}\right) V_{GT}$$

$$g_m = \sqrt{2K' \left(\frac{W}{L}\right) I_d}$$

$$g_m = \frac{2I_d}{V_{GT}}$$

To increase g_m
should we increase V_{GT}?
or decrease it?
Is g_m linearly dependent on transistor size?
dependent on its square root?
or is it independent of transistor size?

In fact, which formula should be applied depends on how the transistor is biased and sized. If size and VGT are known, the first formula applies. If the drain current and size are known, the second one does. If gate voltage and drain current are given and the transistor is accordingly sized, the third formula should be used.

Which formula?

$$g_m = K' \left(\frac{W}{L}\right) V_{GT}$$

$$g_m = \sqrt{2K' \left(\frac{W}{L}\right) I_d}$$

$$g_m = \frac{2I_d}{V_{GT}}$$

To increase g_m
should we increase V_{GT}?
or decrease it?
Is g_m linearly dependent on transistor size?
dependent on its square root?
or is it independent of transistor size?

In fact, which formula should be applied depends on how the transistor is biased and sized. If size and VGT are known, the first formula applies. If the drain current and size are known, the second one does. If gate voltage and drain current are given and the transistor is accordingly sized, the third formula should be used.

Which formula?

$$g_m = K' \left(\frac{W}{L}\right) V_{GT}$$

$$g_m = \sqrt{2K' \left(\frac{W}{L}\right) I_d}$$

$$g_m = \frac{2I_d}{V_{GT}}$$

To increase g_m
should we increase V_{GT}?
or decrease it?

Is g_m linearly dependent on transistor size?

dependent on its square root?
or is it independent of transistor size?

In fact, which formula should be applied depends on how the transistor is biased and sized. If size and VGT are known, the first formula applies. If the drain current and size are known, the second one does. If gate voltage and drain current are given and the transistor is accordingly sized, the third formula should be used.

Which formula?

$$g_m = K' \left(\frac{W}{L}\right) V_{GT}$$

$$g_m = \sqrt{2K' \left(\frac{W}{L}\right) I_d}$$

$$g_m = \frac{2I_d}{V_{GT}}$$

To increase g_m
should we increase V_{GT}?
or decrease it?
Is g_m linearly dependent on transistor size?
dependent on its square root?
or is it independent of transistor size?

In fact, which formula should be applied depends on how the transistor is biased and sized. If size and VGT are known, the first formula applies. If the drain current and size are known, the second one does. If gate voltage and drain current are given and the transistor is accordingly sized, the third formula should be used.

Which formula?

$$g_m = K' \left(\frac{W}{L}\right) V_{GT}$$

$$g_m = \sqrt{2K' \left(\frac{W}{L}\right) I_d}$$

$$g_m = \frac{2I_d}{V_{GT}}$$

To increase g_m
should we increase V_{GT}?
or decrease it?
Is g_m linearly dependent on transistor size?
dependent on its square root?
or is it independent of transistor size?

In fact, which formula should be applied depends on how the transistor is biased and sized. If size and VGT are known, the first formula applies. If the drain current and size are known, the second one does. If gate voltage and drain current are given and the transistor is accordingly sized, the third formula should be used.

Which formula?

$$g_m = K' \left(\frac{W}{L}\right) V_{GT}$$

$$g_m = \sqrt{2K' \left(\frac{W}{L}\right) I_d}$$

$$g_m = \frac{2I_d}{V_{GT}}$$

To increase g_m
should we increase V_{GT}?
or decrease it?
Is g_m linearly dependent on transistor size?
dependent on its square root?
or is it independent of transistor size?

In fact, which formula should be applied depends on how the transistor is biased and sized. If size and VGT are known, the first formula applies. If the drain current and size are known, the second one does. If gate voltage and drain current are given and the transistor is accordingly sized, the third formula should be used.

Output conductance

Assuming a simple Early effect like model, we can write for g_o:

$$g_o \simeq \lambda' I_d / L$$

where L is the channel length and λ is a technology dependent parameter. In terms of geometry and V_{GT}, we can write:

$$g_o = \frac{\lambda' K'}{2} \frac{W}{L^2} V_{GT}^2$$

The Early Voltage V_A is L/λ'. So,

$$g_o \simeq I_d / V_A = \frac{K' W}{2\lambda'} \left(\frac{V_{GT}}{V_A} \right)^2$$

Voltage Gain

The voltage gain in terms of geometry and V_{GT}:

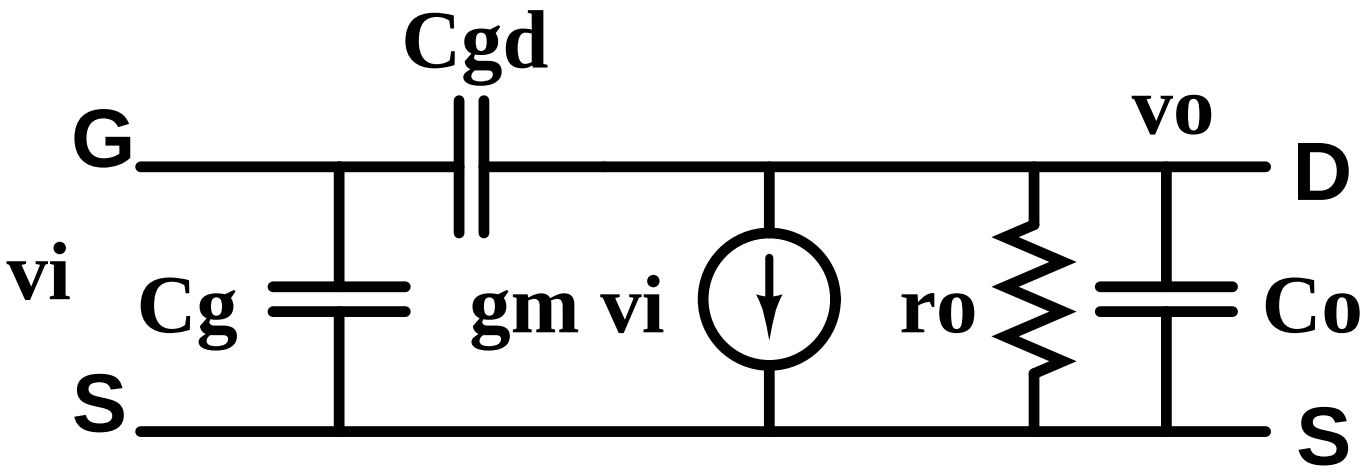
$$A_o = \frac{2L}{\lambda' V_{GT}}$$

In terms of drain current and geometry:

$$A_o = \frac{1}{\lambda'} \sqrt{\frac{2K'WL}{I_d}}$$

Thus, if the transistor is biased at constant current, the DC gain is determined by the square root of the gate *area*.

AC Behaviour



$$sC_{gd}(v_i - v_o) - g_m v_i - \frac{v_o}{r_o} - sC_o v_o = 0$$

$$v_i\left(sC_{gd} - g_m\right) - v_o\left(sC_{gd} + \frac{1}{r_o} + sC_o\right) = 0$$

So the AC gain $A_1 = \frac{v_o}{v_i} = -g_m r_o \frac{1 - sC_{gd}/g_m}{1 + sr_o(c_{gd} + c_o)}$

Bandwidth

$$A_1 = -g_m r_o \frac{1 - sC_{gd}/g_m}{1 + sr_o(c_{gd} + c_o)}$$

$$\text{Let } C_{tot} \equiv C_{gd} + C_o$$

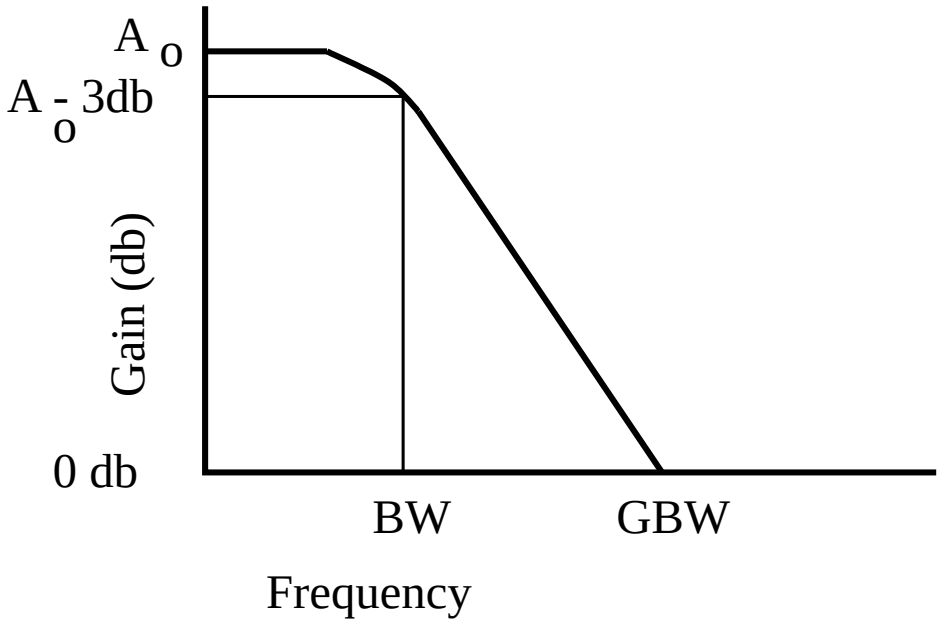
$$\text{Then, } A_1 = A_o \frac{1 - sC_{gd}/g_m}{1 + sr_o C_{tot}}$$

Normally, $\omega C_{gd}/g_m << 1$

$$\text{Therefore, } A_1 \simeq \frac{A_o}{1 + sr_o C_{tot}}$$

This describes the frequency response of a system with one dominant pole. The bandwidth is given by $1/r_o C_{tot}$.

Gain Bandwidth Product



$$GBW = g_m r_o \cdot \frac{1}{r_o C_{tot}} = \frac{g_m}{C_{tot}}$$

The gain bandwidth product (or the cutoff frequency) is independent of r_o.

