Operational Amplifier (OPAMP)

Operational Amplifiers

- Analog ICs include
  - Operational amplifier
  - Filters
  - Analog-to-digital converter (ADC)
  - Digital-to-analog converter (DAC)
  - Analog modulator
  - Phase-locked loop
  - Power management
  - Others
- Basic building blocks of analog ICs
  - Single-stage amplifier
  - Differential pairs
  - Current mirrors
  - MOS switches
  - Others

Operational Amplifiers (Cont.)

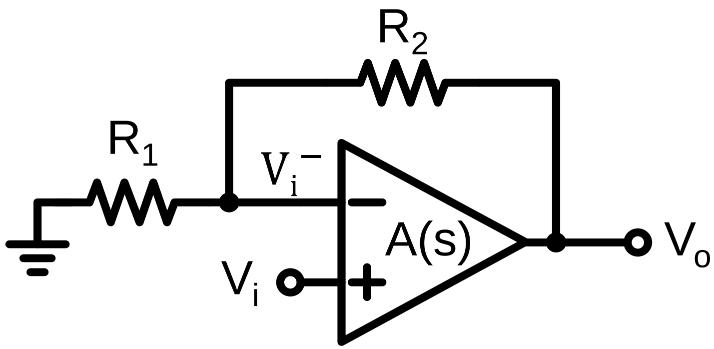
- OPAMP design
  - CMOS OPAMPs are adequate for VLSI implementation.
    - Main stream
    - Two-stage and folded-cascode OPAMPs will be introduced
  - Bipolar OPAMPs
    - Can achieve better performance than CMOS OPAMPs.
    - Less popular
    - 741 OPAMP will be introduced.
  - BiCMOS OPAMPs
    - Combine the advantages of bipolar and CMOS devices.
    - Less popular
    - First published by H. C. Lin in 1960's.

- Two-stage → I guess, it's for 70% applications

- Folded-cascode → I guess, it's for 20% applications.

Operational Amplifiers (Cont.)

- Operational amplifier with negative feedback



$$V_i^-(s) = \beta V_o(s)$$

$$V_o(s) = A(s)(V_i(s) - V_i^-(s))$$

$$A_f(s) = \frac{V_o(s)}{V_i(s)} = \frac{A(s)}{1 + \beta A\,(s)}$$

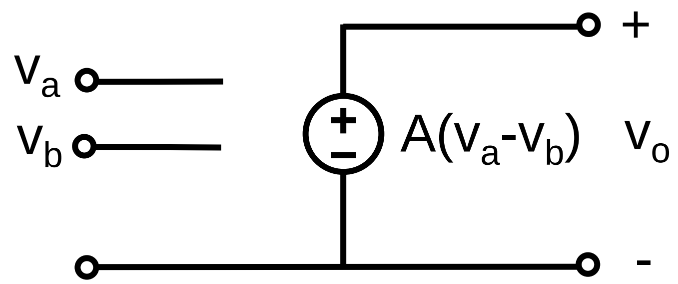
for A_f is closed-loop gain, A is open-loop gain, βA is loop gain

$$\beta = \frac{R_1}{R_1 + R_2}$$

  - Open-loop: Always stable (no internal feedback)
  - Closed-loop: Stability depends on βA(s)
- For stable system, the real part of all poles must be negative.
  - Gain margin = $20\log|\beta A(j\omega_{180})|$
  - Unity-gain frequency ω_t
  - Phase Margin = $\measuredangle \beta A(j\omega_t) + 180°$
    - At least 45º ~ 60º (or larger) margin is preferred
    - This will also give a desirable (i.e., small or no ringing) step response for the closed-loop amplifier

Operational Amplifiers (Cont.)

- Ideal voltage op-amp
  - Voltage-controlled voltage source
  - Infinite voltage gain
  - Infinite input impedance
  - Zero output impedance
  - No noise
  - Infinite bandwidth
  - No offset voltage
  - Infinite CMRR



- Differences between the ideal op-amp and real op-amp
  - Finite gain (practical op-amps, $A \approx 10^2 \sim 10^4$, i.e., 40~80dB)
  - Finite linear range($V_{DD} > V_o > GND$)

Operational Amplifiers (Cont.)

- Offset voltage:
  - Ideal op-amps $V_a=V_b \Rightarrow V_o=0$
  - For real op-amps, this isn't exactly true and $V_o \neq 0$ is always occurred.
  - Input offset voltage V is defined as the differential input voltage needed to restore $V_o=0$.
  - For MOS op-amps, V_{offset} is about 5-15 mV.

    For BJT op-amps, V_{offset} is about 1-2 mV.

- Common Mode Rejection Ratio(CMRR)
  - The CMRR measure how much the op-amp can suppress common-mode signal at its input.
  - Typically CMRR=60~80dB common-mode input voltage: $V_{in,c}=(V_a+V_b)/2$

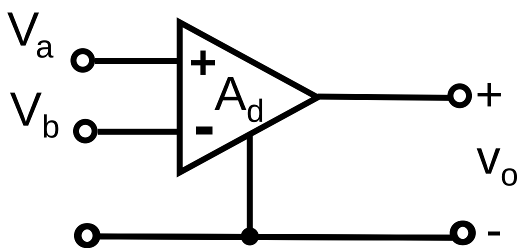
Operational Amplifiers (Cont.)

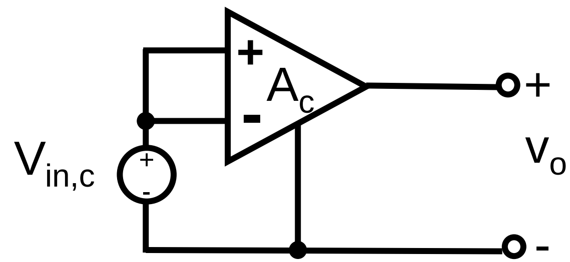
Differential-mode input voltage: $V_{in,d} = V_a - V_b$

Differential gain: $A_d = \frac{V_o}{V_{in,d}}$

Common-mode gain: $A_c = \frac{V_o}{V_{in,c}}$

CMRR = (A_d/A_c) or $20\log_{10}(A_d/A_c)$ in dB

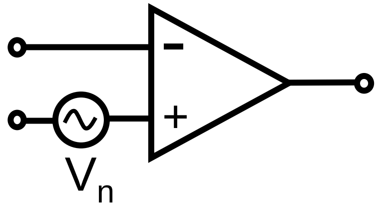




- Frequency response
  - Limited bandwidth (10GHz unity-gain bandwidth is typical)
  - Reasons of gain decreasing at high frequency
    - Stray capacitances
    - Finite carrier mobilities

Operational Amplifiers (Cont.)

- Slew Rate (typically, for MOS op-amps, 1~50V/μs)
  - The maximum rate of output change dV_o/dt
  - For a large input voltage, some transistors may be driven out of their saturation regions or completely cut off. As a result, the output will follow the input at a slower finite rate.
- Nonzero Output Resistance
  - 0.1~5kΩ → typical value
  - Large R will limit frequency response(i.e., speed) when a capacitor is connected to its output.
- Noise
  - Noisy transistors in op-amps give rise to a noise voltage V_{on} at the output of op-amp.
  - Equivalent input noise voltage=$V_{on}/A=V_n$



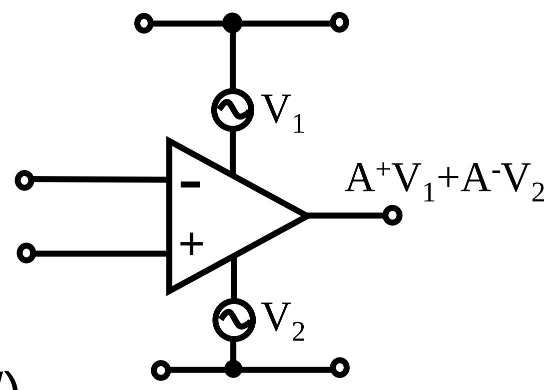
Operational Amplifiers (Cont.)

- Dynamic Range(DR) = $20\log_{10}(\frac{V_{in,max}}{V_{in,min}})$
  - Open loop~30-40dB

    $$V_{in,min} \approx \sqrt{\overline{V_n^{\ 2}}} \sim 30\mu V$$

    $$V_{in,max} \approx \frac{V_{dd}}{A}$$

  - Close loop~100dB has larger DR than open loop.
  - Can be increased by using correlated double sampling (CDS)
- PSRR (Power supply rejection ratio)
  - PSRR⁺= $20\log_{10}(\frac{A_d}{A^+})$
  - PSRR⁻= $20\log_{10}(\frac{A_d}{A^-})$

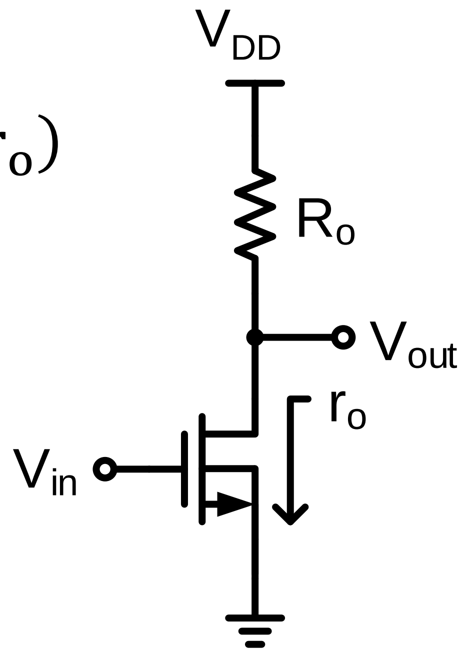


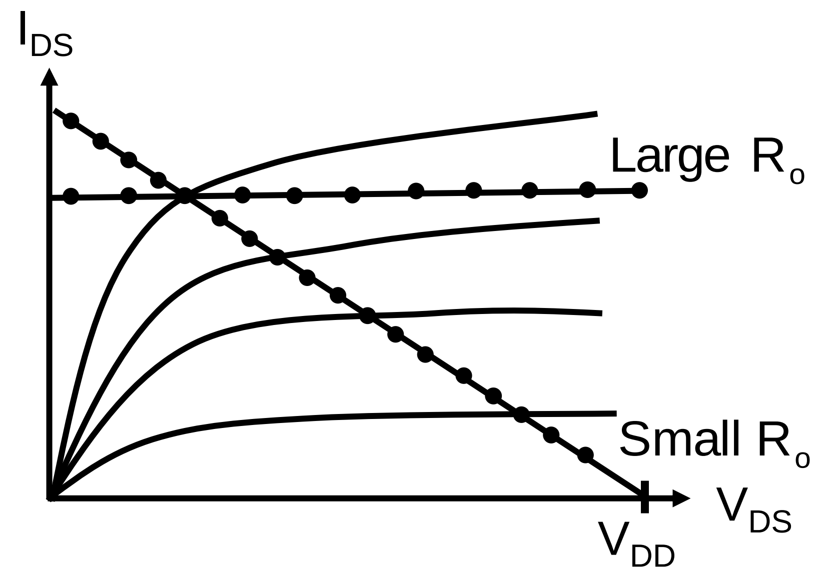
- DC Power Dissipation(10μW~100μW)

CMOS Amplifier with Resistive Load

- Resistor Load

$$A = -g_m(R_o//r_o) \approx -g_m R_o \propto \frac{-I_D R_o}{V_{OV}}$$

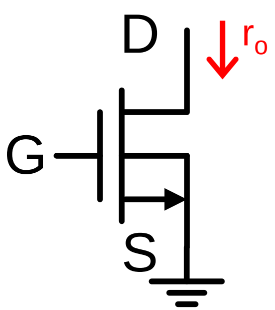


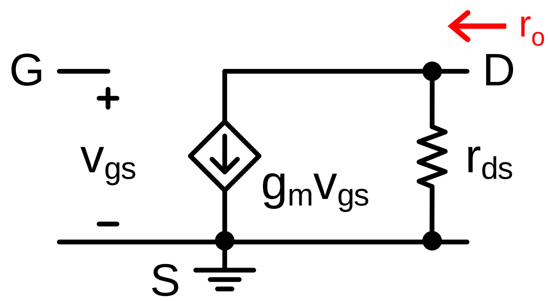


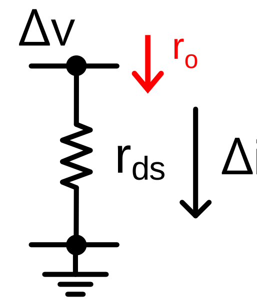
- For high gain
  - High $I_D R_o$
    - High $I_D R_o$ means large voltage drop on R_o
    - Large power supply
  - High R_o reduces speed
  - Use active loads to overcome the above problems

Resistance of Active Load

- Small signal model of NMOS

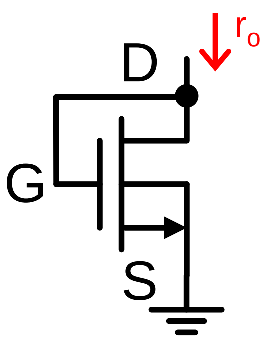


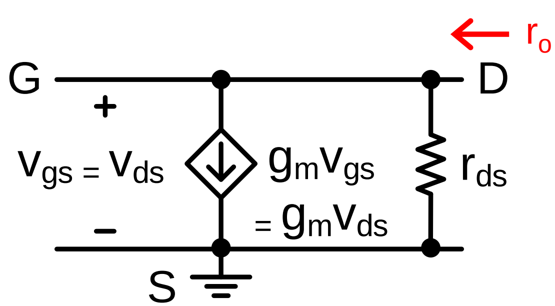


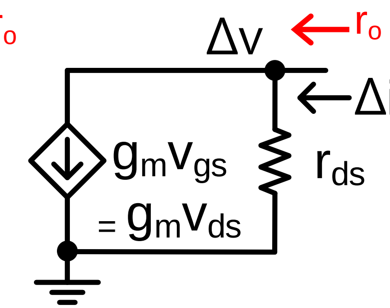


$$r_o = \frac{\Delta v}{\Delta i} = r_{ds}$$

- Small signal model of diode-connected NMOS



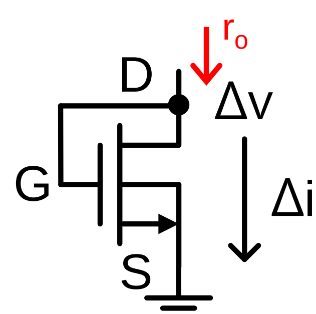




$$\Delta i = g_m \Delta v + \frac{\Delta v}{r_{ds}}$$

$$r_o = \frac{\Delta v}{\Delta i} = r_{ds} || \frac{1}{g_m}$$

  - ◆ Same analysis method



$$r_o = \frac{\Delta v}{\Delta i} = r_{ds} || \frac{1}{g_m}$$

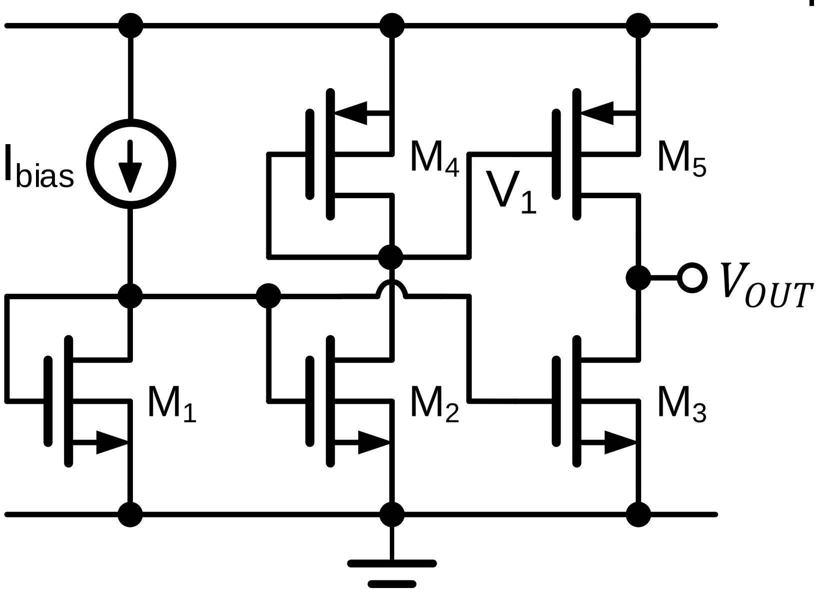
Variation of Quiescent Point

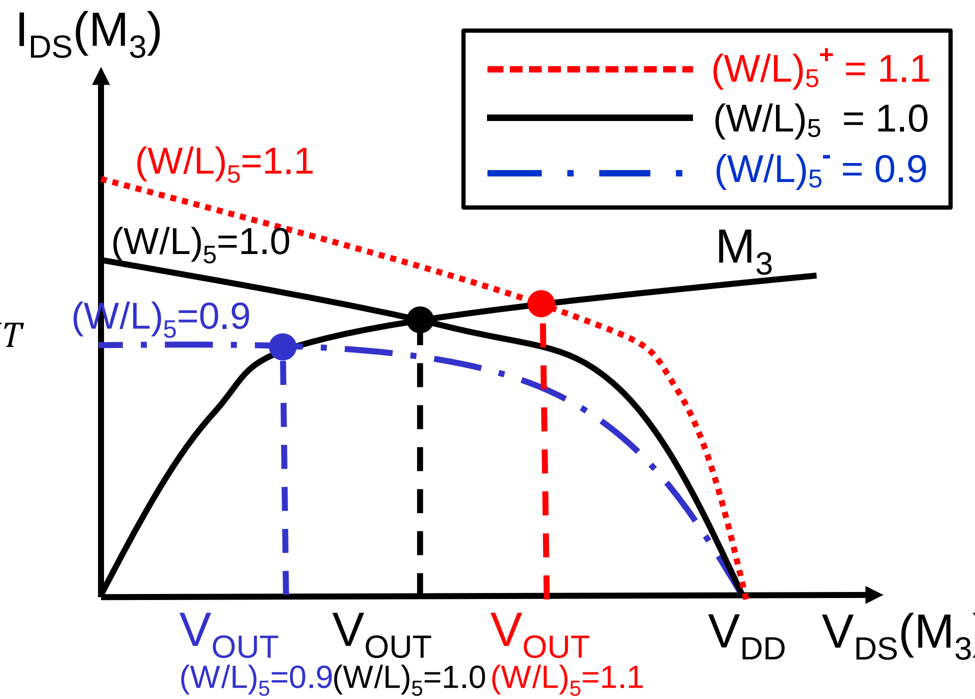
- The effect of process variations on quiescent point V_{OUT}

  Design: $\left(\frac{W}{L}\right)_1 = \left(\frac{W}{L}\right)_2 = \left(\frac{W}{L}\right)_3 = \left(\frac{W}{L}\right)_4 = \left(\frac{W}{L}\right)_5 = 1$

  if $\left(\frac{W}{L}\right)_1 = \left(\frac{W}{L}\right)_2 = \left(\frac{W}{L}\right)_3 = \left(\frac{W}{L}\right)_4 = 1$, $\left(\frac{W}{L}\right)_5 = 1 \pm 10\%$ Due to process variations

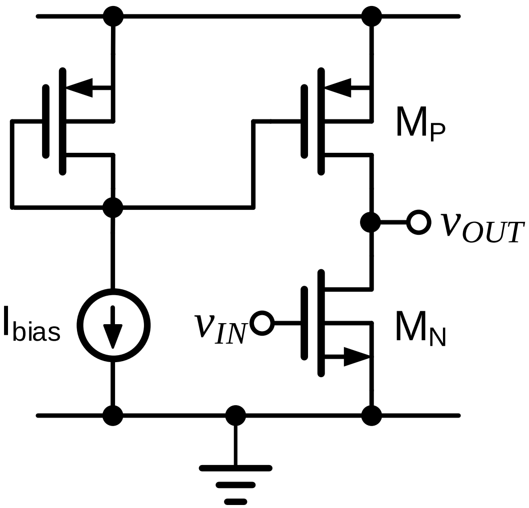
  → V_{OUT} is determined by the actual values of $M_1 \sim M_5$

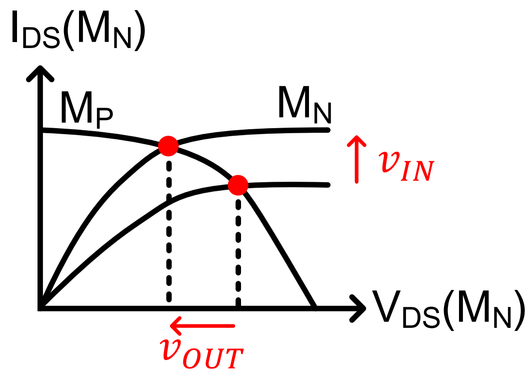




Single-Ended Amplifier with Active Load

- N-input common-source amplifier



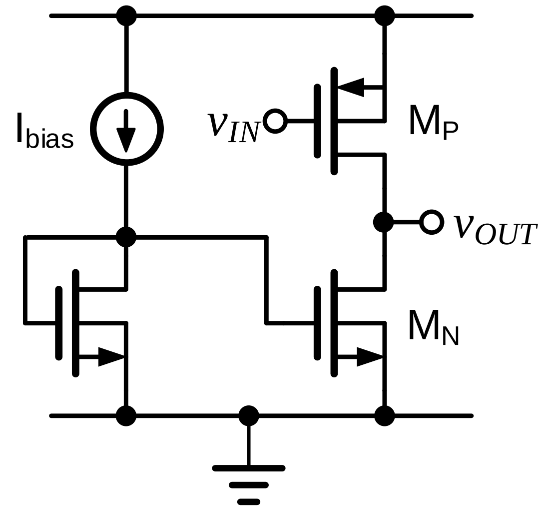


$$v_{IN} = V_{IN} + C\, sin\, \omega t$$

$$v_{OUT} = V_{OUT} + AC\, sin\, \omega t$$

where $A = \frac{v_{out}}{v_{in}} = -g_{mn}\left(r_{dsp} \parallel r_{dsn}\right)$

- P-input common-source amplifier



$$v_{IN} = V_{IN} + C\, sin\, \omega t$$

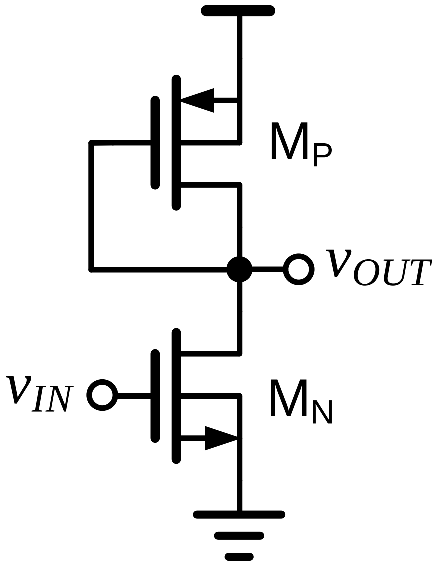
$$v_{OUT} = V_{OUT} + AC\, sin\, \omega t$$

where $A = \frac{v_{out}}{v_{in}} = -g_{mp}\left(r_{dsp} \parallel r_{dsn}\right)$

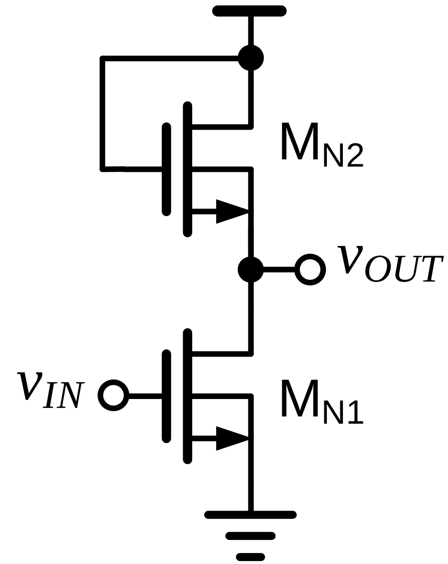
→ Quiescent point V_{OUT} is hard to be determined with active load

Single-Ended Amplifier with Diode-Connected Load

- Input and load transistor type
  - Different type
  - Same type



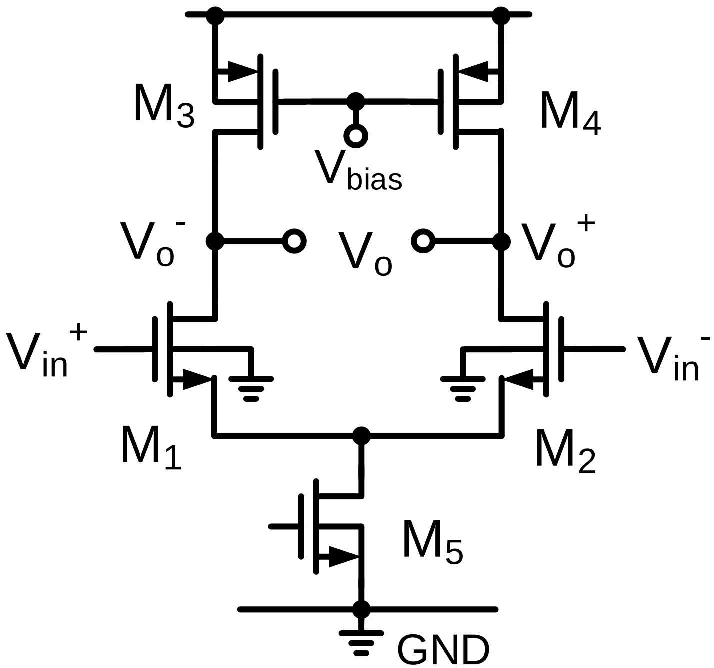
$$v_{IN} = V_{IN} + C\ sin\ \omega t$$
$$v_{OUT} = V_{OUT} + AC\ sin\ \omega t$$
$$A = \frac{v_{out}}{v_{in}} = -g_{mn}\left(r_{dsp} \parallel r_{dsn} \parallel \frac{1}{g_{mp}}\right)$$
$$= -\frac{g_{mn}}{g_{mp}} \text{ (gain with small } \sqrt{\frac{\mu_n}{\mu_p}} \text{ variation)}$$

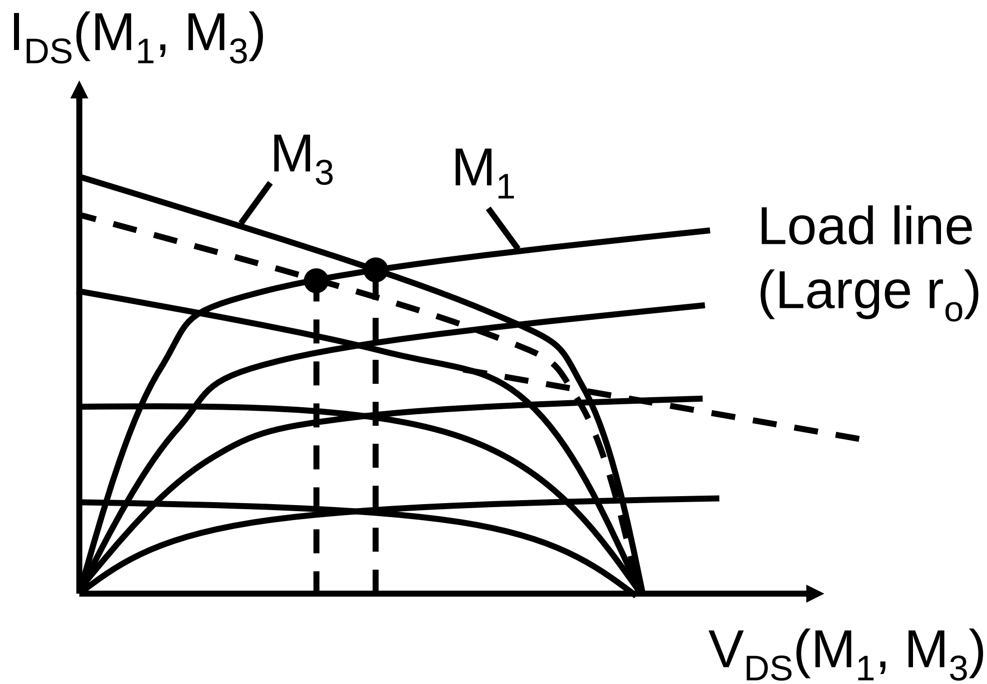


$$v_{IN} = V_{IN} + C\ sin\ \omega t$$
$$v_{OUT} = V_{OUT} + AC\ sin\ \omega t$$
$$A = \frac{v_{out}}{v_{in}} = -g_{mn1}\left(r_{dsn1} \parallel r_{dsn2} \parallel \frac{1}{g_{mn2}}\right)$$
$$= -\frac{g_{mn1}}{g_{mn2}} \approx \sqrt{\frac{(W/\mathrm{L})_{\mathrm{MN1}}}{(W/\mathrm{L})_{\mathrm{MN2}}}} \text{ (accurate gain)}$$

Differential Amplifier with Active Load

- With external bias





- Why not ?
  → Quiescent point of V_o^+ & V_o^- can't be determined due to process variations

CMOS Amplifier with Active Load (Cont.)

- Self-biased active load: quiescent V_o less sensitive to M_1 ~ M_4 variations
- Performs differential gain and differential to single-ended

  $g_{m,M1}, g_{m,M2}, g_{m,M3}, g_{m,M4} \gg \frac{1}{r_{ds}}$; $r_{out} \approx r_{ds2} \parallel r_{ds4}$
- Differential gain A_{dm} $(v_{i1} = -v_{i2} = \frac{1}{2} v_{in})$

  $A_{dm} \approx g_{m1}(r_{ds2} \parallel r_{ds4})$ at node B = A_{dmB}

  Node A: $A_{dmA} \approx -\frac{1}{2} g_{m1} \cdot \frac{1}{g_{m3}}$

  Node B: $A_{dmB} \approx \left(-A_{dmA} \cdot g_{m4} - \left(-\frac{1}{2} g_{m2}\right)\right) \cdot (r_{ds2} \parallel r_{ds4})$
- Common-mode gain A_{cm} $(v_{i1} = v_{i2} = v_1)$

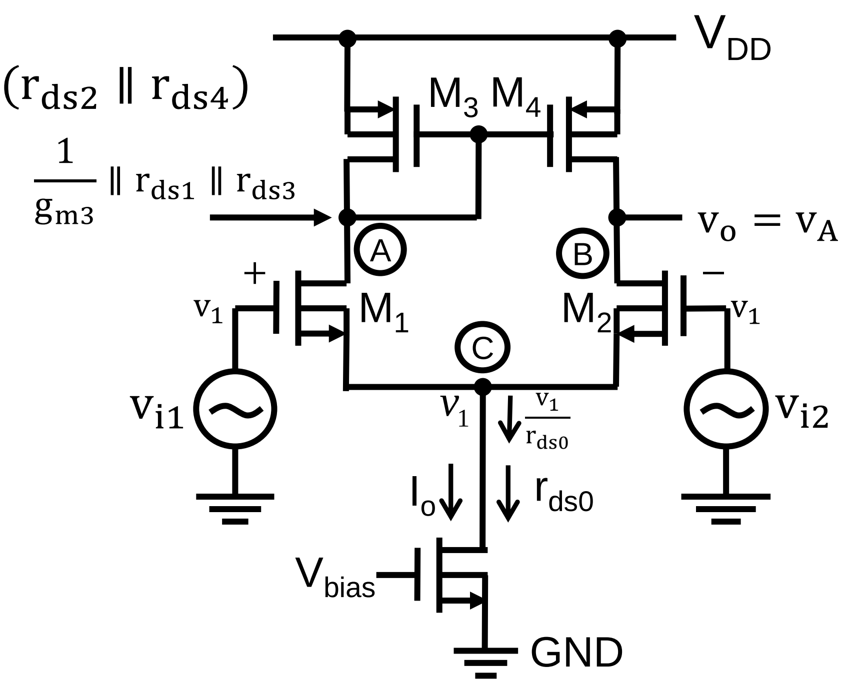
$$v_A = \frac{1}{2}\frac{v_1}{r_{ds0}}\left(\frac{1}{g_{m3}} \parallel r_{ds1} \parallel r_{ds3}\right)$$

$$A_{cm} \approx \frac{1}{2}\frac{1}{r_{ds0}}\left(\frac{1}{g_{m3}} \parallel r_{ds1} \parallel r_{ds3}\right) \approx \frac{1}{2 g_{m3} r_{ds0}}$$

- CMRR(Common-Mode Rejection Ratio)

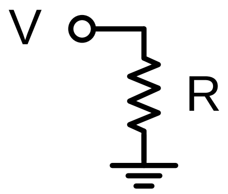
$$CMRR = \frac{A_{dm}}{A_{cm}} \approx 2 g_{m1}(r_{ds2} \parallel r_{ds4}) g_{m3} r_{ds0}$$

◆ Model of A_{dm}/A_{cm}



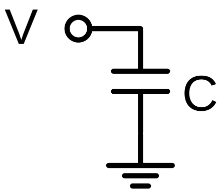
Miller Effect

- Resistor



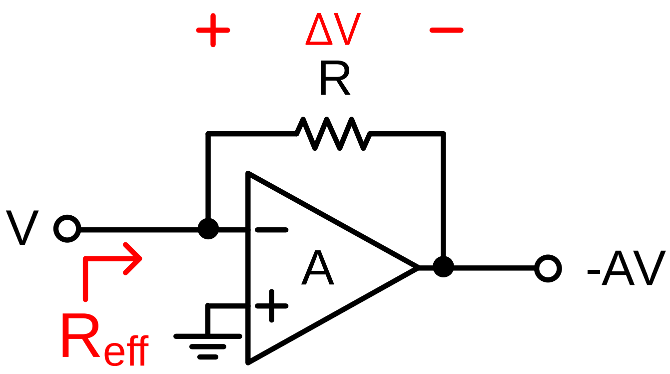
$$I = \frac{V}{R}$$

- Capacitor



$$Q = CV$$

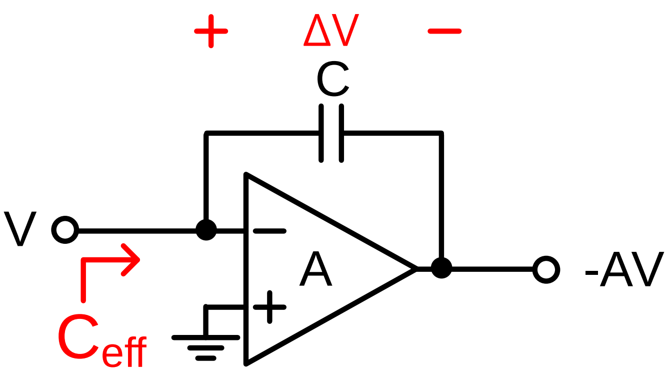
- Miller effect on resistor



$$I = \frac{\Delta V}{R} = \frac{(1 + A)V}{R}$$

$$R_{eff} = \frac{R}{1 + A}$$

- Miller effect on capacitor



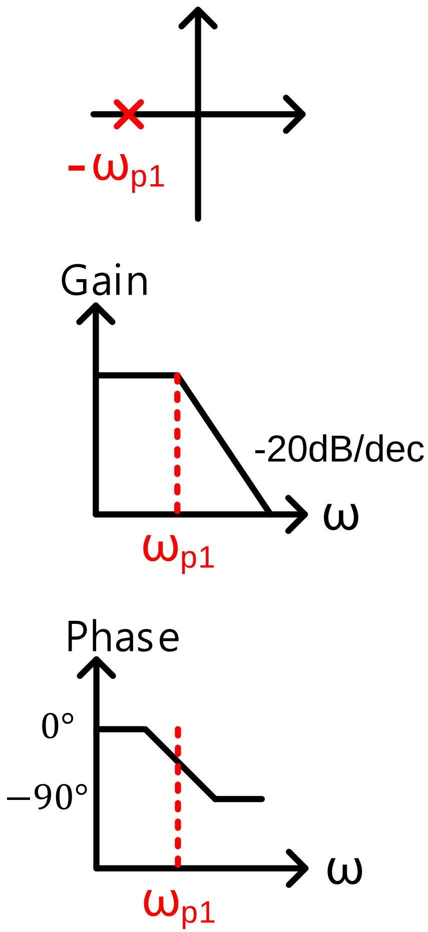
$$Q = C \cdot \Delta V = C \cdot (1 + A)V$$