Maximum GBW

GBW is max. when there is no load connected and the load is entirely due to the device capacitance itself. Then the load capacitance is proportional to the device width.

$C_{tot} = \chi W$ where χ is a technological parameter.

$$GBW_{max} = \frac{g_m}{\chi W}$$

$$\begin{aligned} GBW_{max} &= \frac{K' V_{GT}}{\chi L} \\ &= \frac{1}{\chi}\sqrt{\frac{2K' I_d}{WL}} \\ &= \frac{2I_d}{\chi W V_{GT}} \end{aligned}$$

Summary

Parameters	Free Design Variables:		
	W, L, V_{GT}	W, L, I_d	L, V_{GT}, I_d
g_m	$K' \frac{W}{L} V_{GT}$	$\sqrt{2K' \frac{W}{L} I_d}$	$\frac{2I_d}{V_{GT}}$
g_o	$\frac{\lambda' K' W V_{GT}^2}{2L^2}$	$\frac{\lambda' I_d}{L}$	$\frac{\lambda' I_d}{L}$
A_o	$\frac{2L}{\lambda' V_{GT}}$	$\frac{1}{\lambda'} \sqrt{\frac{2K' WL}{I_d}}$	$\frac{2L}{\lambda' V_{GT}}$
GBW	$\frac{K' W V_{GT}}{L C_{tot}}$	$\sqrt{\frac{2K' W I_d}{L}} \frac{1}{C_{tot}}$	$\frac{2I_d}{V_{GT} C_{tot}}$
GBW_{max}	$\frac{K' V_{GT}}{\chi L}$	$\frac{1}{\chi} \sqrt{\frac{2K' I_d}{WL}}$	$\frac{K' V_{GT}}{\chi L}$

Technological Constraint

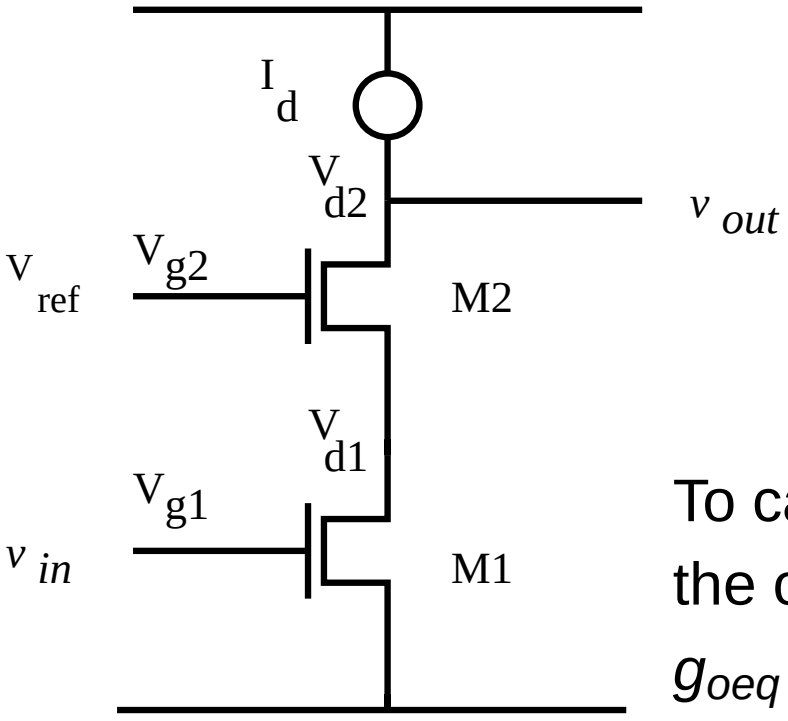
$$A_o \cdot GBW_{max} = \frac{2L}{\lambda' V_{GT}} \cdot \frac{K' V_{GT}}{\chi L} = \frac{1}{\lambda'} \sqrt{\frac{2K'WL}{I_d}} \cdot \frac{1}{\chi} \sqrt{\frac{2K' I_d}{WL}}$$

$$\text{So } A_o \cdot GBW_{max} = \frac{2K'}{\lambda' \chi}$$

Therefore, this quantity is a technological constant and the designer has no control over it.

What if an application requires a Gain-GBW product higher than this value?

Cascode Amplifier



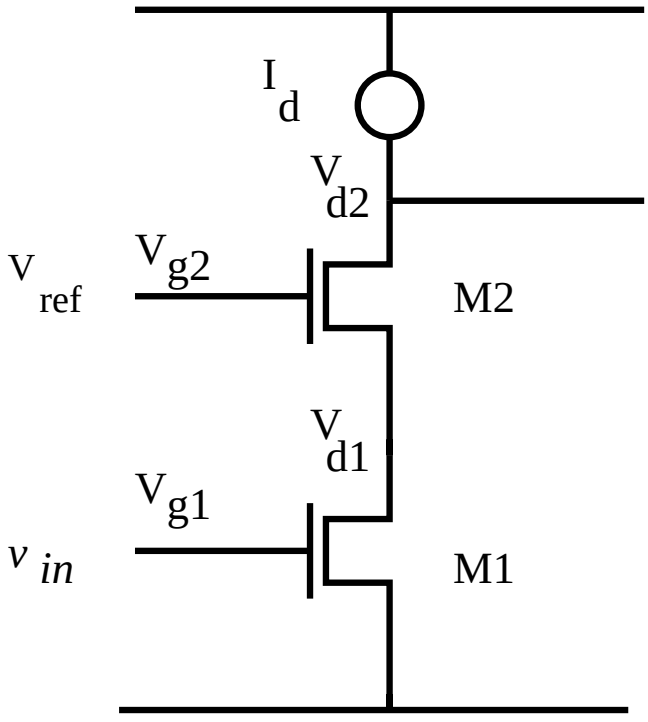
$$dI_d = g_{meq} dV_{g1} + g_{oeq} dV_{d2}$$

$$\text{So } g_{meq} = \frac{\partial I_d}{\partial V_{g1}} \text{ with } dV_{d2} = 0$$

$$\text{and } g_{oeq} = \frac{\partial I_d}{\partial V_{d2}} \text{ with } dV_{g1} = 0$$

To calculate g_{meq}, we put a voltage source at the output node and calculate $\frac{\partial I_d}{\partial V_{g1}}$.
$g_{oeq}$is calculated by putting a voltage source at vg1 and calculating $\frac{\partial I_d}{\partial V_{d2}}$.

Equivalent gm of Cascode



$$g_{meq} = \frac{\partial I_d}{\partial V_{g1}} \quad \text{with } dV_{d2} = 0$$

$$dV_{ds2} = -dV_{d1}, \quad dV_{gs2} = -dV_{d1}$$

v_{out}

$$i_d = g_{m1}v_{g1} + g_{o1}v_{d1}$$

$$i_d = -g_{m2}v_{d1} - g_{o2}v_{d1}$$

$$\text{So } v_{d1} = -\frac{id}{g_{m2} + g_{o2}}$$

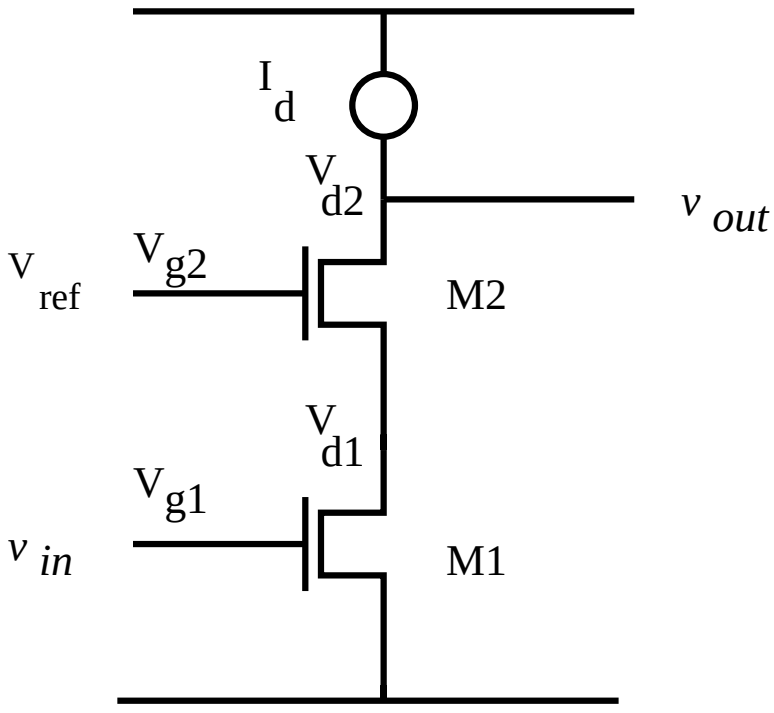
$$i_d = g_{m1}v_{g1} - i_d \frac{g_{o1}}{g_{m2} + g_{o2}}$$

$$g_{meq} = \frac{i_d}{v_{g1}} = g_{m1}\frac{g_{m2} + g_{o2}}{g_{o1} + g_{o2} + g_{m2}} \simeq g_{m1}$$

Equivalent go of Cascode

$$g_{oeq} = \frac{\partial I_d}{\partial V_{d2}} \qquad \text{with } dV_{g1} = 0$$

$$dV_{gs1} = 0, \qquad dV_{gs2} = -dV_{d1}, \qquad dV_{ds2} = dV_{d2} - dV_{d1}$$



$$i_d = 0 + g_{o1} v_{d1}, \qquad \text{so} v_{d1} = \frac{i_d}{g_{o1}}$$

$$i_d = -g_{m2} v_{d1} + g_{o2}(v_{d2} - v_{d1})$$

$$i_d = -i_d \frac{g_{m2} + g_{o2}}{g_{o1}} + g_{o2} v_{d2}$$

$$g_{oeq} = \frac{i_d}{v_{d2}} = \frac{g_{o1} g_{o2}}{g_{o1} + g_{o2} + g_{m2}}$$

$$g_{oeq} \simeq g_{o1} \frac{g_{o2}}{g_{m2}}$$

DC gain of Cascode

$$A_o = -\frac{g_{meq}}{g_{oeq}} = -\frac{g_{m1}(g_{m2}+g_{o2})}{g_{01}+g_{02}+g_{m2}} \cdot \frac{g_{01}+g_{02}+g_{m2}}{g_{01}g_{02}}$$