$$C_{eff} = (1 + A)C$$

Pole and Zero

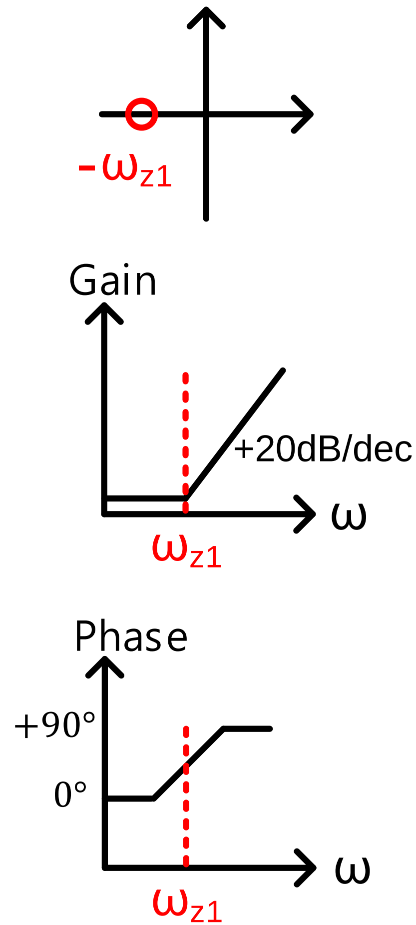
- LHP pole

$$H(s) = \frac{1}{1 + s/\omega_{p1}}$$



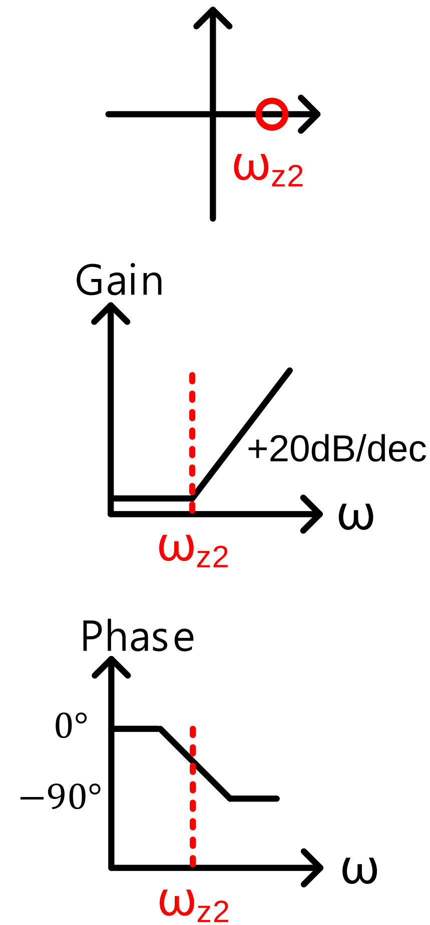
- LHP zero

$$H(s) = 1 + s/\omega_{z1}$$



- RHP zero

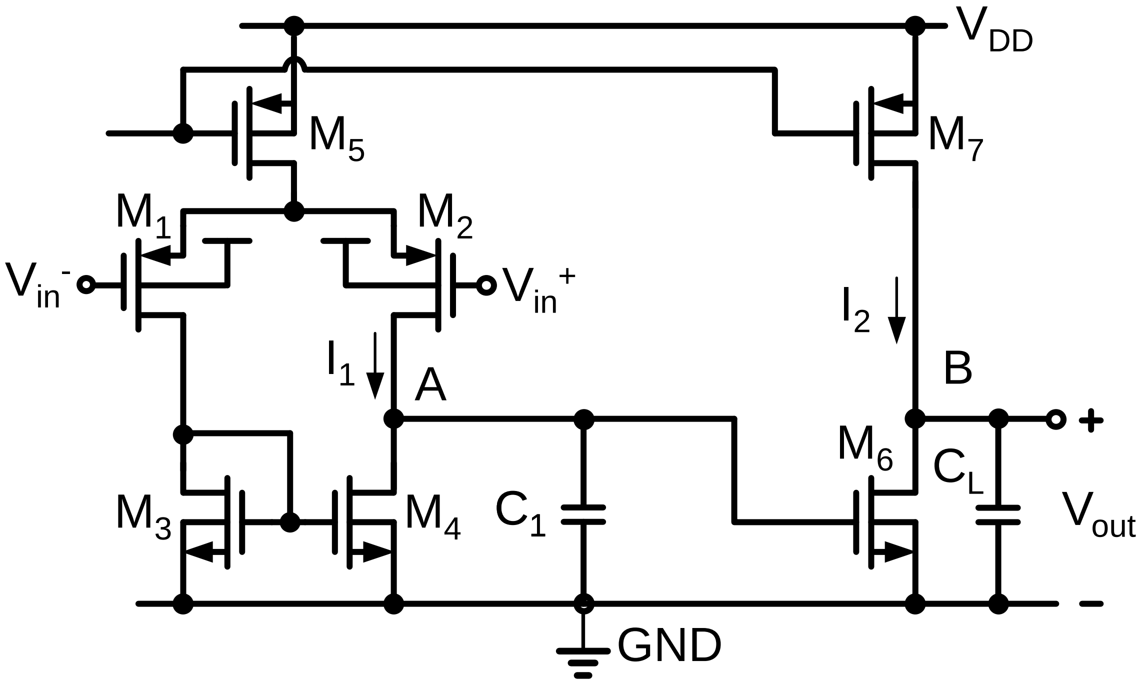
$$H(s) = 1 - s/\omega_{z2}$$



LHP: Left-hand plane, RHP: Right-hand plane

Uncompensated CMOS OPAMP

- Basic building blocks of an operational amplifier



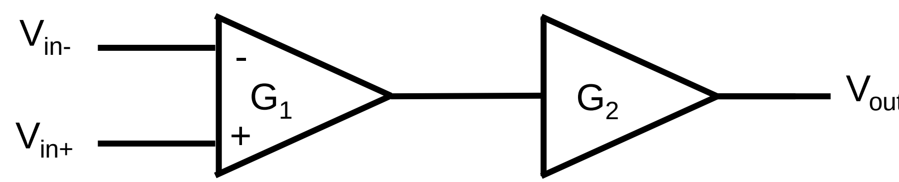
$$A_{V1} = -g_{m1}R_{o1} = -g_{m1}\left(r_{ds4} // r_{ds2}\right)$$

$$A_{V2} = -g_{m6}R_{o2} = -g_{m6}\left(r_{ds6} // r_{ds7}\right)$$

R_{o1} :Low freq. output impedance of A

R_{o2} :Low freq. output impedance of B

$C_A(C_B)$: Capacitive loading at A(B)

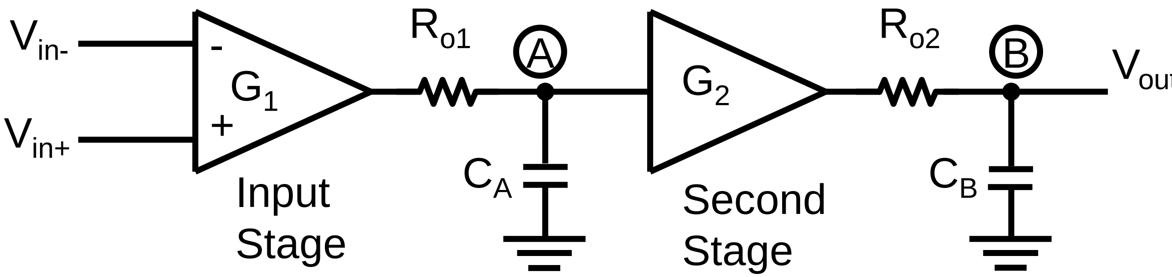


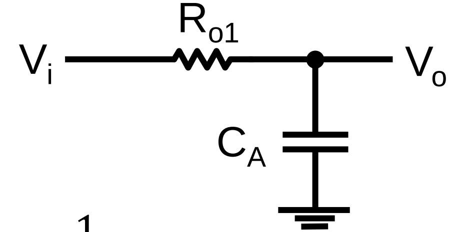
Differential stage
+
Differential-to-single-ended converter

Single-ended gain-stage

Uncompensated CMOS OPAMP (Cont.)

- Block diagram showing the origin of the dominant poles





$$\frac{V_o(s)}{V_i(s)} = \frac{\frac{1}{sC_A}}{R_{o1} + \frac{1}{sC_A}} = \frac{1}{1 + sR_{o1}C_A}$$

$$A_V(s) = \frac{V_{out}(s)}{V_{in}^+(s) - V_{in}^-(s)} = A_v(0)\frac{\frac{1}{sC_A}}{R_{o1} + \frac{1}{sC_A}}\frac{\frac{1}{sC_B}}{R_{o2} + \frac{1}{sC_B}} = A_v(0)\frac{1}{(1 + \frac{s}{S_A})(1 + \frac{s}{S_B})}$$

- ω_A and ω_B are dominant poles since R_{o1} and R_{o2} are normally large.

$$\omega_A = \frac{-1}{R_{o1}C_A} \qquad \omega_B = \frac{-1}{R_{o2}C_B}$$

- The effects of other poles are usually negligible.

Uncompensated CMOS OPAMP (Cont.)

$$A(s) = \frac{A_0}{(1+s/\omega_{P1})(1+s/\omega_{P2})}$$

$$A_0 = g_{m1} R_{01} g_{m6} R_{02}$$

$$g_{m1} = \sqrt{2\mu_p C_{OX} (W/L)_{M_1} I_1}$$

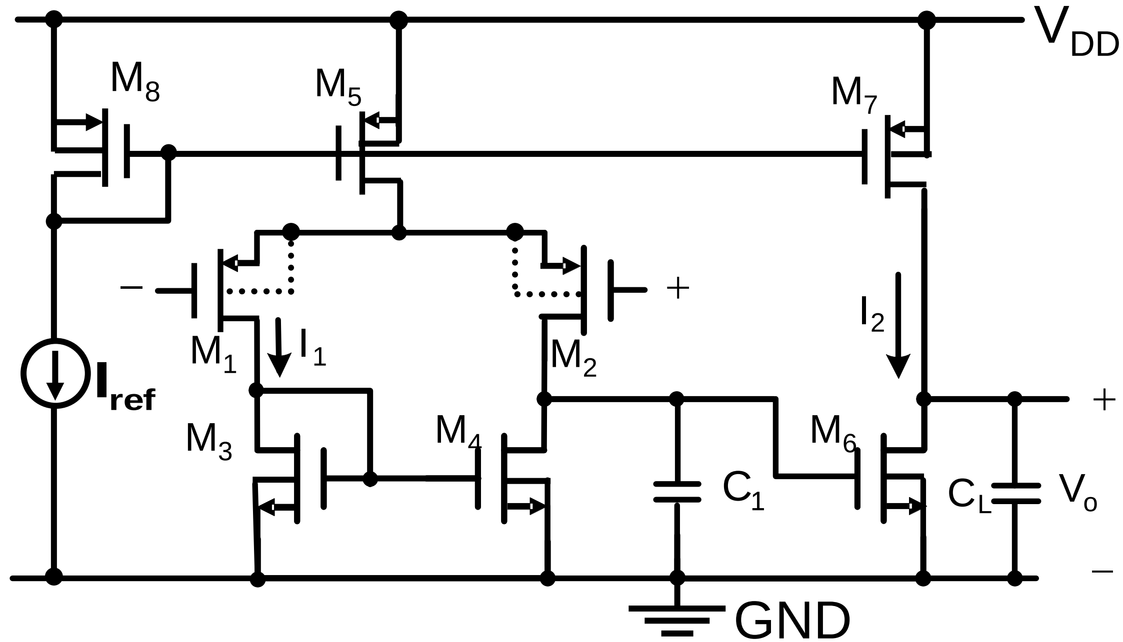
$$g_{m6} = \sqrt{2\mu_n C_{OX} (W/L)_{M_6} I_6}$$

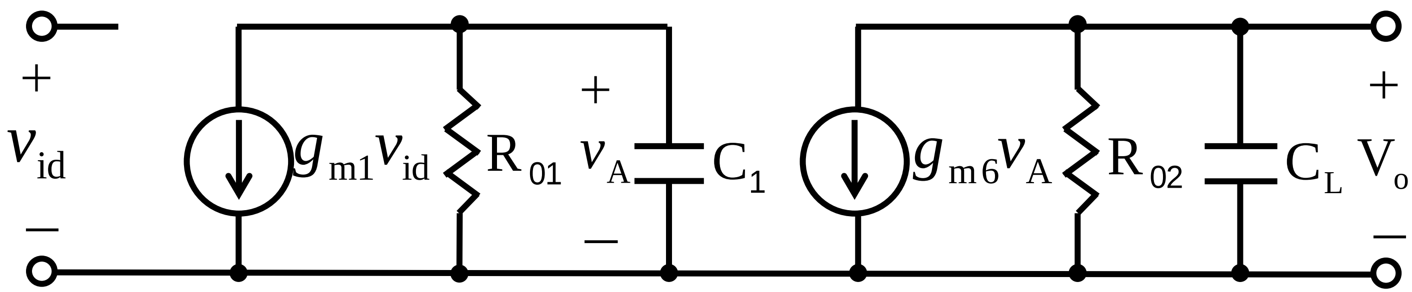
$$R_{01} = r_{ds2} // r_{ds4}$$

$$R_{02} = r_{ds6} // r_{ds7}$$

$$\omega_{P1} = -1/R_{01}C_1$$

$$\omega_{P2} = -1/R_{02}C_L$$





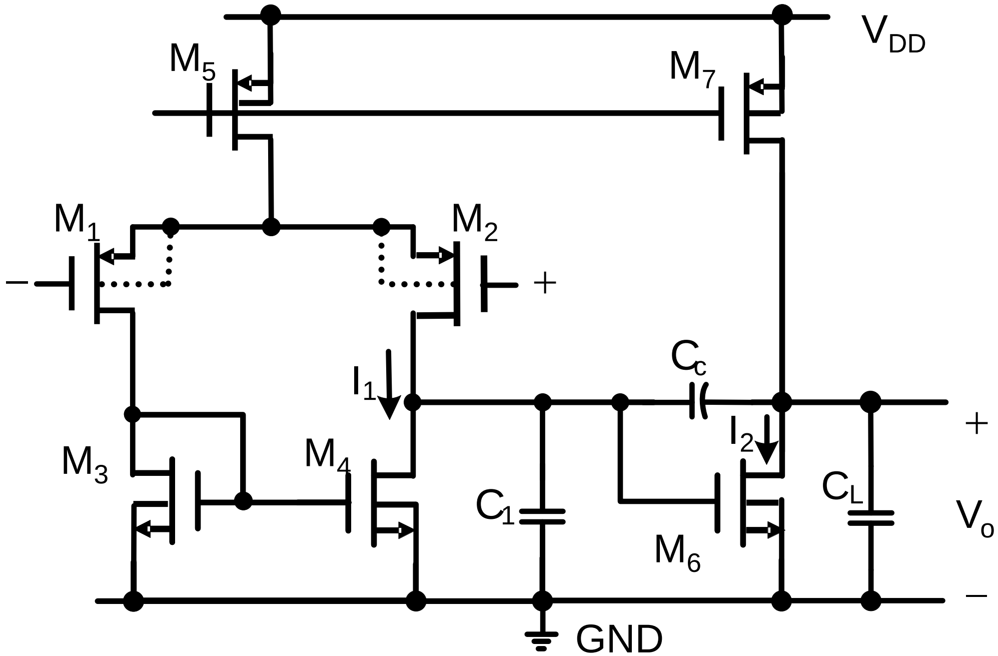
- P_1 & P_2 are dominant poles since R_{01} and R_{02} are normally large. The effects of other poles are usually negligible.

Uncompensated CMOS OPAMP (Cont.)

- For low frequency
  - $A(j\omega) = A(0) \approx A_0$
- For high frequency
  - $A(j\omega) \approx -\frac{g_{m1}g_{m6}}{\omega^2 C_1 C_L}$
  - The amplifier inverts the input voltage.
  - If feedback is used, then positive feedback occurs.
- Two dominant poles
  - Phase margin is not large enough
  - Pole-splitting technique to solve this problem

Pole-Splitting of Two-Stage CMOS OPAMP

- Reduce ω_{P1} and increase ω_{P2}



$$A(s) = \frac{A_0(1 - s/\omega_Z)}{(1 + s/\omega_{P1})(1 + s/\omega_{P2})}$$

$$A_0 = g_{m1}R_{01}g_{m6}R_{02}$$

$$g_{m1} = \sqrt{2\mu_p C_{OX}(W/L)_{M_1} I_1}$$

$$g_{m6} = \sqrt{2\mu_n C_{OX}(W/L)_{M_6} I_6}$$

$$R_{01} = r_{ds2} // r_{ds4}$$

$$R_{02} = r_{ds6} // r_{ds7}$$

$$\omega_Z \approx \frac{g_{m6}}{C_C}$$

If $g_{m6}R_{02} >> 1$

$$\omega_{P1} = \frac{-1}{(1 + g_{m6}R_{02})C_C R_{01}} \approx \frac{-g_{m1}}{A_0 C_C}$$

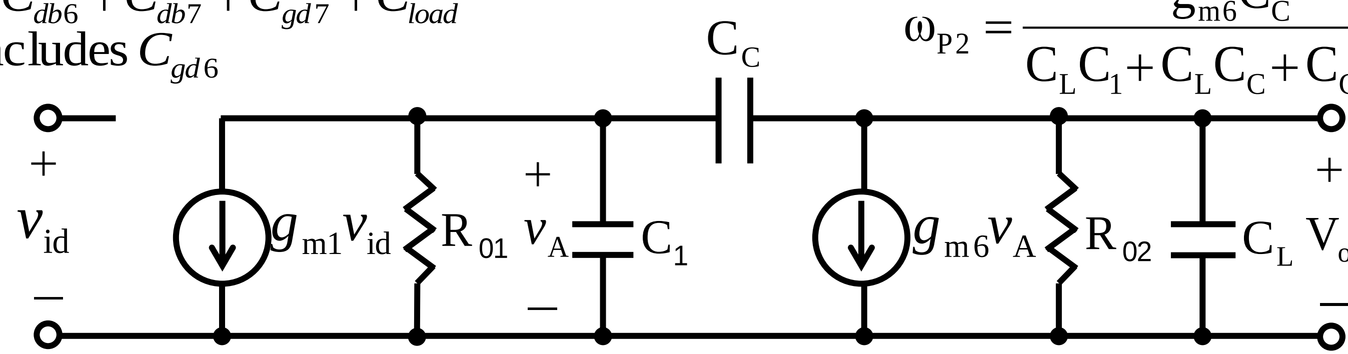
If C_C & $C_L >> C_1$

$$\omega_{P2} = \frac{-g_{m6}C_C}{C_L C_1 + C_L C_C + C_C C_1} \approx \frac{-g_{m6}}{C_L}$$

$$C_1 = C_{gd2} + C_{db2} + C_{gd4} + C_{db4} + C_{gs6}$$

$$C_L = C_{db6} + C_{db7} + C_{gd7} + C_{load}$$

C_C includes C_{gd6}



$\Rightarrow$ Right plane zero causes slower gain drop but quick phase drop

Pole-Splitting of Two-Stage CMOS OPAMP (Cont.)

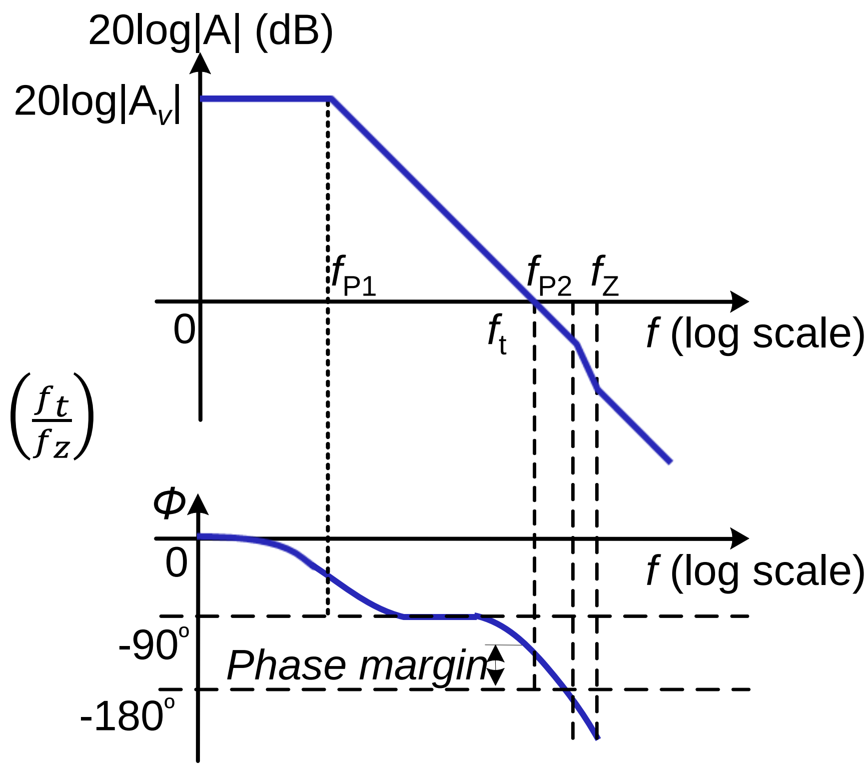
- Unity-gain frequency → $f_t \text{ (or } f_u) = |A_0| \frac{\omega_{P1}}{2\pi} = \frac{1}{2\pi} \frac{g_{m1}}{C_C}$
- To achieve an uniform -20dB/dec gain rolloff down to 0dB, the following two conditions must be satisfied
  - $f_t < f_{p2} \Longrightarrow \frac{g_{m1}}{C_C} < \frac{g_{m6}}{C_L}$
  - $f_t < f_z \Longrightarrow g_{m1} < g_{m6}$
- At unity-gain frequency f_t (or f_u)

$$\Phi_{total} = \tan^{-1}\left(\frac{f_t}{f_{p1}}\right) + \tan^{-1}\left(\frac{f_t}{f_{p2}}\right) + \tan^{-1}\left(\frac{f_t}{f_z}\right)$$

where $\tan^{-1}\left(\frac{f_t}{f_{p1}}\right) \cong 90°$

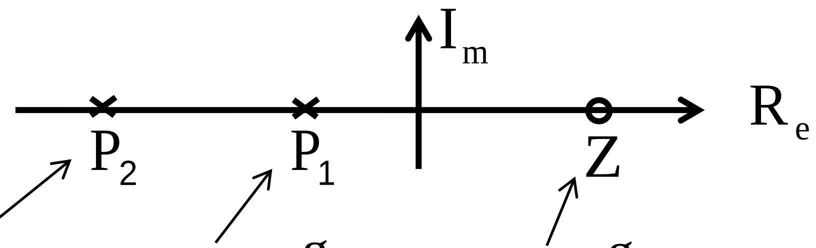
$$\text{Phase margin} = 180° - \Phi_{total}$$

$$= 90° - \tan^{-1}\left(\frac{f_t}{f_{p2}}\right) - \tan^{-1}\left(\frac{f_t}{f_z}\right)$$

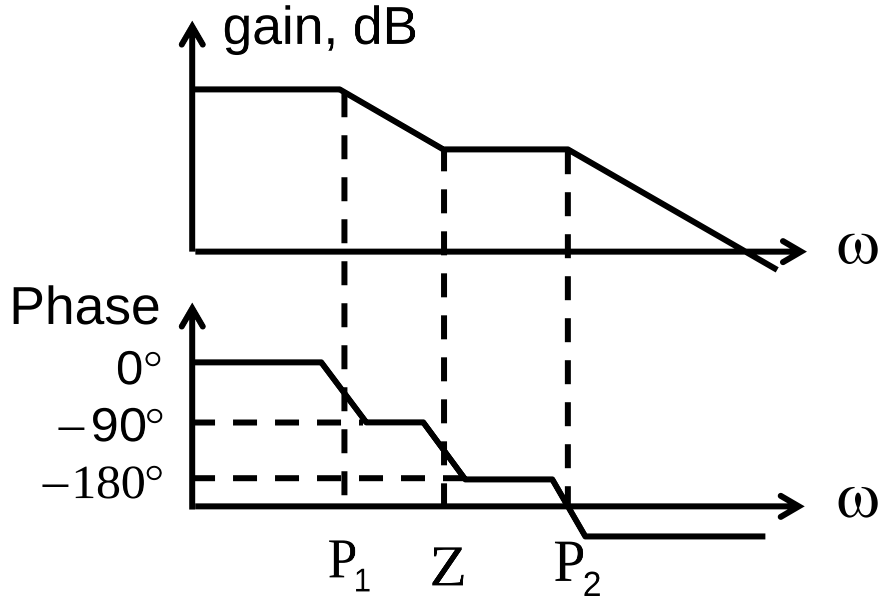


Right Plane Zero

- Cause slower gain drop but quick phase drop
  - Usually moved away if phase margin is not large enough

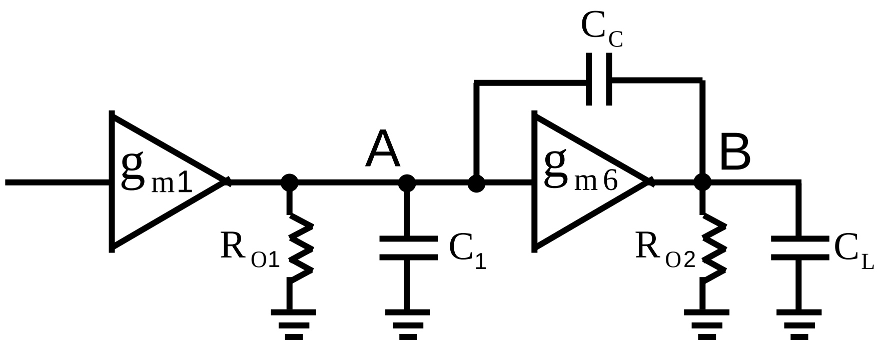


$$P_2 \approx -\frac{g_{m6}}{C_L} \qquad P_1 \approx -\frac{g_{m1}}{A_0 C_C} \qquad \omega_Z \approx \frac{g_{m6}}{C_C}$$



Right-Plane Zero (Cont.)

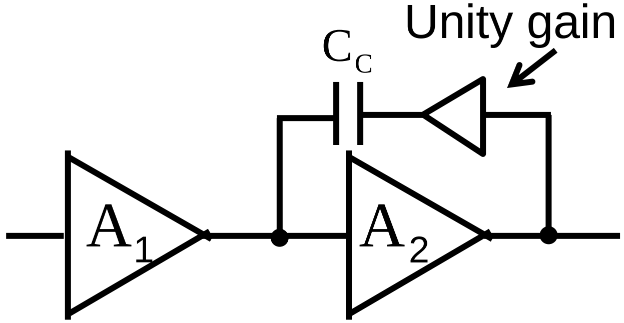
- The zero is due to the existence of two path through which the signal can propagate from node A to node B



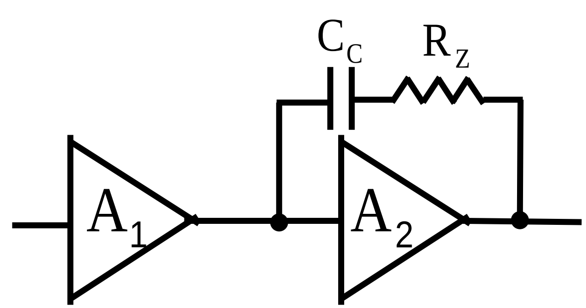
1. through C_C
2. through the controlled source $g_{m6}V_A$

- To eliminate zero ω_z

1. Method-1

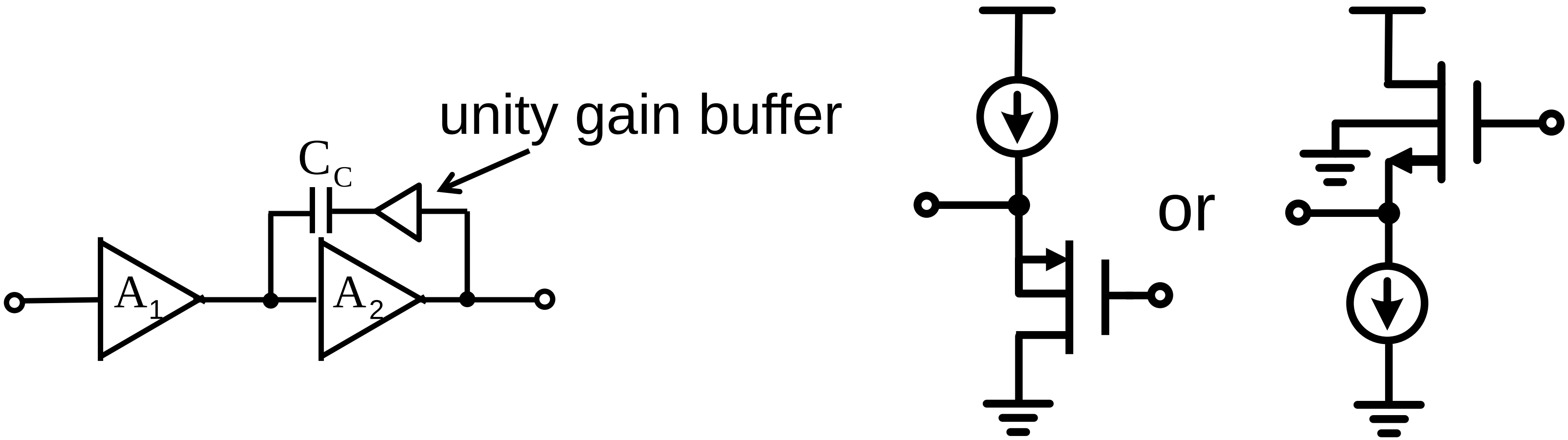


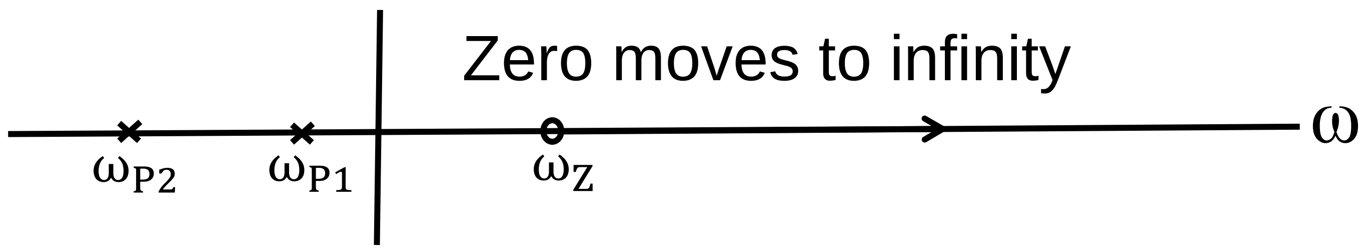
2. Method-2



Elimination of Right-Plane Zero

- Method-1: Use unity-gain buffer → Zero moves to infinity





$$A(s) = \frac{A_0}{(1+\frac{s}{\omega_{P1}})(1+\frac{s}{\omega_{P2}})} \quad \text{where} \quad \omega_{P1} \approx -\frac{g_{m1}}{A_0 C_C}, \ \omega_{P2} \approx \frac{-g_{m6}}{C_L}$$

Elimination of Right-Plane Zero (Cont.)

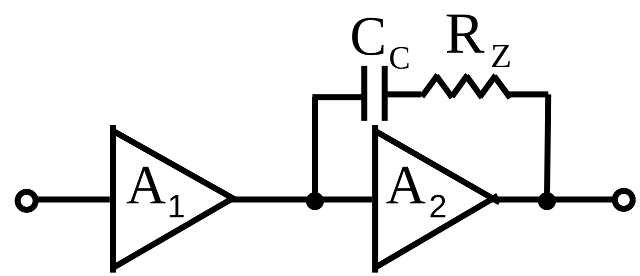
- Method-2: Using R instead of buffer
  - Elimination of zero → Let $R_Z = \frac{1}{g_{m6}}$
  - Pole-zero cancellation → Let $\omega_Z = \omega_{P2}$

$$\omega_{P1} \approx -\frac{g_{m1}}{A_0 C_C}$$

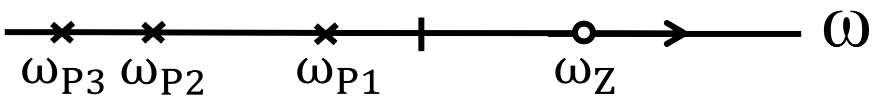
$$\omega_{P2} \approx -\frac{g_{m6}}{C_L}$$

$$\omega_{P3} \approx -\frac{1}{R_Z}\left(\frac{1}{C_C} + \frac{1}{C_1} + \frac{1}{C_L}\right)$$

$$\omega_Z = -\frac{1}{\left[R_Z - \left(\frac{1}{g_{m6}}\right)\right] C_C}$$

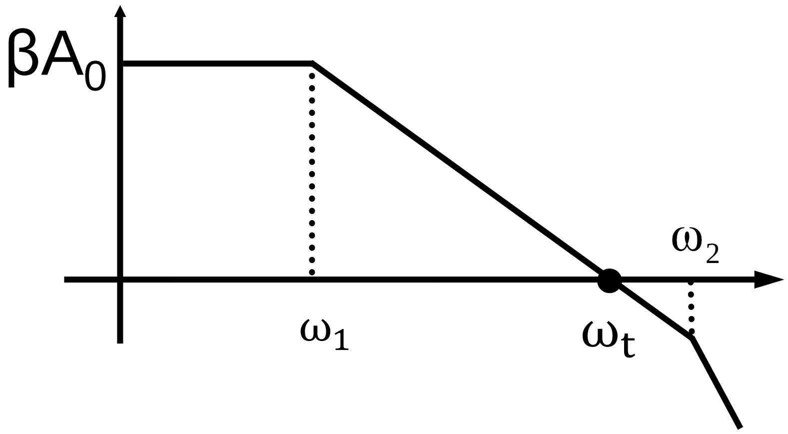


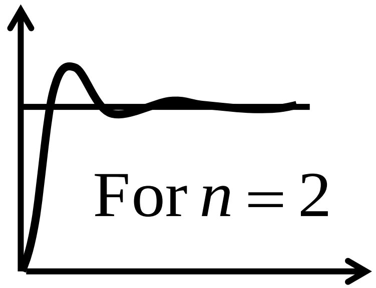
Zero moves toward the left plane as R_Z increases

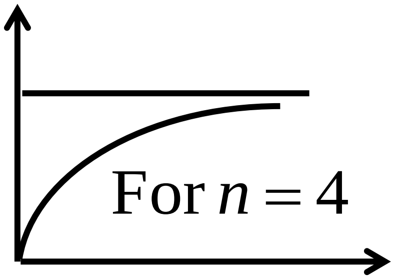


Pole Separation vs. Phase Margin and Speed

- $\omega_t = \frac{1}{n}\omega_2 = \beta A_0 \omega_1$
- Step response (**with fixed ω_2**)
  - n=2
    - Phase margin = 63º
    - Fast
  - n=3
    - Phase margin = 71º
  - n=4
    - Phase margin = 76º
    - Critically damped