So $$A_o = -\frac{g_{m1}(g_{m2}+g_{o2})}{g_{01}g_{02}} = -\frac{g_{m1}}{g_{01}} \cdot \left(1+\frac{g_{m2}}{g_{02}}\right)$$

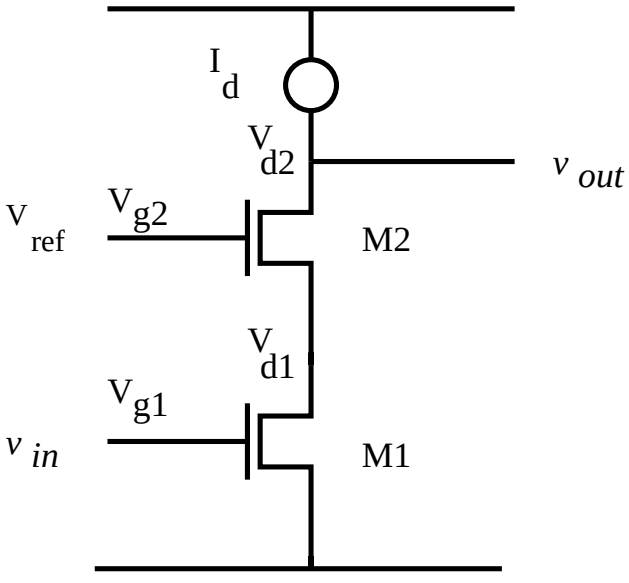
Let $A_{01} \equiv -\frac{g_{m1}}{g_{01}}$ common source gain

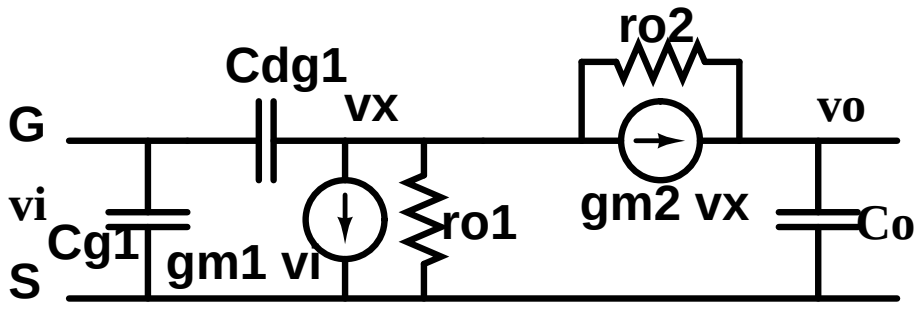
And $A_{02} \equiv 1+\frac{g_{m2}}{g_{02}}$ common gate gain

Then, $A_o = -A_{01} \cdot A_{02}$

DC gain = the product of the DC gain of the two transistors.

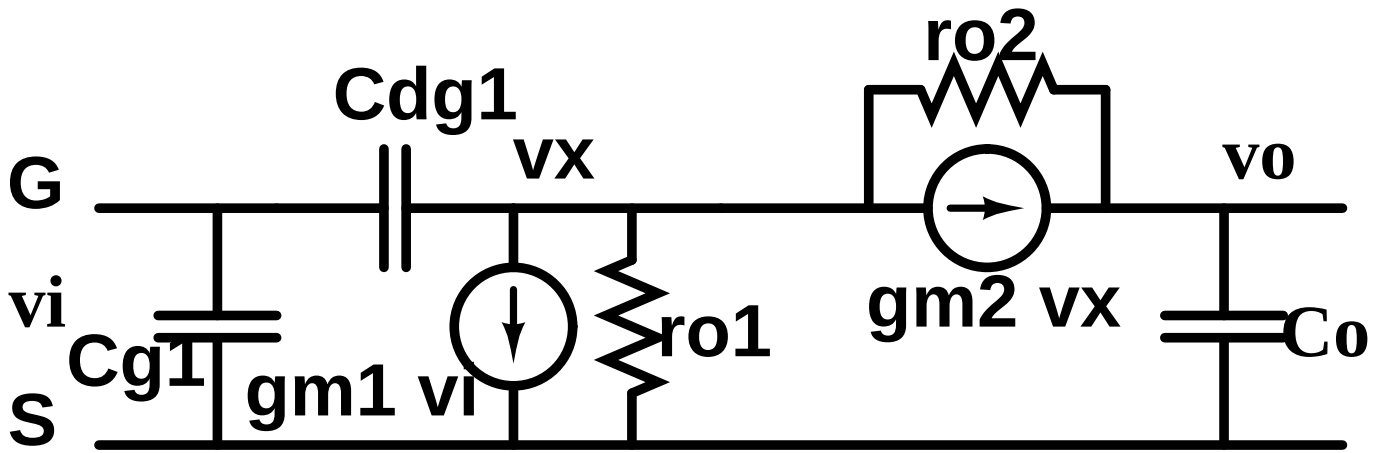
AC Behaviour of Cascode





We shall see presently that v_x is quite small.

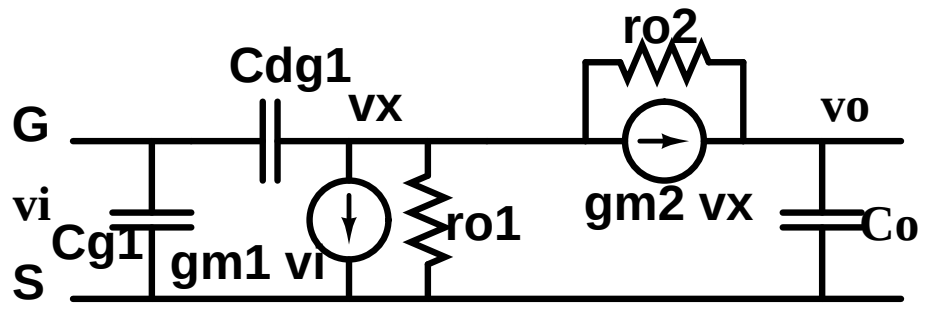
Initially, we shall ignore the effect of the drain capacitance of the lower transistor and the gate capacitance of the upper one. If necessary, we can always replace r_{o1} by $r_{o1} \| C_{ds1} \| C_{g2}$.



$$g_{m2}v_x + \frac{v_x - v_o}{r_{o2}} = sC_o v_o$$

$$v_x = \frac{1 + sr_{o2}C_o}{1 + g_{m2}r_{o2}} v_o = \frac{1 + sr_{o2}C_o}{A_2} v_o$$

Since A_2 is quite large, v_x is very small compared to v_o.



$$sC_{dg1}(v_i - v_x) = g_{m1}v_x + \frac{v_x}{r_{o1}} + sC_o v_o$$

$$\frac{v_o}{v_i} = -\frac{(A_1 - sr_{o1}C_{dg})A_2}{(1 + sr_{o2}C_o)(1 + sr_{o1}C_{dg}) + A_2 sC_o r_{o1}}$$

If $sr_{o1}C_{dg}$ is small,

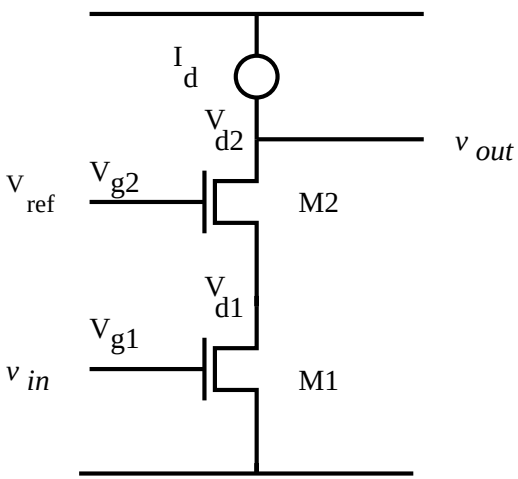
$$\text{Voltage gain} = \frac{v_o}{v_i} = -\frac{A_1 A_2}{1 + sr_{o1}C_o(A_2 + r_{o2}/r_{o1})}$$

This shows that the DC gain is multiplied by A_2 and the bandwidth is reduced by roughly the same factor.

Example Cascode Design

We want to design a cascode amplifier with the following specifications:

- DC gain = 2500
- Gain-Bandwidth product = 100MHz.
- Load capacitance = 1 pF



The two transistors in cascode configuration have identical geometries and the load is an ideal current source. Assume the following technological parameters:
$K_n' = 150\mu A/V^2, V_{Tn} = 0.5V, V_E = 20V$
Assume the supply voltage to be 3.3V.

Calculation of g_m

The gain bandwidth product is given by $\frac{g_m}{C}$. So,

$$2\pi \times 10^8 = \frac{g_m}{C} = \frac{g_{m1}}{10^{-12}}$$

$$\text{So } g_{m1} = 628.3\mu S$$

Since the same current flows through the two transistors and they have the same geometry, $g_{m1} = g_{m2}, g_{o1} = g_{o2}$.

$$\text{Let } A = \frac{g_{m1}}{g_{o1}} = \frac{g_{m2}}{g_{o2}}$$

Therefore,

$$2500 = \frac{g_{m1}}{g_{o1}} \cdot \left(1 + \frac{g_{m2}}{g_{o2}}\right) = A(A+1)$$

This gives $A \simeq 49.5$.

Calculation of bias current and geometry

$$49.5 = \frac{628.3 \times 10^{-6}}{g_{o1}} \qquad \text{so } g_{o1} = 12.7\mu S$$

$$\text{Therefore } g_{o1} = 12.7 \times 10^{-6} = \frac{I_d}{V_E} = \frac{I_d}{20}$$

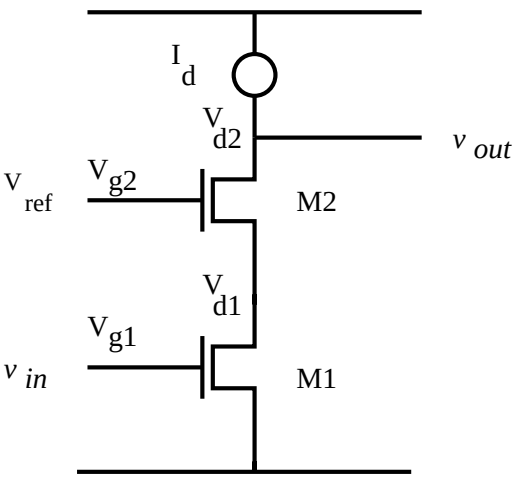
From where, the drain current is 254μA.

$$\text{Since } g_{m1} = \sqrt{2K'\frac{W}{L}I_d}, \qquad \frac{W}{L} = \frac{628.3^2 \times 10^{-12}}{300 \times 10^{-}6 \times 254 \times 10^{-6}} \simeq 5.2$$

$$\text{Therefore } g_m = 628.3\mu S, \frac{W}{L} = 5.2, I_d = 254\mu A$$

Bias Voltages

$$I_d = \frac{1}{2}K'\frac{W}{L}V_{GT}^2 \quad \text{So } V_{GT} = \sqrt{\frac{2 \times 254}{150 \times 5.2}} = .81V$$



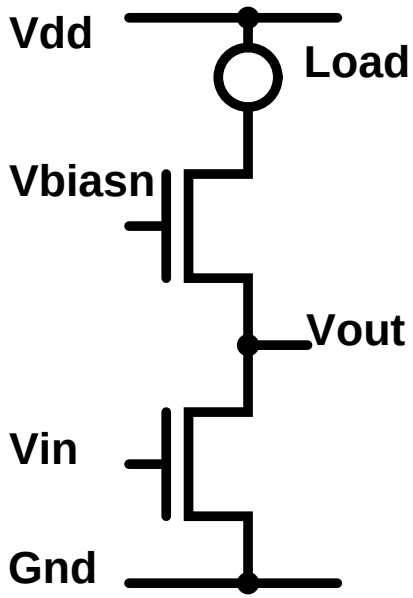
$V_{g1} \geq V_{Tn} + V_{GT} = 0.5 + 0.81 = 1.31V$
M1 will be in saturation when
$V_{d1} = V_{S2} \geq 0.81V$,
So $V_{g2} \geq 0.81 + 0.5 + 0.81 = 2.12V$.
For M2to be in saturation,
$V_{d2} \geq 2.12 - 0.5 = 1.62V$.

Thus the maximum output swing is from 1.62V to V_{dd}.

DC level incompatibility

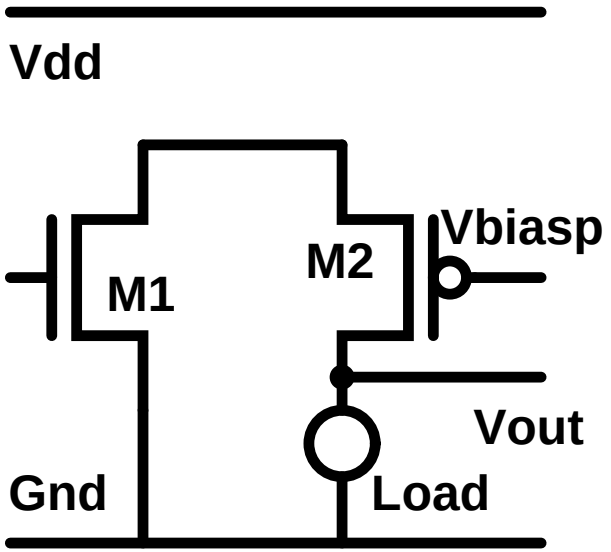
The output DC level of a cascode amplifier is higher than the input DC level. This causes problems with direct connection to the next stage, or with DC feed back to itself.

- These problems can be reduced if we use a complementary arrangement of n and p channel transistors for cascoding.
- The upper transistor of the cascode arrangement can be thought of as a source follower to its bias voltage, which keeps the drain voltage of the lower amplifier transistor (nearly) constant.
- Can we use a p channel transistor as a source follower?

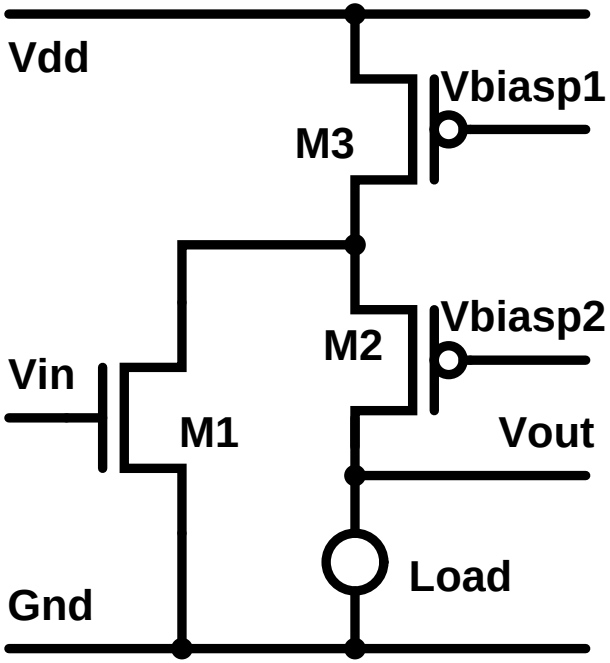


Alternative Cascode

- The p source follower will keep the drain voltage of the amplifier at $\simeq V_{biasp} + |V_{Tp}|$, allowing the cascode action as before.
- Unfortunately, the circuit won't work as there is no path between V_{dd} and ground!
- We can rectify this problem by providing a current source p load to the amplifier transistor M1.



Folded Cascode



This arrangement is called a folded cascode.
M3 provides the bias current.
M2 and M3 keep the drain voltage of M1 nearly fixed
Id3 - Id1 flows through the p channel cascoding transistor M2, which provides amplification in a common gate configuration.

$$r_{out} = (1 + g_{m2}r_{o2})(r_{o1}||r_{o3}) + r_{o2}$$

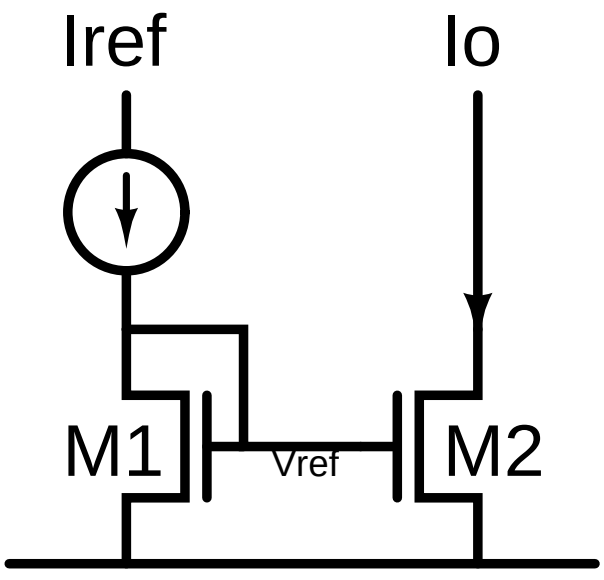
This is lower than the output resistance of the telescopic cascode stage, because of the paralleling of r_{o1} and r_{o3}. However, it is much higher than the single transistor output resistance.

Current Source Loads

Up to now we have assumed current source loads. How do we implement these?

- A transistor in saturation has a (nearly) constant drain current.
- Therefore single transistors (preferably with long channels) can be used as current sources/sinks.
- These act as current sources/sinks only over some voltage range — not for all voltages.
- There is a weak dependence on voltage due to nonzero output conductance.
- This dependence can be reduced by using a cascode stage.

A simple Current Mirror



For M1, $V_{ds} = V_{gs} > V_{gs} - V_T$
Therefore M1 is saturated.

$$I_{ref} = \frac{K}{2}(V_{ref} - V_T)^2$$

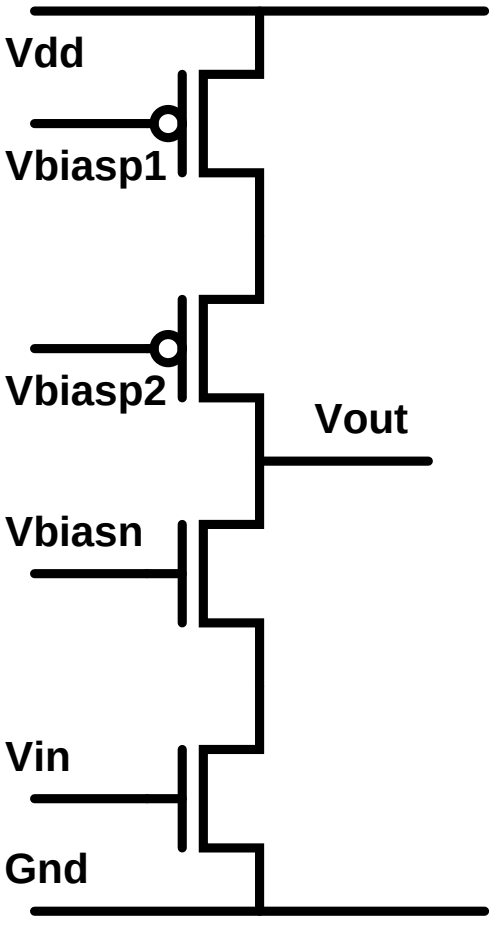
Therefore $V_{ref} = V_T + \sqrt{\frac{2I_{ref}}{K}}$

If M2 is also saturated, $I_o = I_{ref}$

Thus M2 can act as a current source load

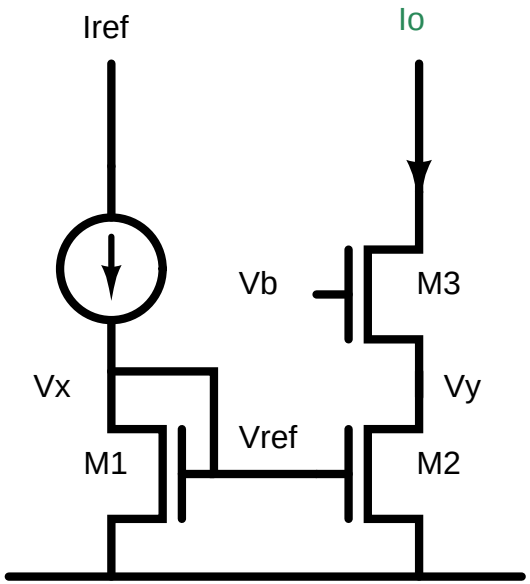
if $V_o > V_{ref} - V_T$ i.e. $V_o > \sqrt{\frac{2I_{ref}}{K}}$

Load for a Cascode stage



The output resistance of the load appears in parallel with that of the amplifying stage.
If we use a single transistor current load for a cascode, the output resistance of the load will be $\approx r_o$ while that of the cascode stage will be $\approx A \times r_o$.
The effective output resistance will thus be dominated by the much lower resistance of the load and we shall lose the advantages of the cascode stage.
It is important, therefore, that the load also should be a current source made from a cascode pair.