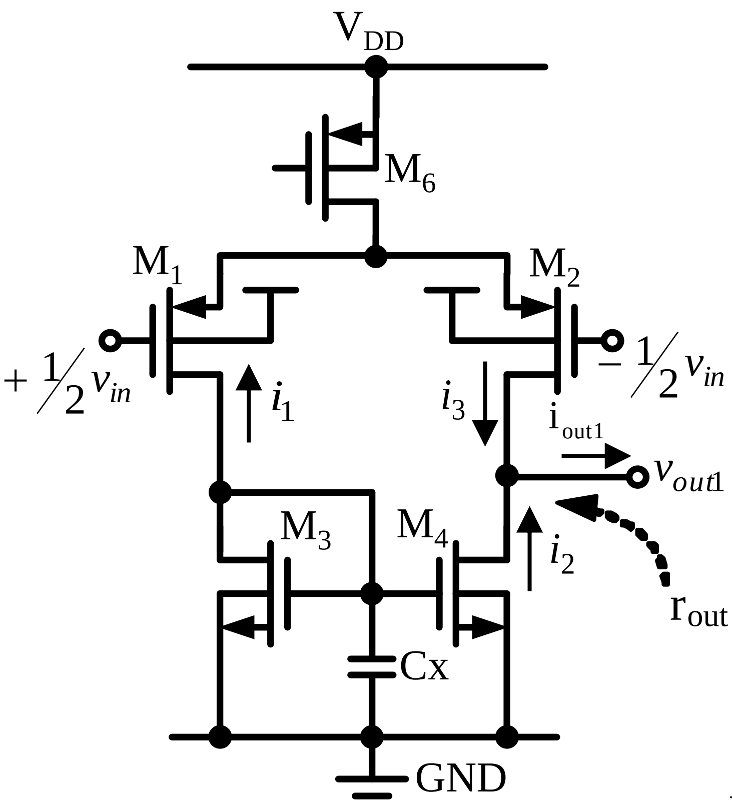






Pole-Zero Doublet

- First stage of a two-stage opamp

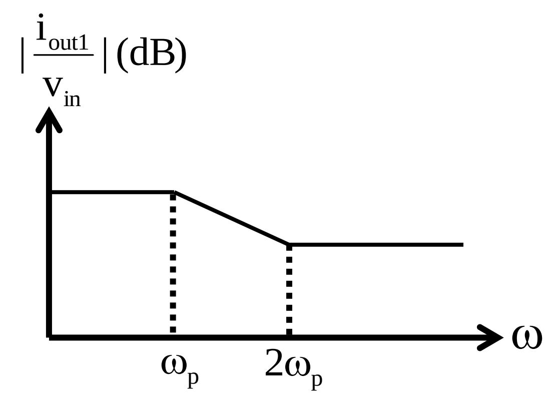


Assume $g_{m_1} = g_{m_2},\ g_{m_3} = g_{m_4}$

$$C_X = C_{gs3} + C_{gs4} + C_{db1} + C_{db3} + C_{gd1} + C_{gd4}$$

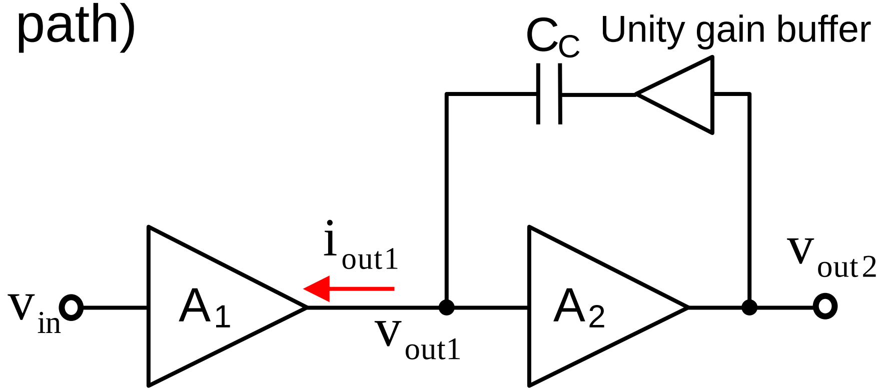
$$\begin{aligned}
i_{out1} &= (i_2 + i_3) \\
&= \left[\frac{1}{2}v_{in} \cdot g_{m1} \cdot \left(\frac{1}{g_{m3}} // \frac{1}{sC_X}\right) \cdot g_{m4} + \frac{1}{2}v_{in} \cdot g_{m2}\right] \\
&= \frac{1}{2} g_{m1} v_{in} \left(\frac{1}{1 + s/\omega_p} + 1\right) \qquad \text{where } \omega_p = \frac{g_{m3}}{C_X} \\
&= g_{m1} v_{in} \cdot \left(\frac{1 + {}^{s}/_{2\omega_p}}{1 + {}^{s}/_{\omega_p}}\right)
\end{aligned}$$

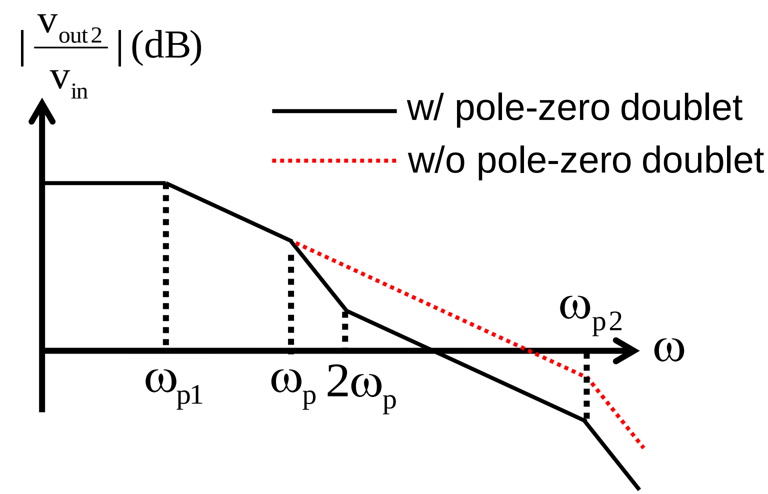
$$\frac{i_{out1}}{v_{in}} = g_{m1} \cdot \left(\frac{1 + {}^{s}/_{2\omega_p}}{1 + {}^{s}/_{\omega_p}}\right)$$



Pole-Zero Doublet (Cont.)

- Equivalent circuit of two-stage opamp (example: buffer in feedback path)





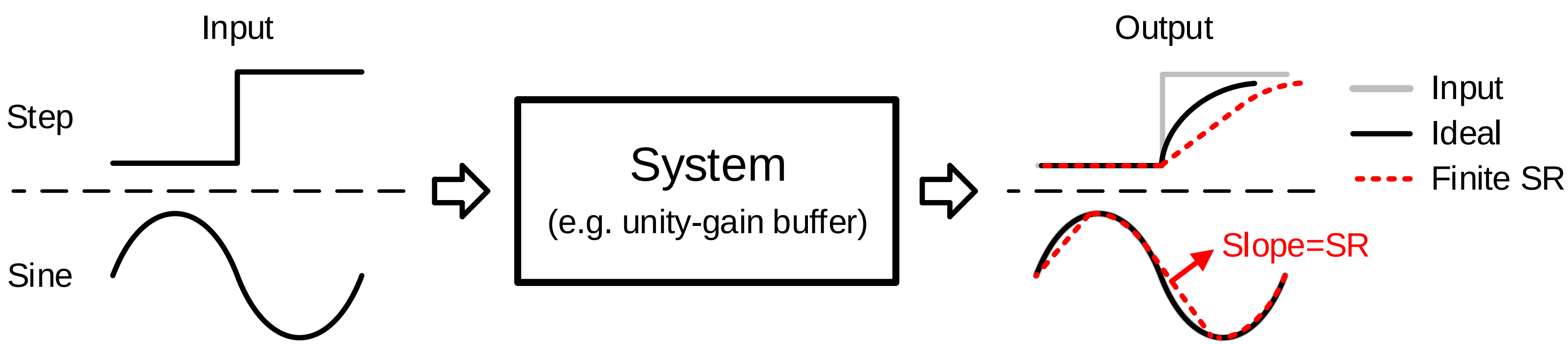
$$\Rightarrow \frac{i_{out1}}{v_{in}} = g_{m1}\cdot\left(\frac{1+\frac{s}{2\omega_p}}{1+\frac{s}{\omega_p}}\right),\ \text{and}\ \frac{v_{out2}}{i_{out1}} = \frac{R_{o1}g_{m5}R_{o2}}{\left(1+\frac{s}{\omega_{P1}}\right)\left(1+\frac{s}{\omega_{P2}}\right)}\ \ \text{where}\ \omega_{P1}\approx\frac{1}{R_{o1}g_{m5}R_{o2}Cc}=\frac{g_{m1}}{A_0Cc},\ \omega_{P2}\approx\frac{g_{m5}}{C_L}$$

$$\Rightarrow \frac{v_{out2}}{v_{in}} = \frac{i_{out1}}{v_{in}}\cdot\frac{v_{out2}}{i_{out1}} = \frac{g_{m1}R_{o1}g_{m5}R_{o2}}{\left(1+\frac{s}{\omega_{P1}}\right)\left(1+\frac{s}{\omega_{P2}}\right)}\cdot\left(\frac{1+\frac{s}{2\omega_p}}{1+\frac{s}{\omega_p}}\right) = A_0\frac{\left(1+\frac{s}{2\omega_p}\right)}{\left(1+\frac{s}{\omega_{P1}}\right)\left(1+\frac{s}{\omega_{P2}}\right)\left(1+\frac{s}{\omega_p}\right)}$$

- ◆ The parasitic capacitance C_x creates a pole at ω_p and a zero at $2\omega_p$.
- ◆ It may affect OPAMP stability if ω_p close to unity gain frequency.

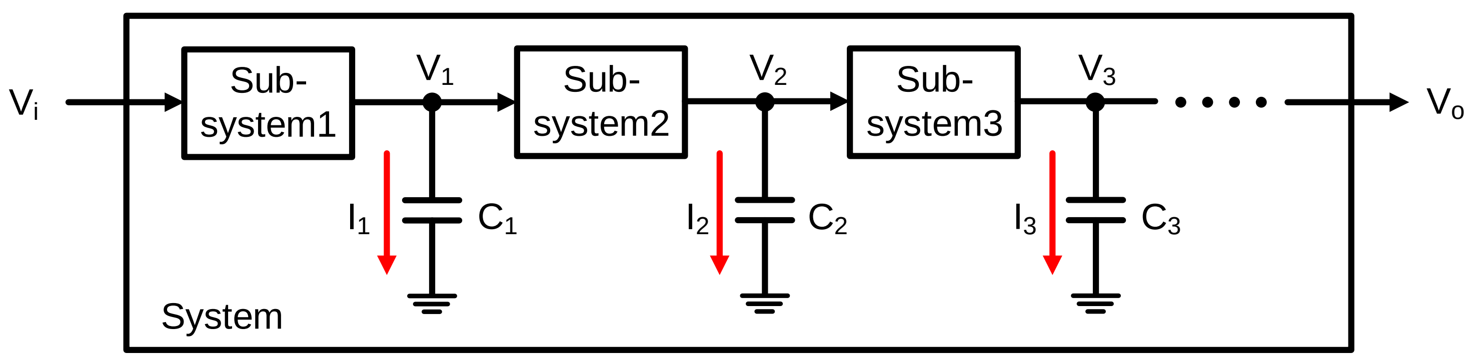
Introduction of Slew Rate (SR)

- Definition: Maximum change rate of voltage



- SR depends on system driving currents and capacitive loads
- SR should be considered at all nodes in a circuit, for example:

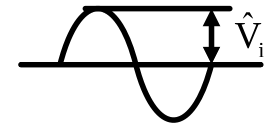
$$SR = \left.\frac{dv_o}{dt}\right|_{max} = \left.\frac{I_i}{C_i}\right|_{min} \quad i = 1,2,3....$$



SR Effect on Sinusoidal Response

- Voltage change rate without SR limitation

$$v_i = v_o = \hat{V}_i \sin\omega t \;\Rightarrow\; \frac{dv_o}{dt} = \omega \hat{V}_i \cos\omega t \;\Rightarrow\; \left.\frac{dv_o}{dt}\right|_{max} = \omega \hat{V}_i \cos 0 = \omega \hat{V}_i$$

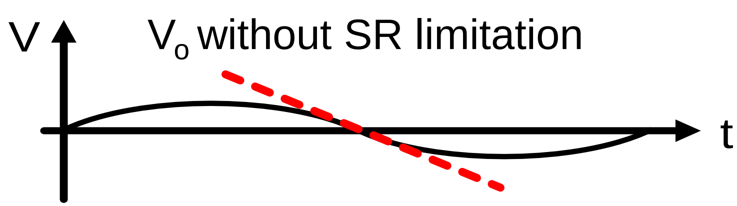


- Full-power bandwidth (f_M)

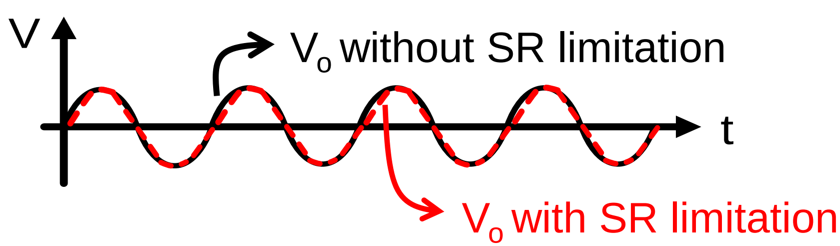
$$SR = \omega_M V_{o\_max} \Rightarrow f_M = \frac{SR}{2\pi V_{o\_max}}$$

V_{o_max}: rated opamp output voltage
ω_M : maximum input frequency without distortion

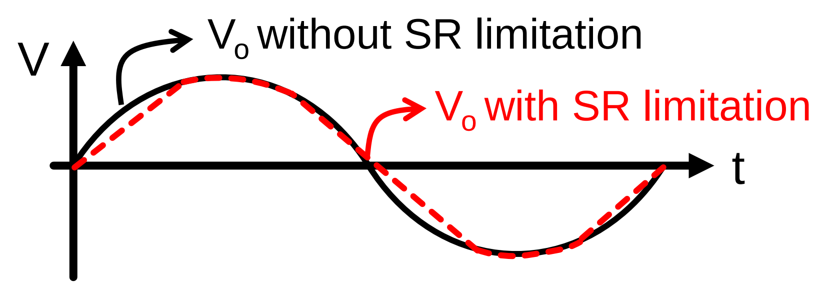
- SR effect on sine waves
  - Small amplitude, low freq.



  - Small amplitude, high freq.



  - Large amplitude, low freq.



SR limitation depends on amplitude and frequency

SR Effect on Step Response of a One-Pole System

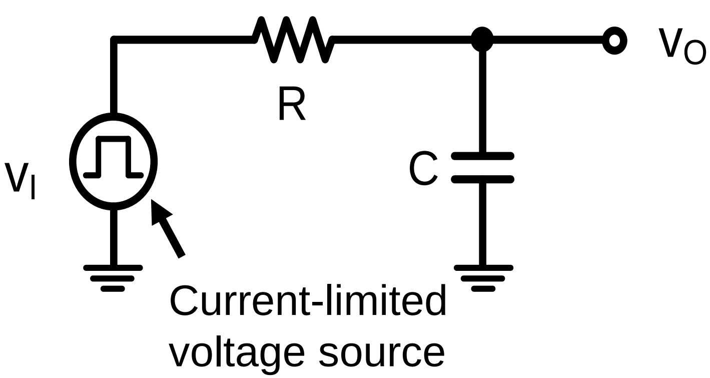
- Step response of a one-pole system
  - Ideal response: Exponential output $v_{O,Ideal}(t)$

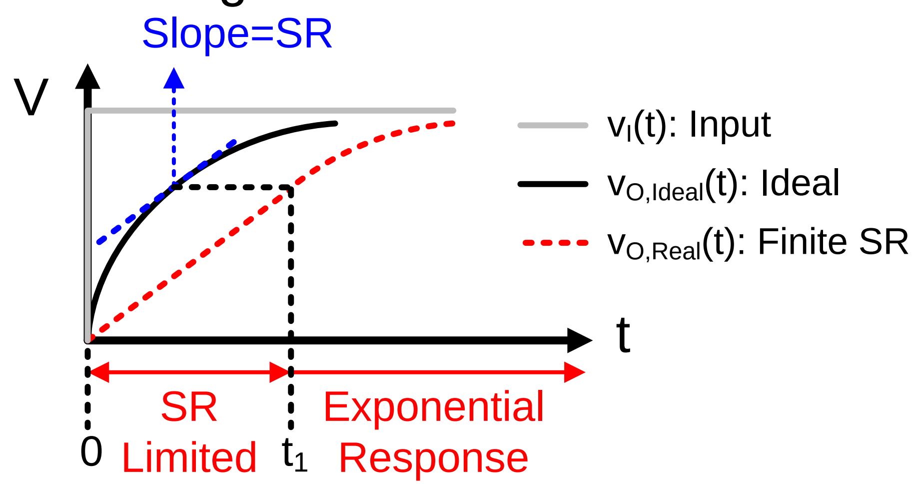
$$v_{O,Ideal}(t) = V_i\left(1 - e^{-t/\tau}\right) \Rightarrow \frac{d}{dt}\left(v_{O,Ideal}(t)\right) = \frac{V_i}{\tau} e^{-t/\tau}$$

  - Without large enough system SR → Slewing happens

$$\text{When SR} < \frac{d}{dt}\left(v_{O,Ideal}(t)\right) \Rightarrow \frac{d}{dt}\left(v_{O,Real}(t)\right) = \text{SR (As } 0\sim t_1 \text{ in the waveform below)}$$

- Example: RC filter with current-limited voltage source



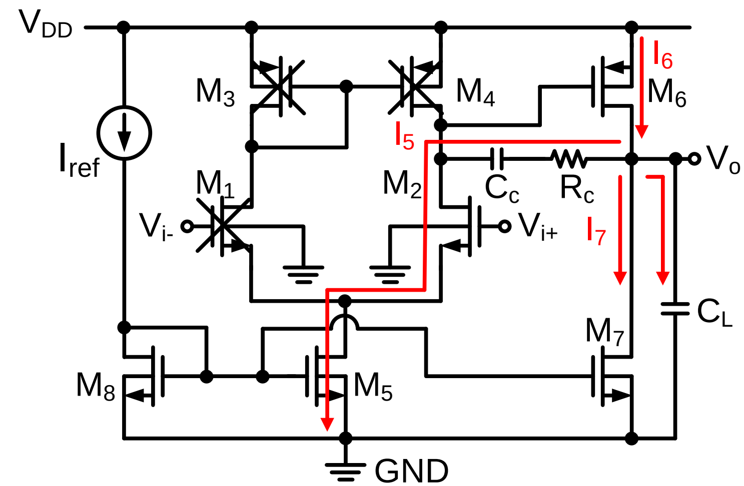


- Please refer to P.5-62~P.5-66 for more detailed description.