A cascode current mirror



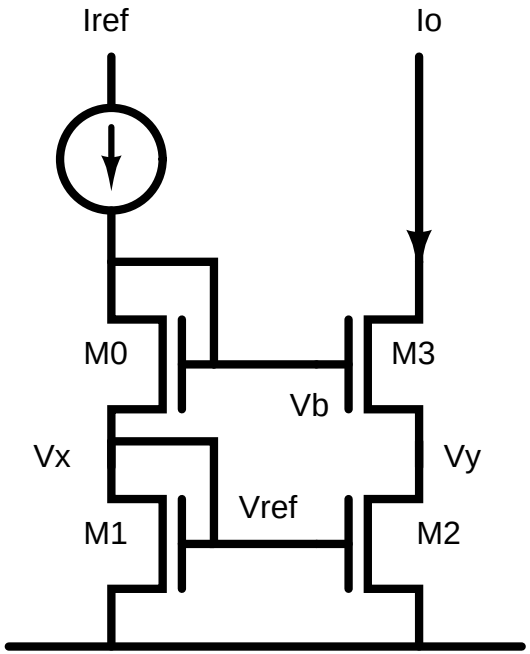
- A single transistor current mirror will have some dependence on the drain voltage due to its output resistance.
- This dependence can be reduced substantially by using a cascode stage.
- However, this reduces the available voltage range over which the transistors are saturated.

For saturation of M2 $V_y \geq V_{ref} - V_T = \sqrt{\frac{2I_{ref}}{K}}$

Therefore $V_b \geq 2\sqrt{\frac{2I_{ref}}{K}} + V_T$

For saturation of M3 $V_o \geq 2\sqrt{\frac{2I_{ref}}{K}}$

Self biased Cascode current mirror



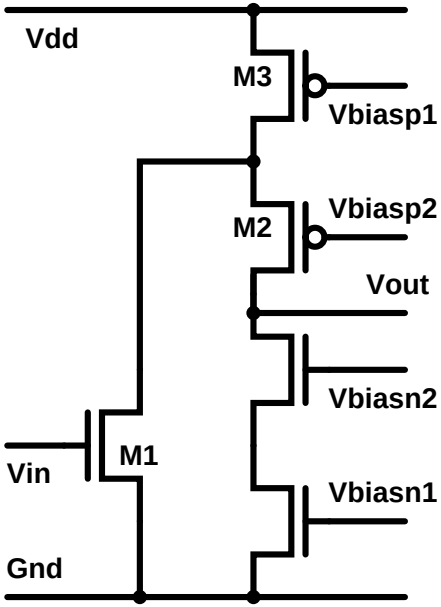
- This circuit does not need an external voltage bias.
- The reference side of the mirror generates the bias voltages for both the transistors of the cascode output side.
- However, this reduces the voltage range over which the the output may swing.

$$V_b = 2\sqrt{\frac{2I_{ref}}{K}} + 2V_T$$

$$\text{For saturation of M3 } V_o \geq 2\sqrt{\frac{2I_{ref}}{K}} + V_T$$

The output voltage needs to be a V_T higher than the minimum.

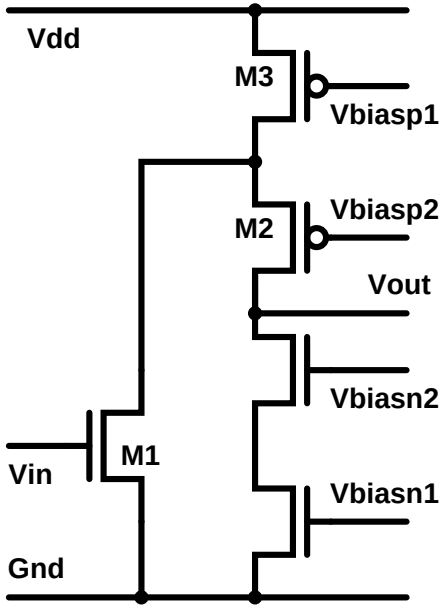
Folded Cascode with load



The load for the folded cascode should also be a cascode pair.
Here two n channel transistors in cascode configuration are used as the load.

One major advantage of the folded cascode is that the output can be directly coupled to the input for negative feedback.

Folded Cascode with Load



The single transistor amplifier can be replaced by any transconductance, of course. In operational amplifiers, the single transistor stage will be replaced by a differential amplifier.

Differential Amplifiers

Circuits which amplify the *difference* of two input voltages (each of which has equal and opposite signal excursions) have many advantages over single ended amplifiers.

- Noise picked up by both inputs gets canceled in the output.
- Input and feedback paths can be isolated.
- If both inputs have the same DC bias, the output is insensitive to changes in the bias.

Some definitions

It is more convenient to represent the two input voltages and the two output voltages by their mean and difference values.

$$\begin{aligned} v_{id} &\equiv v_{i1} - v_{i2} \\ v_{icm} &\equiv \frac{v_{i1} + v_{i2}}{2} \\ v_{od} &\equiv v_{o1} - v_{o2} \\ v_{ocm} &\equiv \frac{v_{o1} + v_{o2}}{2} \end{aligned}$$

The common mode and differential gains are:

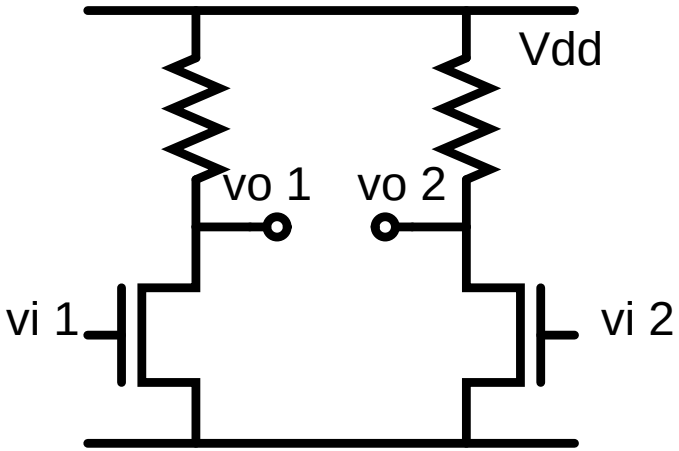
$$\begin{aligned} A_{diff} &\equiv \frac{v_{od}}{v_{id}} \\ A_{cm} &\equiv \frac{v_{ocm}}{v_{icm}} \end{aligned}$$

Common Mode Rejection Ratio

For a good diff amp, the differential gain should be high and independent of input common mode voltage, whereas the common mode gain should be as low as possible. The common mode rejection ratio is:

$$\text{CMRR} \equiv 20 \log \frac{A_{diff}}{A_{cm}} dB$$

Will this do?



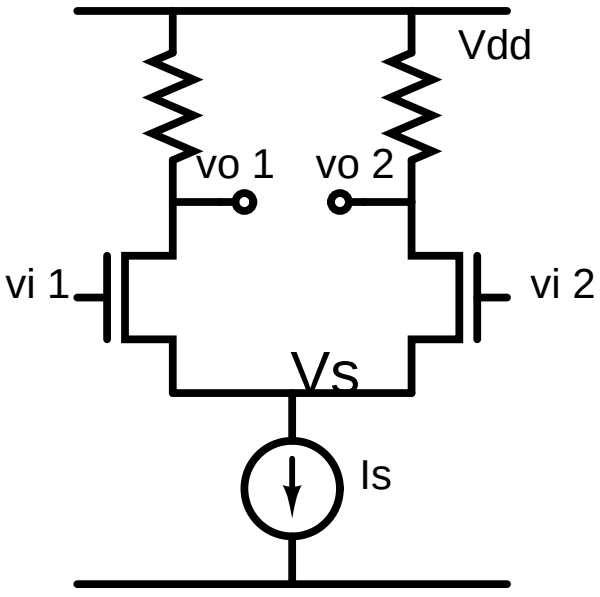
One (not very good) way of implementing a diff amp is to use two single ended amplifiers as shown above.

$$\text{Output} = V_{o1} - V_{o2}$$

Here the transistor currents, and hence the differential gain, will depend on the common mode voltage. This is not desirable as we would like the circuit to ignore the common mode voltage and to amplify just the difference signal.

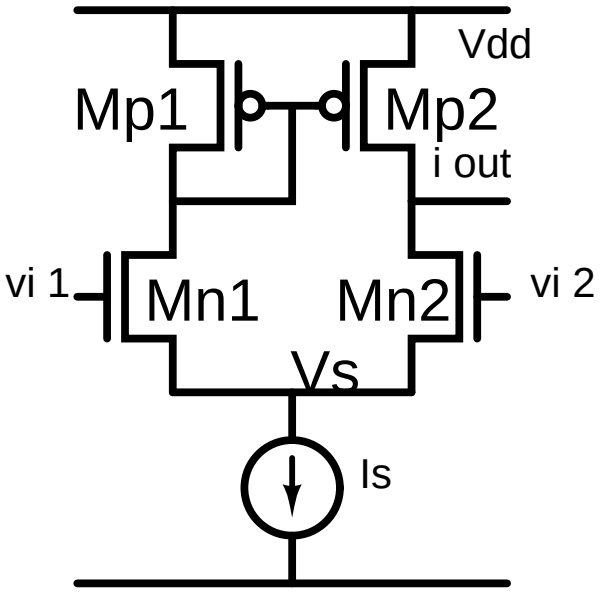
The long tail pair

A better diff amp can be implemented by adding a current source to keep the total current constant.



If the common mode voltage appearing at the two inputs changes, it will only change the voltage at the node where the two sources join (V_S). However, the current remains unchanged due to the current source - and therefore, the differential gain is unaffected by the common mode voltage.

Diff amp with single ended output



$$i_{out} = I(Mp2) - I(Mn2)$$

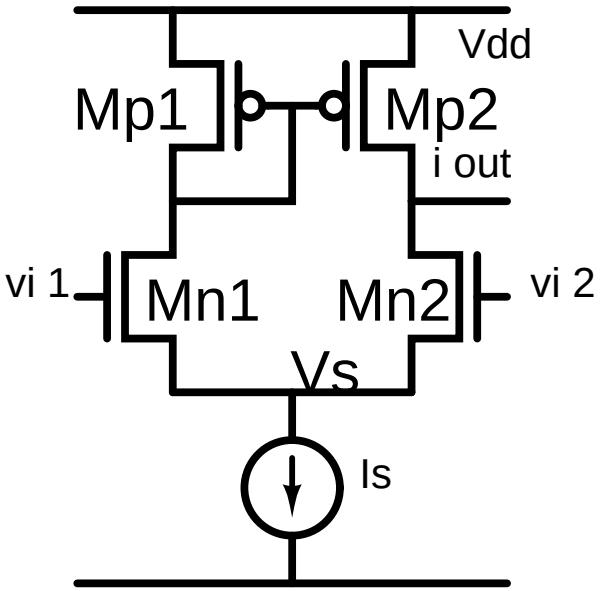
$I(Mp2) = I(Mp1)$ (current mirror)
$I(Mp1) = I(Mn1)$ (series connection)

$$i_{out} = I(Mn1) - I(Mn2) = g_m(vi_1 - vi_2)$$

$$i_{out} \equiv G_m(vi_1 - vi_2) = G_m v_{id}$$

Thus we have a single output which is proportional to the difference of inputs.
The effective G_m is just the g_m of either of the diff-pair transistors.

Gain of the OTA



This circuit is also called an operational transconductance amplifier (OTA) because the output is a current.

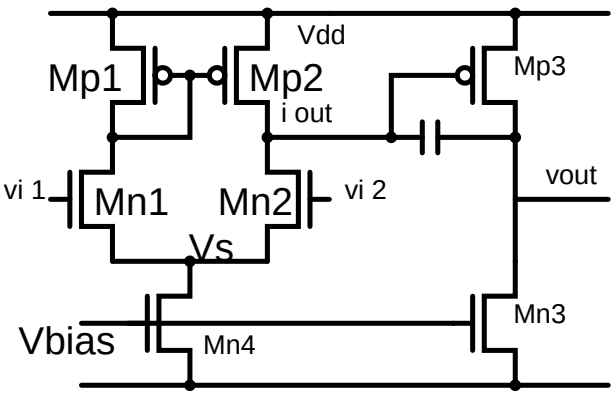
$$R_{out} = r_o(Mn2) \| r_o(Mp2)$$

$$\text{So DC voltage gain} = g_m(r_o(Mn2) \| r_o(Mp2))$$

$$\text{and } GBW = \frac{g_m}{C_L}$$

C_L includes C_{dg} and C_d for Mn2 and Mp2, as well as the load capacitance.

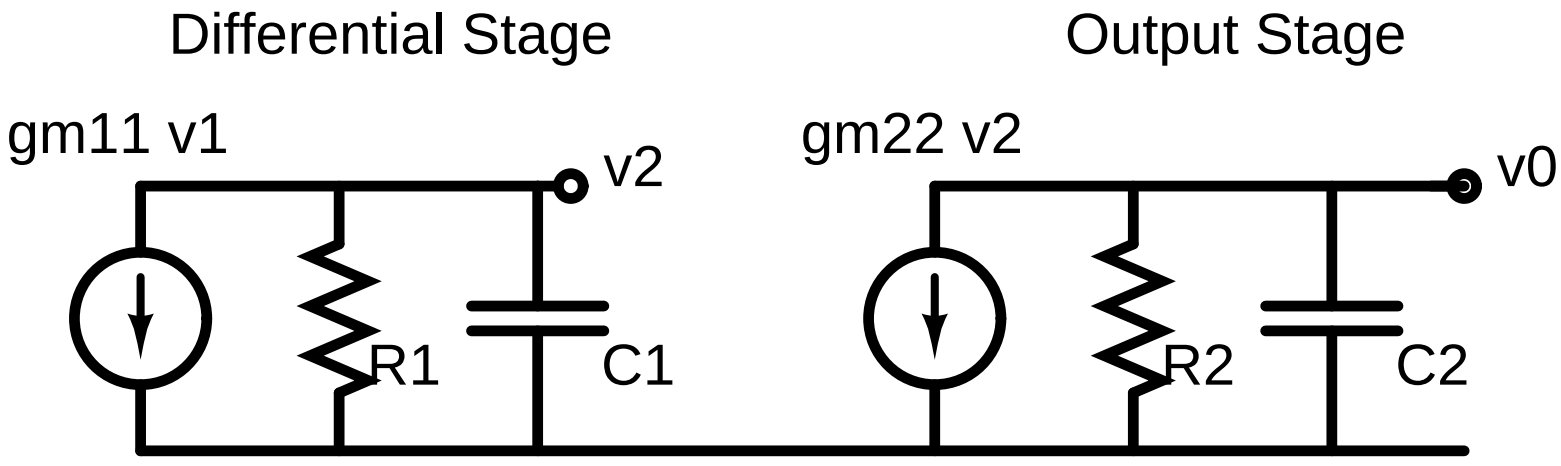
Two stage op-amp



A simple two stage op-amp can be constructed by following the diff amp by a common source stage with a constant current load.
The current source for the diff amp is implemented by an n channel MOS transistor in saturation.

The two stage design permits us to optimize the output stage for driving the load and the input stage for providing good differential gain and CMRR.
A diff amp with n transistors and an output stage with p driver is shown. However, a p type diff amp with n type common source stage is better for low noise operation.

op-amp eq. circuit



Each stage of the opamp can be considered a gain stage with a single pole frequency response.
Notice that the phase of the output of each stage will undergo a phase change of 90^o around its pole frequency.

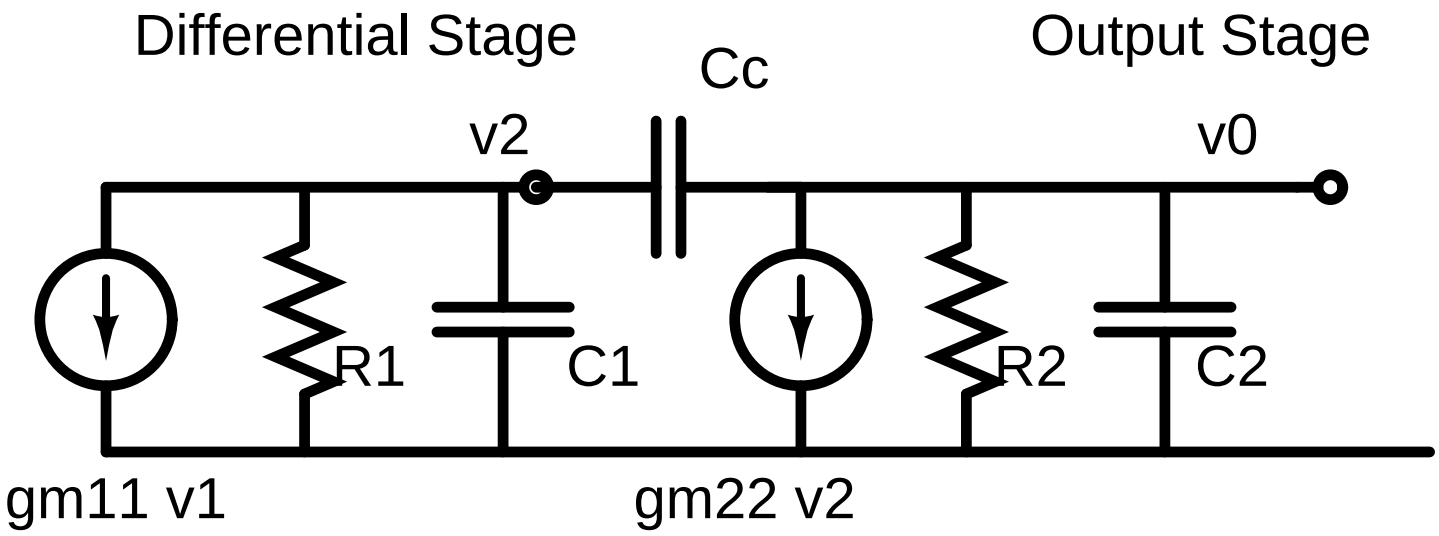
op-amp Compensation

Most opamps are used with negative feedback.
If the opamp stages themselves contribute a phase difference of 180^o, the negative feedback will appear as positive feedback.
If the gain at this frequency is > 1, the circuit will become unstable.
Both stages of the opamp have a single pole frequency response.
The poles for both the stages can be quite close together.
As a result, they can contribute a total of 180^o phase shift over a relatively narrow frequency range.