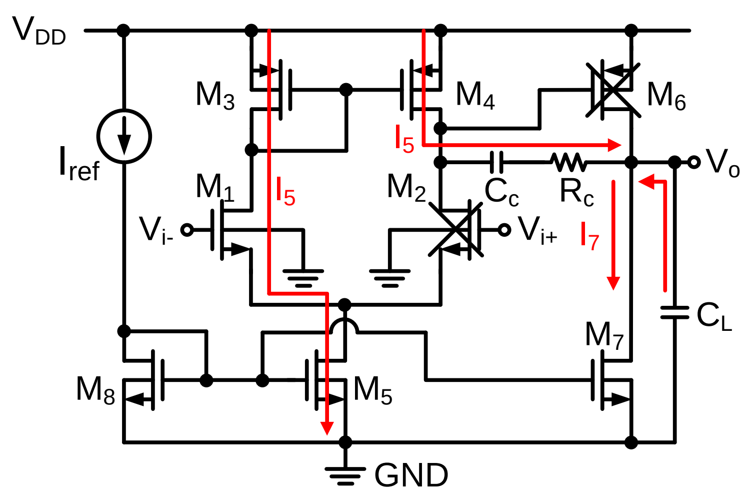
SR Analysis of Two-Stage OPAMP

- V_o rising process
  - Large positive input at V_{i+}
  - M_1 turned off
  - I_5 flows through C_c
  - Driving capability of I_6 is usually large
    - For SR not limited by I_6 , SR = I_5 / C_c

      (For small I_6 , $SR = \frac{I_6 - I_7}{C_C + C_L}$)



- V_o falling process
  - Large positive input at V_{i-}
  - M_2 turned off
  - I_5 flow through C_c
    - I_7 large enough: SR= I_5 / C_c
    - I_7 not large enough: SR= I_7 / (C_c+C_L)



- SR when I_7 is large enough

$$g_{m1} = \sqrt{2\mu C_{ox} \frac{W_{M1}}{L_{M1}} \frac{I_5}{2}}, \quad \omega_t = \frac{g_{m1}}{C_C} \Rightarrow SR = \left.\frac{dv_o(t)}{dt}\right|_{max} = \frac{I_5}{C_C} = \omega_t \sqrt{\frac{I_5 L_{M1}}{\mu_n C_{ox} W_{M1}}} = (V_{GS1} - V_t)\,\omega_t$$

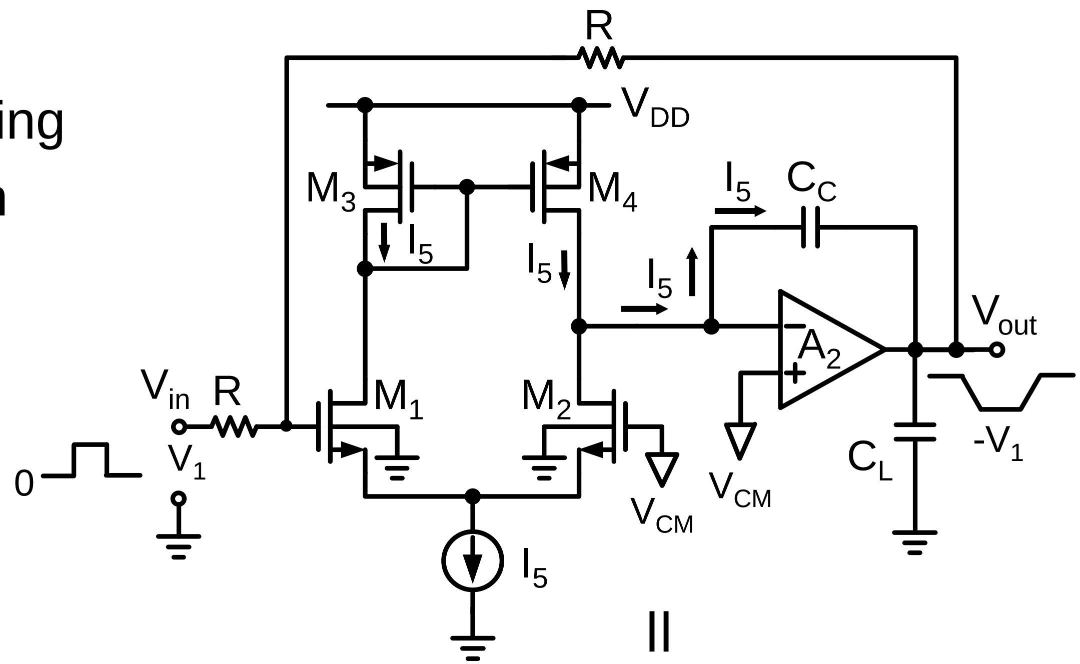
Example - Negative Feedback Amplifier

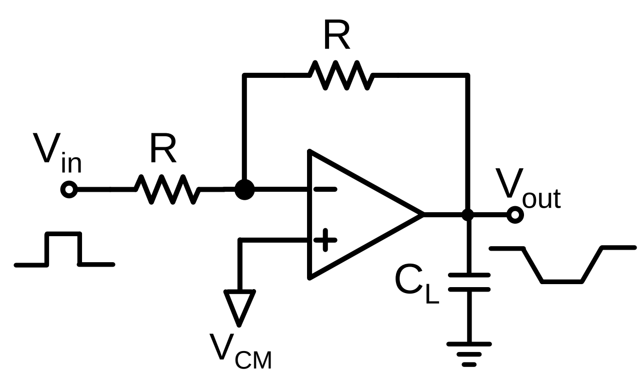
- Slew rate
  - Assume the output driving current is large enough

$$SR = \left|\frac{dV_{out}}{dt}\right| = \left|-\frac{1}{C_C}\frac{dQ_c}{dt}\right| = \frac{I_5}{C_C}$$

$$\text{For}\begin{cases}\omega_t = \dfrac{g_{m1}}{C_C} \Rightarrow C_C = \dfrac{g_{m1}}{\omega_t} \\ g_{m1} = \sqrt{2\mu_n C_{OX}\left(\dfrac{W}{L}\right)_{M1}\dfrac{I_5}{2}}\end{cases}$$

$$\Rightarrow SR = \frac{I_5\omega_t}{g_{m1}} = \omega_t\sqrt{\frac{I_5}{\mu_n C_{OX}\left(\frac{W}{L}\right)_{M1}}}$$





- Slew rate can be increased by
  - Increasing the unity-gain bandwidth
  - Increasing bias current of input stage
  - Decreasing the W/L ratio of the input transistors

Example 2 - Voltage Follower (1/2)

- V_{out} at large positive input ($t_1 \sim t_2$)

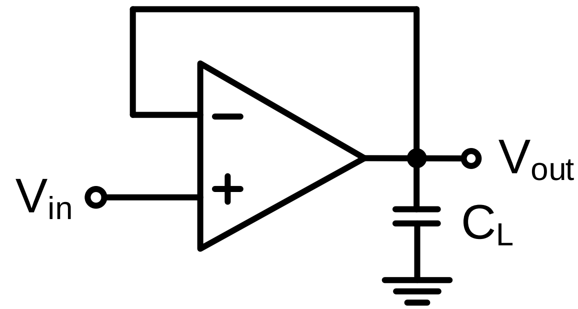
$$V_{in}(t) = V_1 u(t)$$

$$i_W(t) = C_W \frac{dV_W(t)}{dt} \approx C_W \frac{dV_{in}(t)}{dt} = C_W V_1 \delta(t)$$

($dV_W = dV_{in}$ due to source follower)

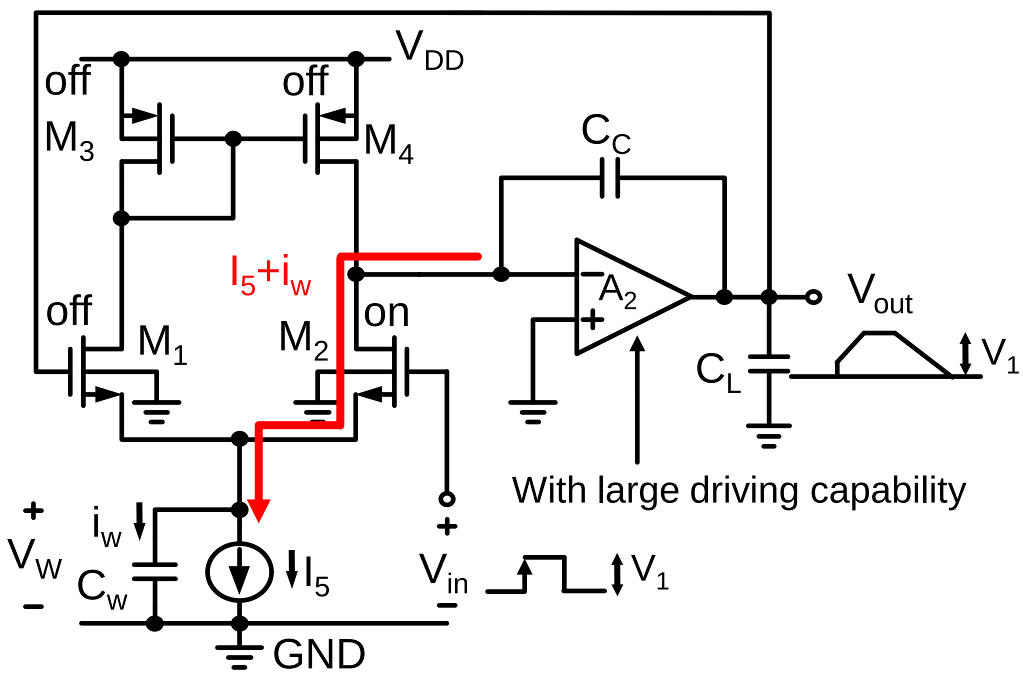
$$V_{out}(t) = \frac{1}{C_C}\int_0^t (I_5 + i_W)\,dt = \frac{I_5 t}{C_C} + \frac{C_W}{C_C}\int_0^t \frac{dV_{in}}{dt}\,dt = \frac{I_5}{C_C}t + \frac{C_W}{C_C}V_1 u(t)$$

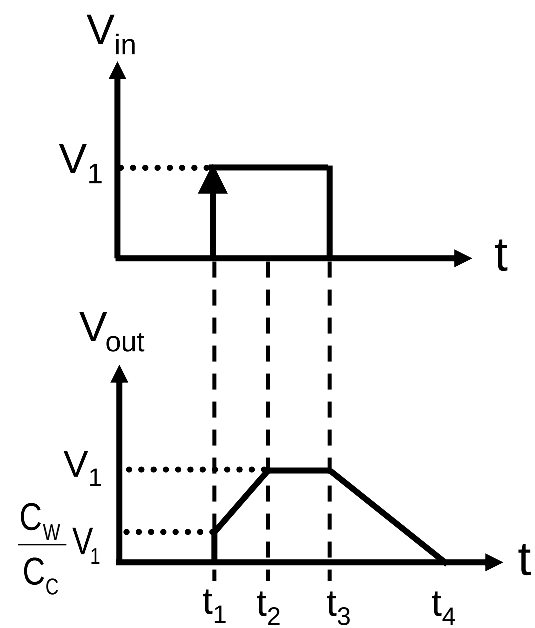
- Circuit diagram



Voltage follower

(Assume the output driving current is large enough)





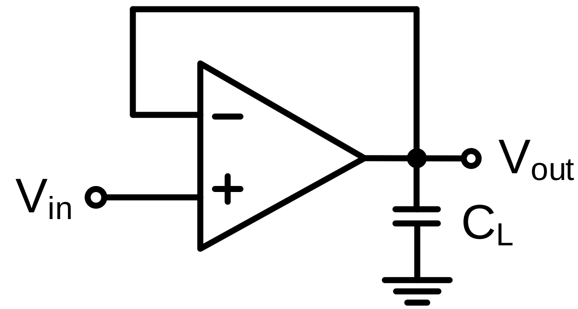
Example 2 - Voltage Follower (2/2)

- V_{out} at large negative input ($t_3 \sim t_4$)

As a source follower, V_W follows $V_{out} \Rightarrow \frac{dV_{out}}{dt} \approx \frac{dV_W}{dt}$

$$\frac{dV_{out}}{dt} = -\frac{I_5 - i_W}{C_C} = -\frac{i_W}{C_W}\left(=\frac{dV_W}{dt}\right)$$

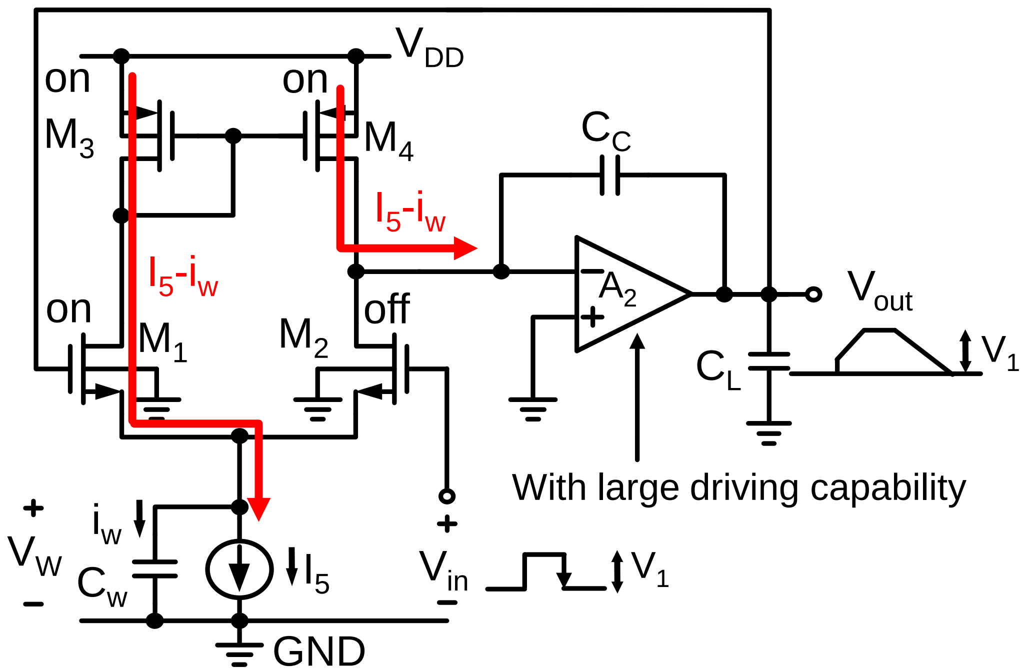
$$\frac{dV_{out}}{dt} = -\frac{I_5}{C_C + C_W} \rightarrow \text{SR reduced by } C_W \text{, from } \frac{I_5}{C_C} \text{ to } \frac{I_5}{C_C + C_W}$$

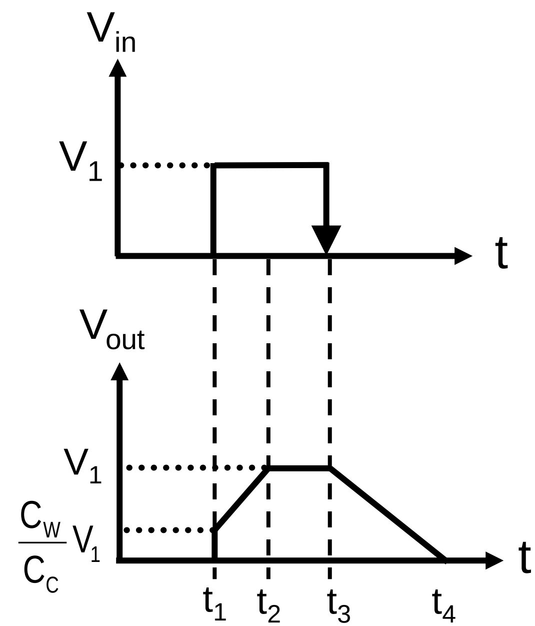


Voltage follower

(Assume the output driving current is large enough)

- Circuit diagram





◆ $dV_w = dV_{in}$ (Due to source follower)

Power-Supply Rejection Ratio (PSRR)

- Mixed-signal circuits combine analog and digital circuits
  - ◆ Switching activity in digital portion results in supply ripple
  - ◆ Add large capacitors between supply rails and ground
    - ➔ Not practical in IC design
  - ➔ High-PSRR analog circuits

- Definition

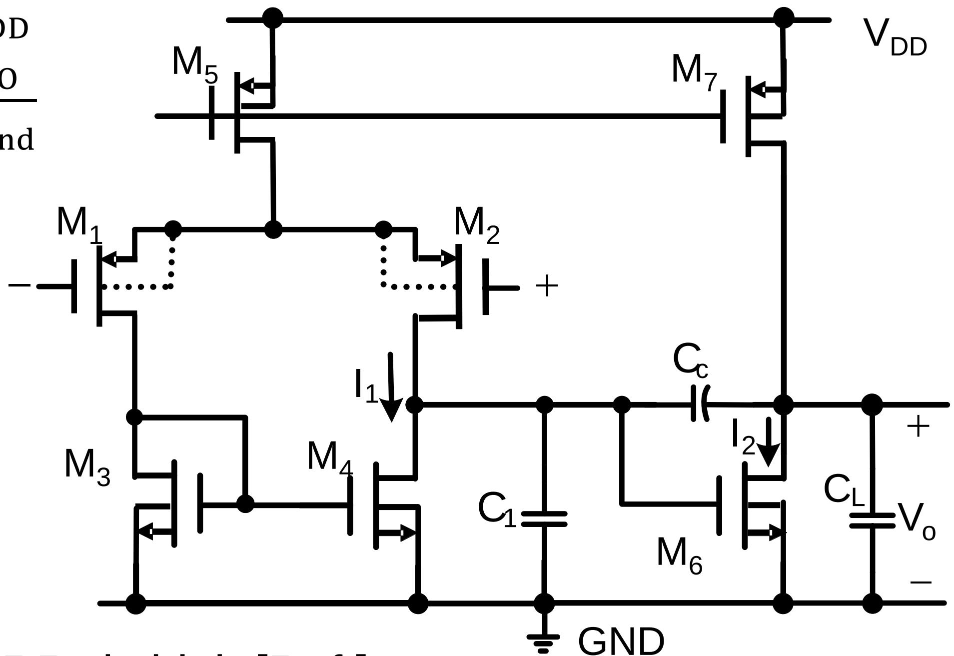
$$\begin{cases} PSRR^{+} \equiv \dfrac{A_d}{A^{+}} \\ PSRR^{-} \equiv \dfrac{A_d}{A^{-}} \end{cases}, \text{where} \begin{cases} A^{+} \equiv \dfrac{V_O}{V_{DD}} \\ A^{-} \equiv \dfrac{V_O}{V_{gnd}} \end{cases}$$

  - ◆ For a two-stage op amp

$$V_o = V_{gnd} \times \frac{r_{o7}}{r_{o6} + r_{o7}}$$

$$\Rightarrow A^{-} \equiv \frac{V_O}{V_{gnd}} = \frac{r_{o7}}{r_{o6} + r_{o7}}$$

$$\Rightarrow PSRR^{-} \equiv \frac{A_d}{A^{-}} = g_{m1}(r_{o2} \parallel r_{o4}) g_{m6} r_{o6}$$



    - ➢ It's insensitive to V_{DD} → $PSRR^{+}$ is high [Ref.]