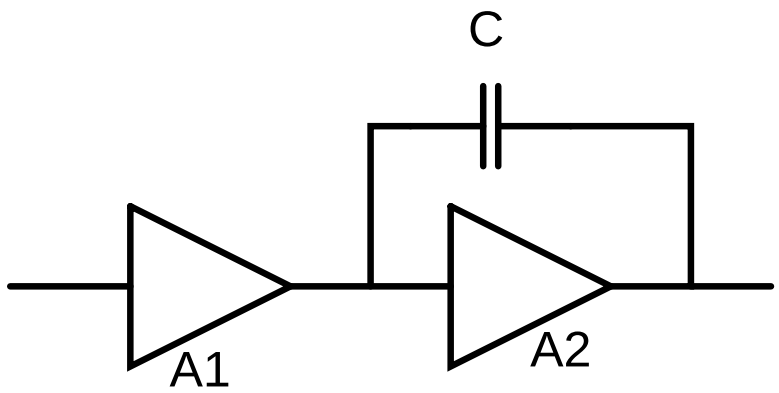
Pole Splitting

- To avoid instability, we would like to arrange things such that the gain drops to below one by the time the phase shift through the opamp becomes 180^o.
  - Even if it means that we have to reduce the bandwidth of the op amp.
- This is often achieved by a technique called pole splitting.
- The lower frequency pole is brought to a low enough frequency, so that the gain diminishes to below one by the time the second pole is reached.
- One way of doing this is to use a Miller capacitor.

Eq. Circuit of compensated Opamp



Miller Compensation



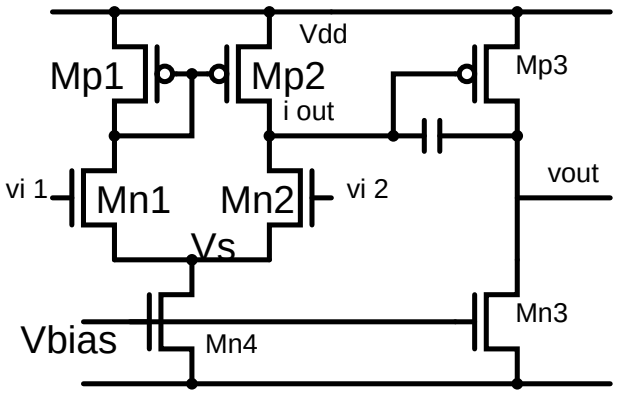
- The diff amp stage sees a load capacitance $A_2 C$.
  This brings its pole to $\frac{1}{r_{o1} A_2 C}$.
- The total DC gain is $A_1 A_2$.
  The bandwidth is set by the diff amp stage.

Therefore the gain-bandwidth product is:

$$\frac{A_1 A_2}{r_{o1} A_2 C} = \frac{A_1}{r_o 1 C}$$

Slew rate

Miller compensation also sets the slew rate of the op amp.



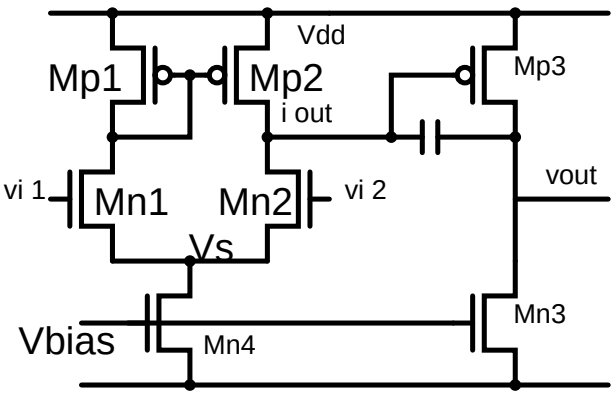
For large signal input, the output current of the OTA = tail current.

The effective load capacitance for this stage is $A_2 \times C$.

$$A_2 \times C \frac{dV}{dt} = I(Mn4)$$

- Output of the OTA slews at a rate $\frac{I(Mn4)}{A_2 \times C}$.
- So the op amp slews at a rate which is A_2 times this value.
- Hence the slew rate of the op amp is $\frac{I(Mn4)}{C}$.

Design Equations-I



All transistors must be saturated

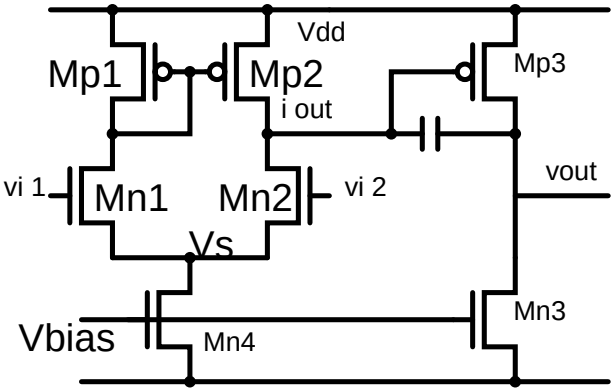
$$I(Mn1) = I(Mn2) = \frac{I(Mn4)}{2}$$

$$I(Mn1) = I(Mp1) \quad \text{(Series connection)}$$

$$I(Mp1) = I(MP2) \quad \text{(Mirror)}$$

Mp1 is always saturated.
Mp1, Mp2 have the same V_s, V_g, I_d
Since W/L(Mp2) = W/L(Mp1), MP2 will have the same V_d as Mp1, and so, will be saturated.

Design Equations-II



Mp3 has the same V_s, V_g as Mp1.

$$\text{If} \quad \frac{I(Mp3)}{I(Mp1)} = \frac{W/L(Mp3)}{W/L(Mp1)}$$

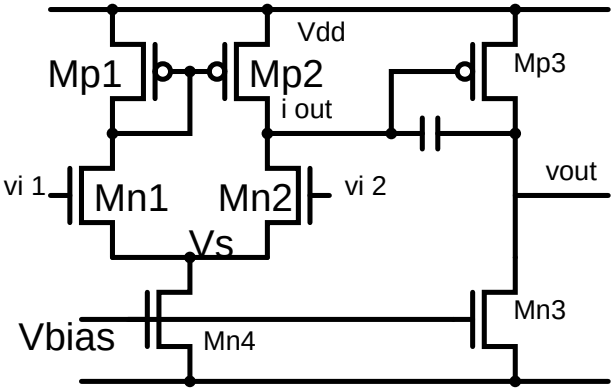
Mp3 will have the same V_d as Mp1 and will be saturated.

The slew rate determines I(Mn4).

$$I(Mn4) = C \times \text{Slew Rate}$$

$$I(Mn1) = I(Mn2) = \frac{I(Mn4)}{2}$$

Design Equations-III



GBW determines g_m of Mn1, Mn2.

$$\text{GBW} = \frac{g_m(\text{Mn2})}{C}$$

Since the current as well as g_m of Mn1 and Mn2 are now known

$$\begin{aligned} g_m(Mn2) &= \sqrt{2K'W/L(Mn2)I(Mn2)} \\ W/L(Mn1) &= W/L(Mn2) \end{aligned}$$

This will determine the geometries of Mn1 and Mn2.

Design Equations-IV

Currents through Mn2,Mp2, Mp3 and Mn3 are known

$$(g_o = I_d/V_A) \qquad \text{where } V_A \text{ is the Early voltage } = L/\lambda'$$

The overall DC gain is given by

$$A = \frac{g_m(Mn2)g_m(Mp3)}{(g_o(Mn2)||g_o(Mp2))(g_o(Mp3)||g_o(Mn3))}$$

As $g_m$for Mn2 and all $g_o$values are known, this determines the $g_m$for MP3.
Once we know the $g_m$as well as the current for Mp3, we can calculate its geometry.

Example Design: Specifications

$$K'(n) = 120\mu\mathrm{A/V}^2, K'(p) = 60\mu\mathrm{A/V}^2$$

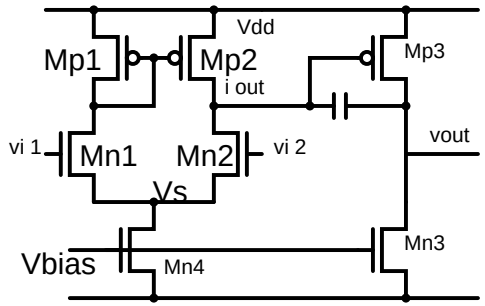
$$V_T(n) = 0.4\mathrm{V}, V_T(p) = -0.4\mathrm{V}$$

Early Voltage $V_A = 20V$ for both p and n channel transistors

Op amp DC gain $= 80dB$ (Voltage gain of 10000)

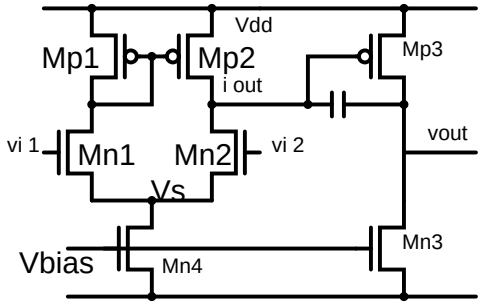
Gain Bandwidth product $= 50MHz$, slew rate $= 20V/\mu s$

Example Design-1



- We choose a compensation capacitor value of 2 pF.
- We shall bias the second stage at 5 times the tail current of the differential stage.
- From the slew rate, I(Mn4) = $2 \times 10^{-12} \times \frac{20}{10^{-6}} = 40\mu A$
- Therefore I(Mn1) = I(Mn2) = I(Mp1) = I(Mp2) $= 20\mu A$ and I(Mp3) = I(Mn3) = $= 200\mu A$

Example Design-2



- From the GBW requirement,

$$2\pi \times 50 \times 10^{6} = \frac{g_m(\mathrm{Mn2})}{2 \times 10^{-12}}$$

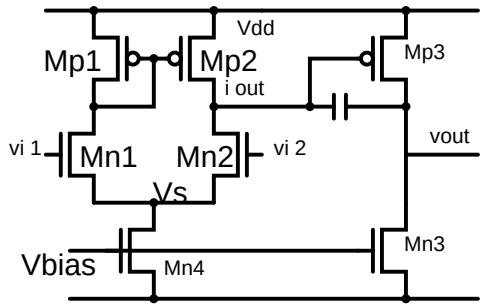
This gives $g_m(\mathrm{Mn2}) \simeq 628\mu\mho$.