[Ref.] P. R. Gray, P. J. Hurst, A. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th ed., New York: John Wiley & Sons, 2009. pp. 430–432

Design Trade-offs

- To increase the differential gain, CMRR, and PSRR for a two-stage op amp
  - Enlarge the length L for the channel of each MOSFET
  - Lower the $|V_{OV}|$ at which each MOSFET is operated

$$A_O = g_m r_{ds} \propto \sqrt{I} \cdot \frac{1}{\lambda I} = \frac{1}{\lambda \sqrt{I}} \propto \frac{L}{\sqrt{I}}, \text{where} \frac{1}{\lambda} \propto L \text{ (roughly)}$$

$$\omega_t = \frac{g_m}{C} \propto \frac{\sqrt{I}}{C} \propto \sqrt{I}$$

- The transition frequency of the MOSFETs can be increased by using a shorter channel and/or a larger $|V_{OV}|$

$$f_t = \frac{g_m}{2\pi(C_{gs} + C_{gd})} = \frac{2 \cdot \frac{1}{2}\mu C_{OX}\frac{W}{L}V_{OV}^2/V_{OV}}{2\pi(\frac{2}{3}WLC_{OX})} \approx \frac{1.5 \cdot \mu \cdot |V_{OV}|}{2\pi L^2}; \text{where} \begin{cases} \mu\text{: carrier mobility} \\ V_{OV}\text{: overdrive voltage} \\ L\text{: channel length} \end{cases}$$

$$g_m = \frac{2I}{V_{OV}}, C_{gs} \approx \frac{2}{3}WLC_{OX} \text{ and } C_{gs} \gg C_{gd}$$

- In conclusion, it's a trade-off between low-frequency performance parameters and high-frequency ones

  → For analog circuits in submicron process operated at 1V~1.5V supplies, 0.1V~0.3V of $|V_{OV}|$ is typically used, and the typical channel lengths are usually at least 1.5~2 times the L_{min}

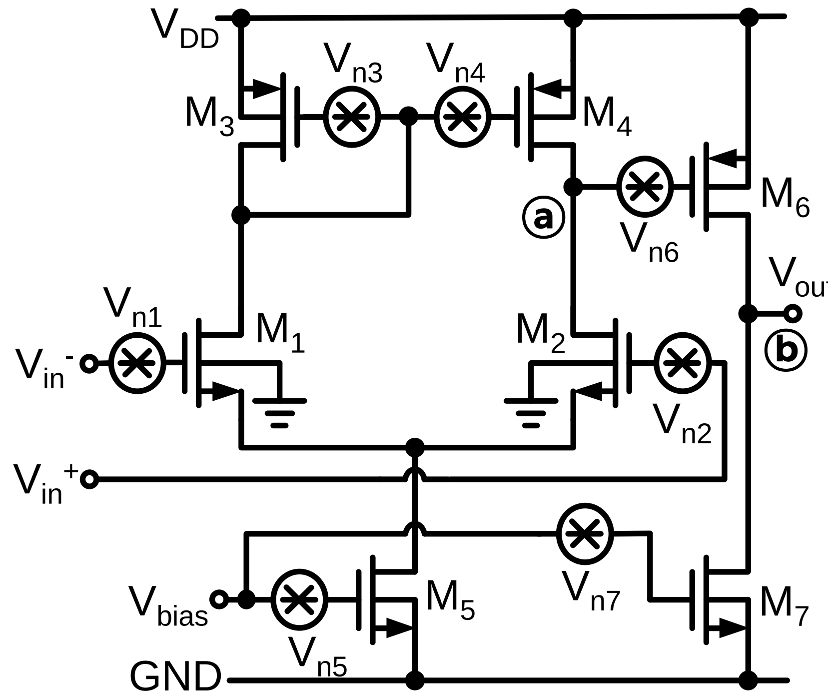
Noise Performance of CMOS OPAMP

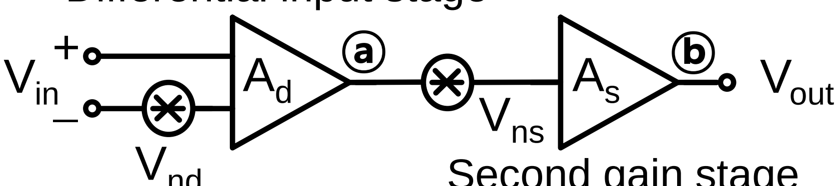
- Noise is fundamental limitation of OPAMP performance. The equivalent noise voltage of MOS OPAMPs may be 10 times larger than that of a comparable bipolar amplifier
- Example
  - Mean-square value at ⓐ

$$\begin{cases}\text{For } V_{n1} \ \& \ V_{n2,} \ A_d = g_{m1}(r_{ds2}//r_{ds4}) \\ \text{For } V_{n3} \ \& \ V_{n4,} \ A_v = g_{m3}(r_{ds2}//r_{ds4})\end{cases}$$

$$\overline{V_A^2} = A_d^2(\overline{V_{n1}^2} + \overline{V_{n2}^2}) + A_v^2(\overline{V_{n3}^2} + \overline{V_{n4}^2})$$

  - Equivalent input noise voltage



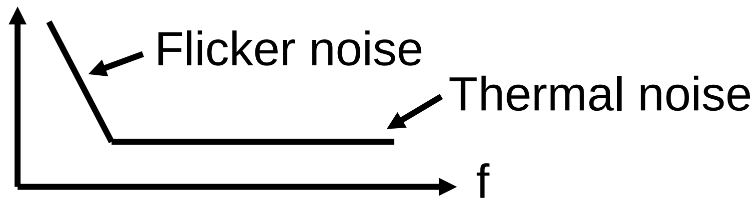


$$\overline{V_{nd}^2} = \frac{\overline{V_A^2}}{A_d^2} = \overline{V_{n1}^2} + \overline{V_{n2}^2} + \left(\frac{g_{m3}}{g_{m1}}\right)^2(\overline{V_{n3}^2} + \overline{V_{n4}^2})$$

$$\overline{V_n^2} = \overline{V_{nd}^2} + \frac{\overline{V_{ns}^2}}{A_d^2} = \overline{V_{n1}^2} + \overline{V_{n2}^2} + (\frac{g_{m3}}{g_{m1}})^2(\overline{V_{n3}^2} + \overline{V_{n4}^2}) + \frac{\overline{V_{n6}^2} + (\frac{g_{m7}}{g_{m6}})^2\overline{V_{n7}^2}}{A_d^2}$$

Noise Performance of CMOS OPAMP (Cont.)

- Device noise



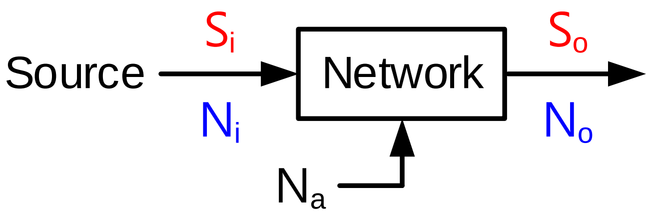
  - 1/f noise component dominates at low frequencies
  - The equivalent input noise voltage is greatest at low frequencies (below 1kHz)

- If $|A_d(\omega)|>>1$, then the input devices M_1 and M_2 tend to be the dominant noise sources and their optimization is the key to low-noise design.

- 1/f noise of OPAMP can be cancelled using
  - Chopper-stabilized technique
    - Dynamic Range over 100dB can be obtained
  - Correlated double sampling (CDS)

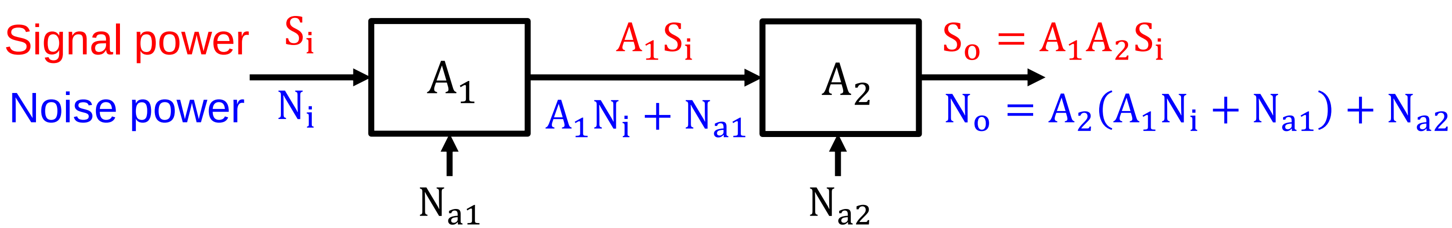
Noise Figure

- Definition of noise figure (NF)
  - The ratio of the signal-to-noise power ratio at the input to the signal-to-noise power ratio at the output



S_i : Signal power at the input
S_o : Signal power at the output
N_i : Noise power at the input
N_o: Noise power at the output
N_a: Additional noise power from the network

  - Example

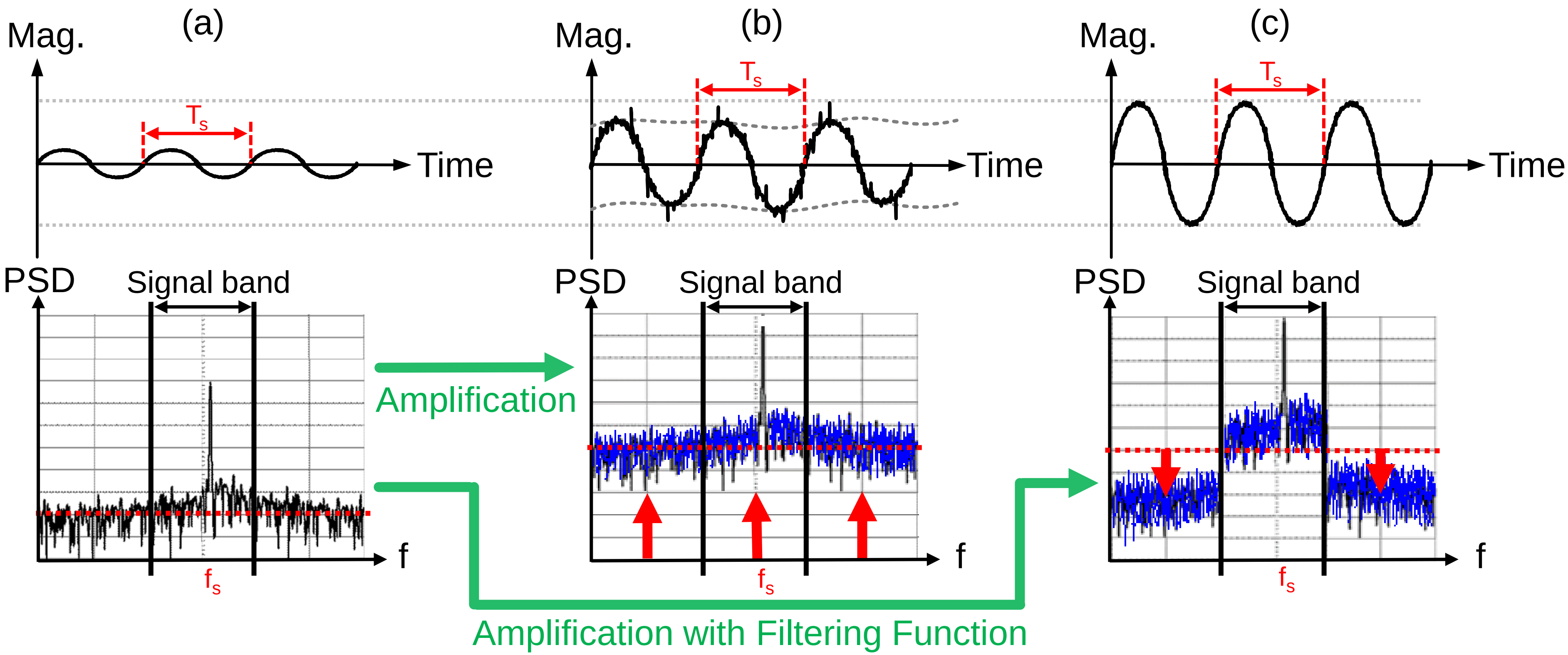


$$\text{Noise Figure (NF)} = \frac{S_i/N_i}{S_o/N_o} = \frac{S_i/N_i}{A_1A_2S_i/[A_2(A_1N_i + N_{a1}) + N_{a2}]} = \frac{N_i + \frac{N_{a1}}{A_1} + \frac{N_{a2}}{A_1A_2}}{N_i}$$

- Ideal NF = 1
- Large NF caused by noisy network
  - Low-noise amplifier (LNA): Higher A_1 with lower N_{a1} → Smaller NF

Noise Figure of Amplifier with Filtering Function

- Illustration in time domain and frequency domain



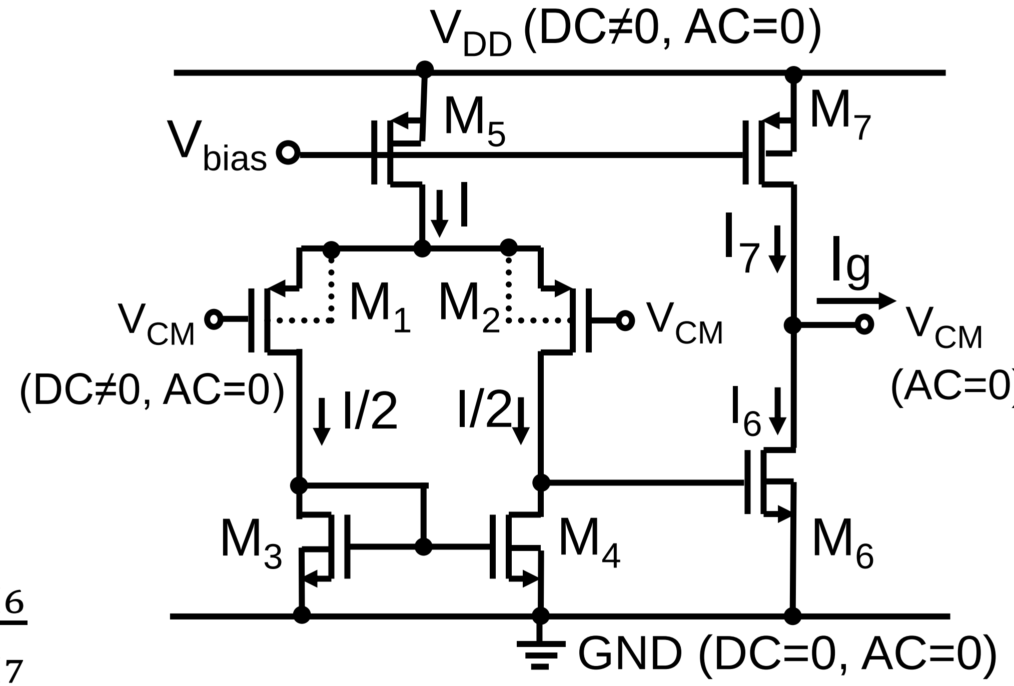
- Signal band: $\mathrm{NF}_{(a)} < \mathrm{NF}_{(b)}$ & $\mathrm{NF}_{(c)}$
- Function of filter: Attenuate out-of-band power
  - Allow larger in-band signal
  - Reduce signal slew rate
  - Avoid interference

Mag. : Magnitude
PSD : Power spectral density
f : Frequency
$f_s = 1/T_s$

Offset Voltage of Two-Stage CMOS OPAMPs

- Input voltage need to restore the output to zero

- Two components
  - Systematic offset
  - Random offset



- To avoid systematic offset, design must follow the rule

$$\frac{(W/L)_{M3}}{(W/L)_{M5}} = \frac{(W/L)_{M4}}{(W/L)_{M5}} = \frac{1}{2}\frac{(W/L)_{M6}}{(W/L)_{M7}}$$

- To minimize random offset
  - $L_1=L_2$, $W_1=W_2$, $L_3=L_4$, $W_3=W_4$, $L_3=L_6$ and $L_5=L_7$ to minimize the offsets of channel length and channel width variations
  - Large L and W such that $\Delta L/L$ and $\Delta W/W$ can be ignored

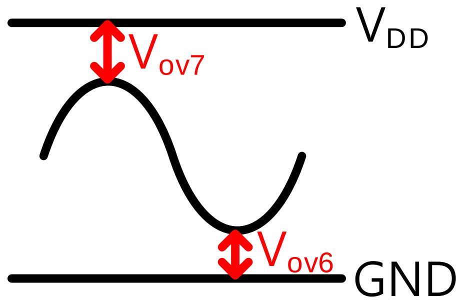
Input Common-Mode Range and Output Swing of Two-Stage CMOS OPAMP

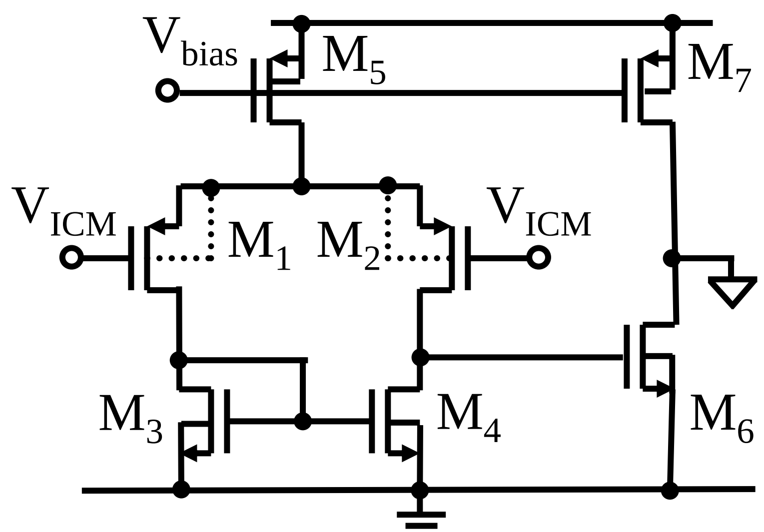
- Input common-mode range, V_{ICM}
  - Minimum V_{ICM}
    - Keep M_1 and M_2 in saturation → $V_{dg1,2} < |V_{tp}|$
    - Hence, $V_{ICM} \geq V_{tn} + V_{OV3} - |V_{tp}|$, where V_{ov} is overdrive voltage
  - Maximum V_{ICM}
    - Keep M_5 in saturation, $V_{ds5} > V_{ov5}$
    - Hence, $V_{ICM} \leq V_{DD} - |V_{OV5}| - |V_{tp}| - |V_{OV1}|$

  → $V_{OV3} + V_{tn} - |V_{tp}| \leq V_{ICM} \leq V_{DD} - |V_{tp}| - |V_{OV1}| - |V_{OV5}|$