- To get a $g_m$of 628 $\mu\mho$ with a current of $20\mu\mathrm{A}$,

$$628 \times 10^{-6} = \sqrt{2 \times 120 \times 10^{-6} \times (W/L) \times 20 \times 10^{-6}}$$

this gives W/L(Mn2) $\approx$ 82 = W/L(Mn1)

Example Design-3

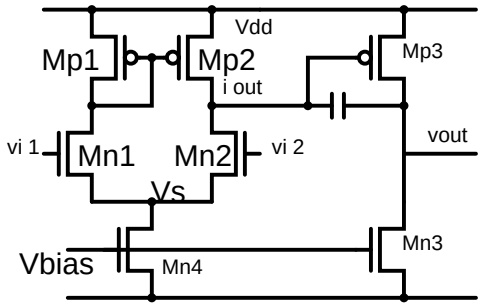


- $g_o$of Mn2 and Mp2 $= 20\mu\text{A}/20\text{V} = 1\mu\mho$.
  Therefore $g_o(\text{Mn2}) \| g_o(\text{Mp2}) = 2\mu\mho$.
  $g_o$of Mn3 and Mp3 is $= 200\mu\text{A}/20\text{V} = 10\mu\mho$.
  Therefore $g_o(\text{Mp3}) \| g_o(\text{Mn3}) = 20\mu\mho$.

$$\text{DC gain} = 10000 = \frac{628\mu\mho}{2\mu\mho} \times \frac{gm(Mp3)}{20\mu\mho}$$

$$\text{So, } g_m(Mp3) \simeq 637\mu\mho$$

Example Design-4

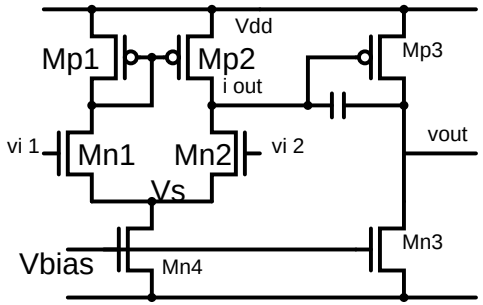


- To get a $g_m$of 637$\mu\mho$ with a drain current of 200μA, we should have

$$637 \times 10^{-6} = \sqrt{2 \times 60 \times 10^{-6} \times (W/L) \times 200 \times 10^{-6}}$$

  which gives the W/L of Mp3 $\approx$ 17.
- Since the geometry of Mp1 and Mp2 has to be in the current ratio with Mp3, W/L of Mp1 and Mp2 should be $\approx$ 1.7.

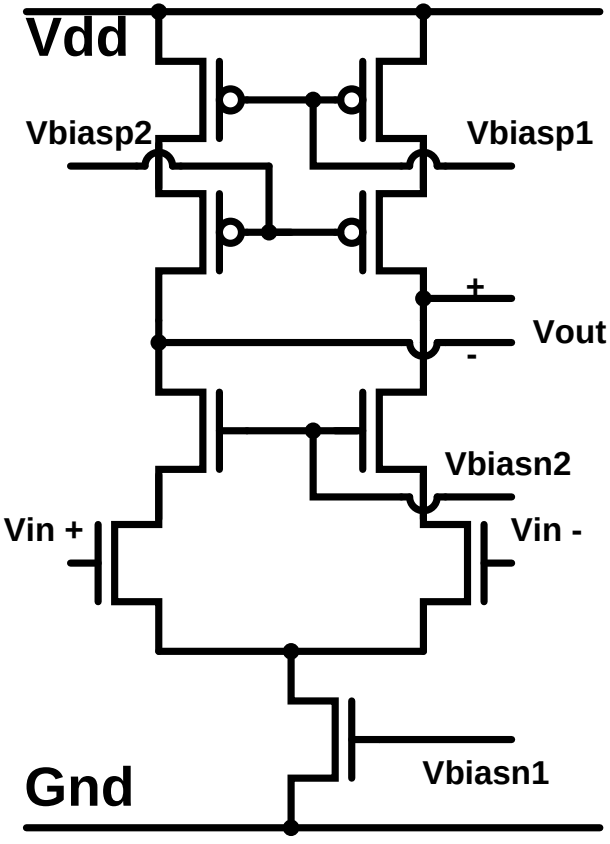
Example Design-5



- Finally, we assume that an n type reference bias transistor of W/L = 4 is available with a current of 10 μA. This will give the W/L of Mn4 and Mn3 as 16 and 80 respectively.

This completes the design for the simple two stage op amp.

Telescopic Cascode Opamp



The telescopic cascode is a differential version of the cascode amplifier discussed earlier.

Its gain is comparable to the two stage op-amp.

The output impedance is (very) high!

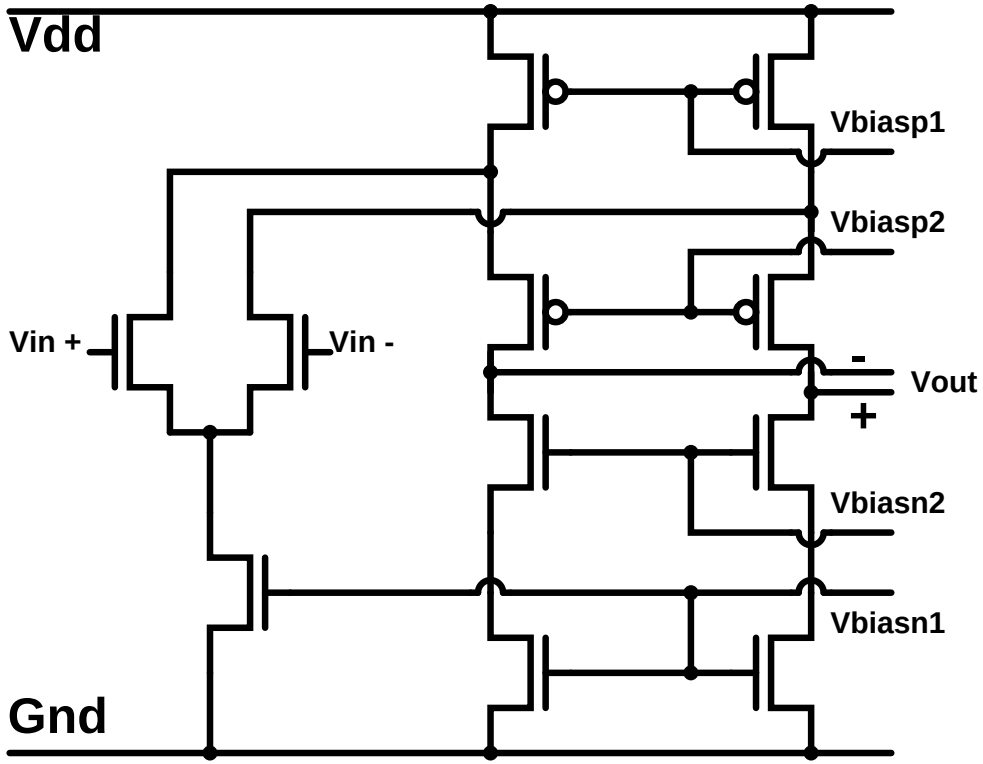
The output impedance in conjunction with the load capacitance constitutes the dominant pole of the system.

Telescopic Cascode Opamp

- Gain is comparable to the two stage opamp (product of two single stage amplifiers).
- It needs a higher supply voltage compared to a two stage opamp.
- The output stage is high impedance, so the dominant pole is at the output.
- Compensation is provided by the load capacitance. So a minimum value of load capacitance is required for stability.
- The output common mode voltage is different from the input common mode voltage range.
- This presents difficulties in direct coupling to the next stage and DC feedback to its own input.

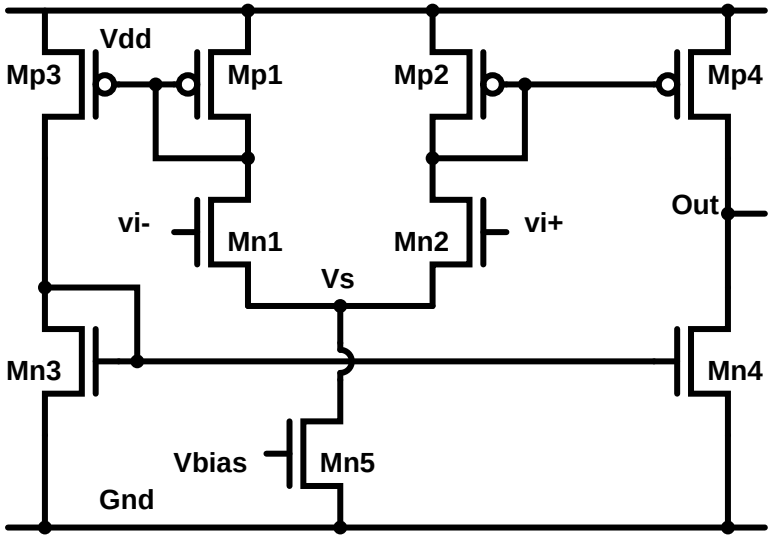
Folded Cascode

The common mode voltage incompatibility of a telescopic cascode can be solved by using a folded cascode.



Push-Pull Op Amp

Differential to single ended conversion can be done in the output stage, by using a push-pull driver. The output loads in the differential stage (Mp1 and Mp2) are diode connected.



- Current through Mp2 is mirrored in the output p transistor Mp4.
- Current through Mp1 is mirrored into a pMOS (Mp3) and passed through a diode connected nMOS (Mn3).
- This current is mirrored in the output stage nMOS (Mn4).
- Mirroring ratio of Mp4 to Mp2 and Mn4 to Mn3 should be identical (and can be large).