- Output swing, V_O
  - Keep M_6 and M_7 in saturation

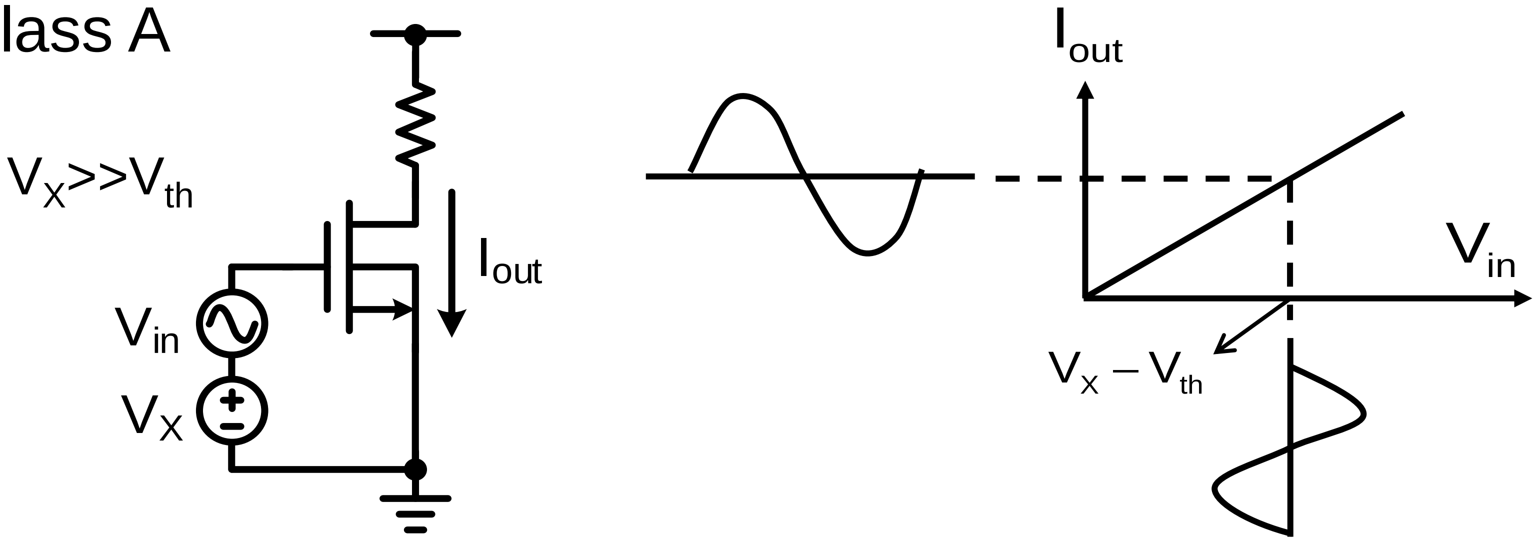
    $V_{OV6} \leq V_o \leq V_{DD} - |V_{OV7}|$



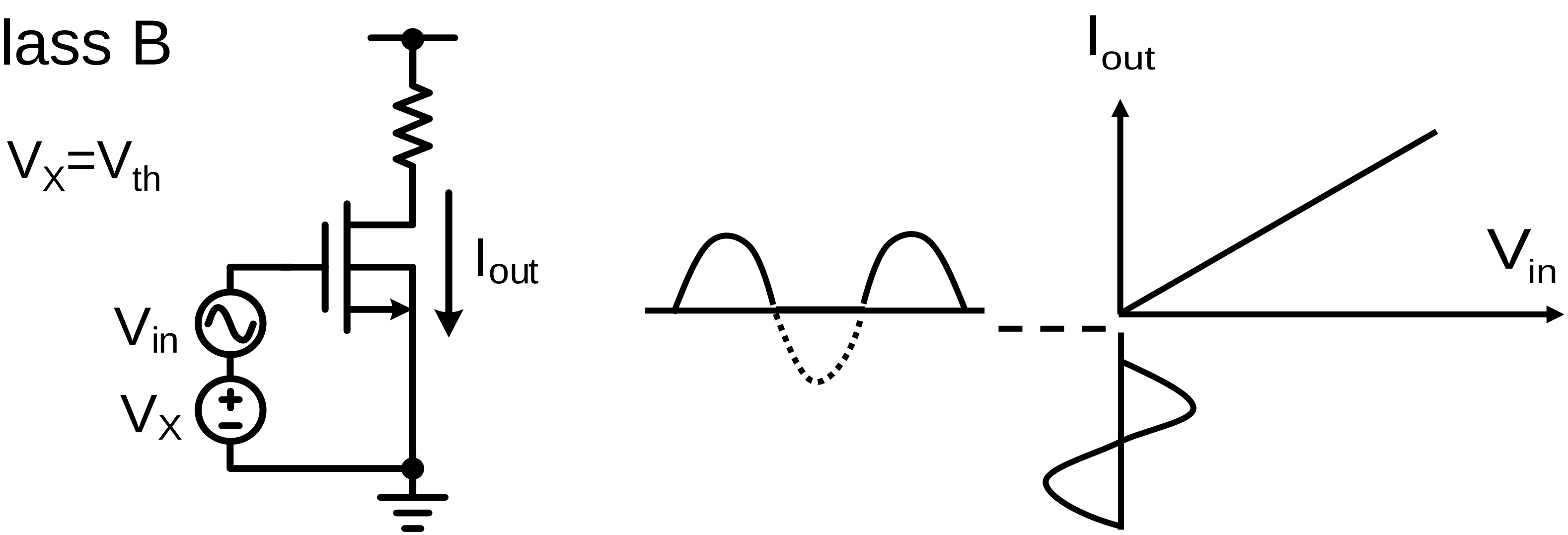


Amplifier Classification

- Class A

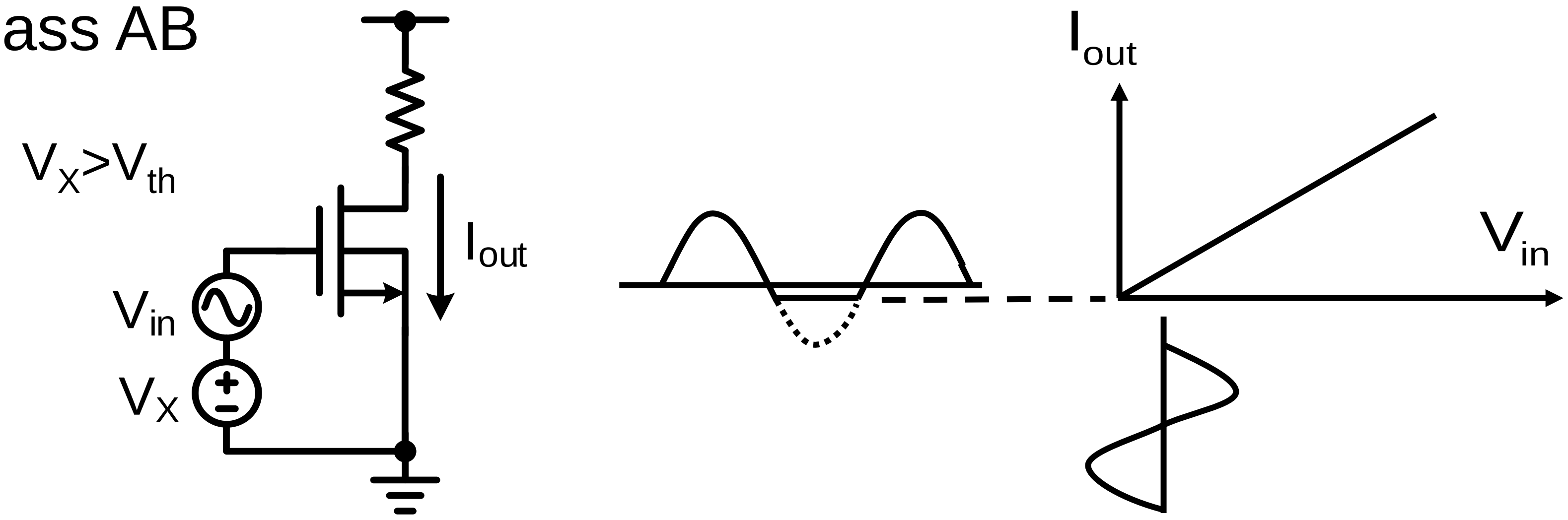


- Class B

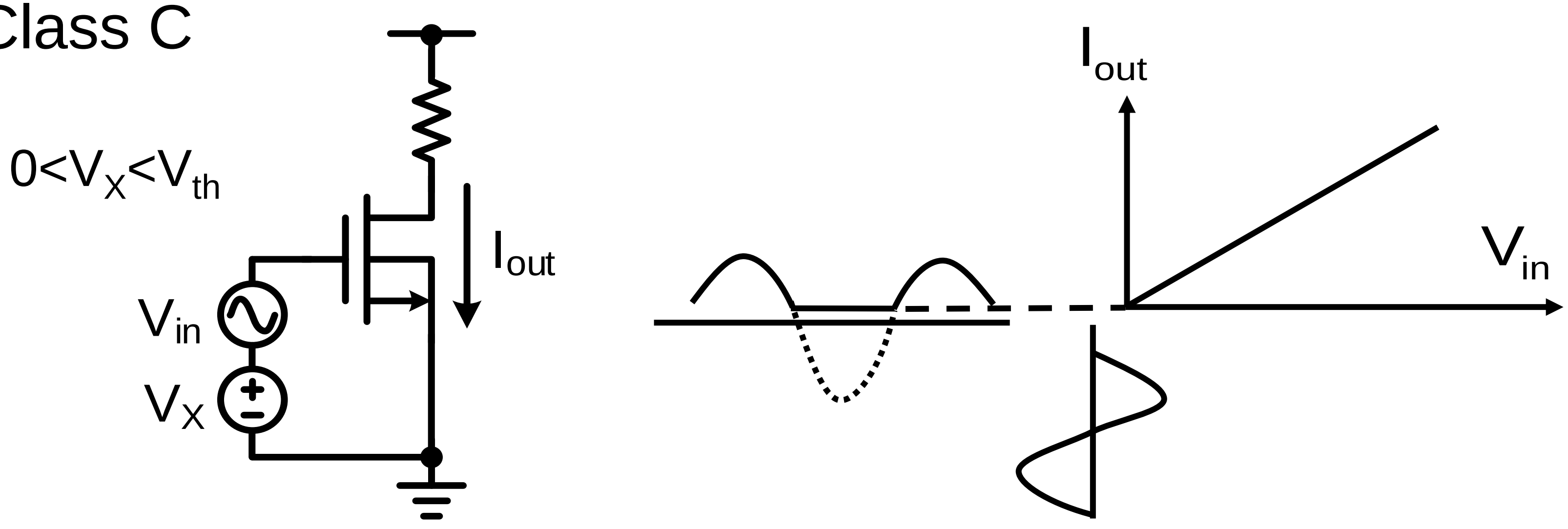


Amplifier Classification (Cont.)

- Class AB

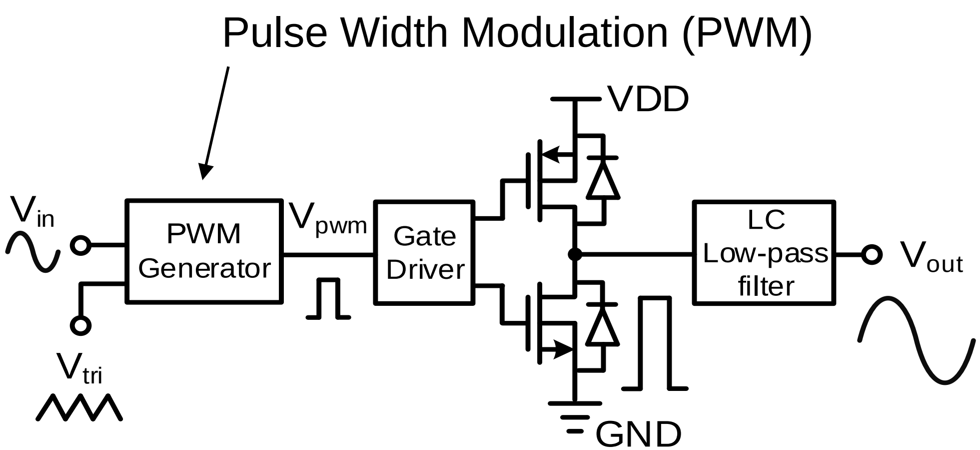


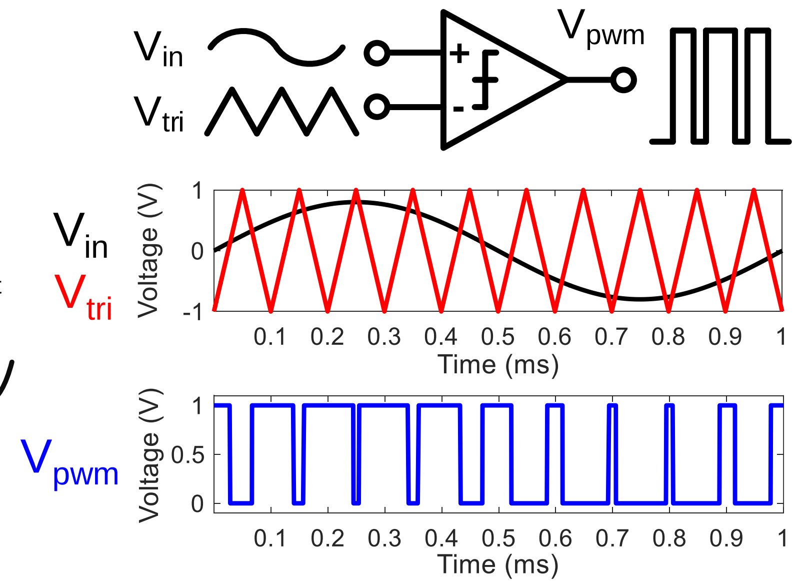
- Class C



Class-D Amplifier

- Class A, B and AB amplifier → Linear amplifier
- Class-D amplifier → Switching amplifier
  - ◆ Block diagram
  - ◆ Common PWM Generator



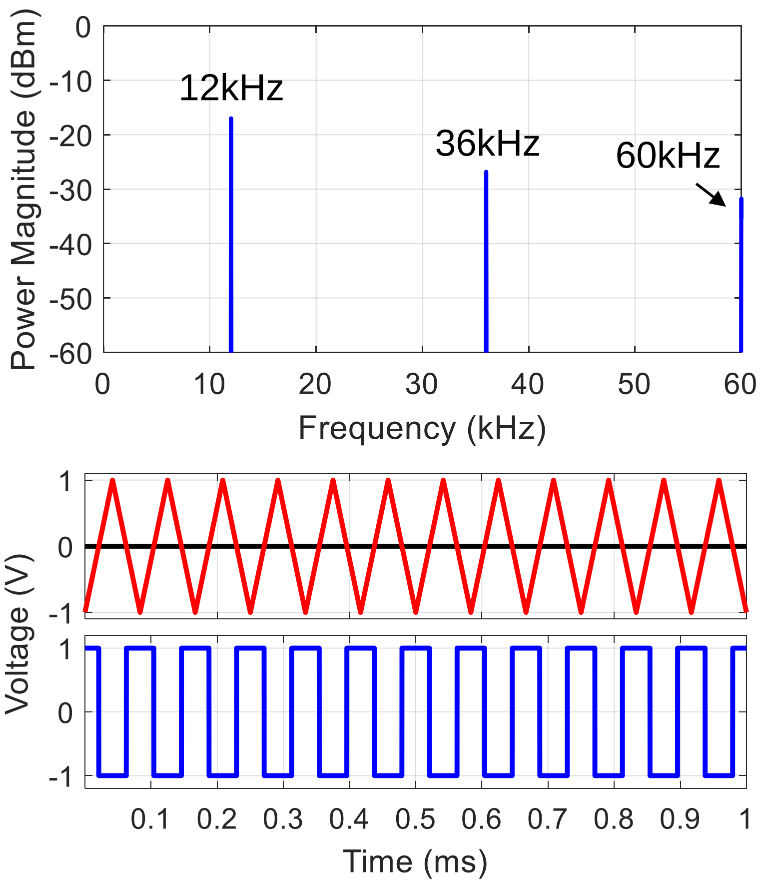


- Characteristics
  - ◆ Low power dissipation → High efficiency
  - ◆ Small heat sink → Small size
  - ◆ Distortion problem due to switching scheme

Class-D Amplifier (Cont.)

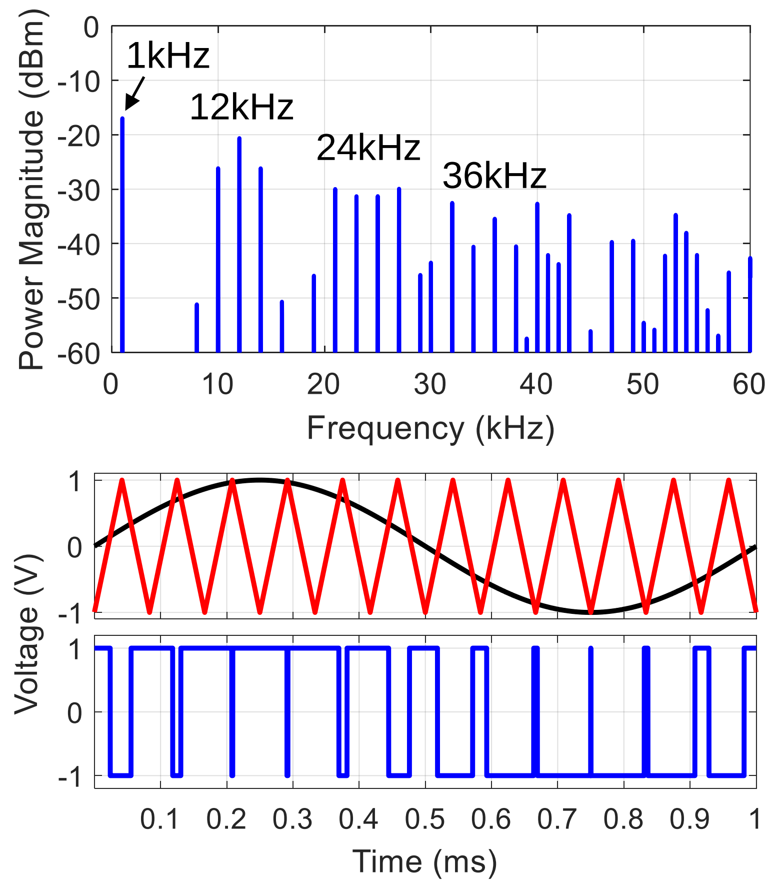
- Spectrum of Class-D signal
  - Square wave

    Triangular wave: V_{pp}=2V, Freq.=12kHz
    Input sine wave: V_{pp}=0V



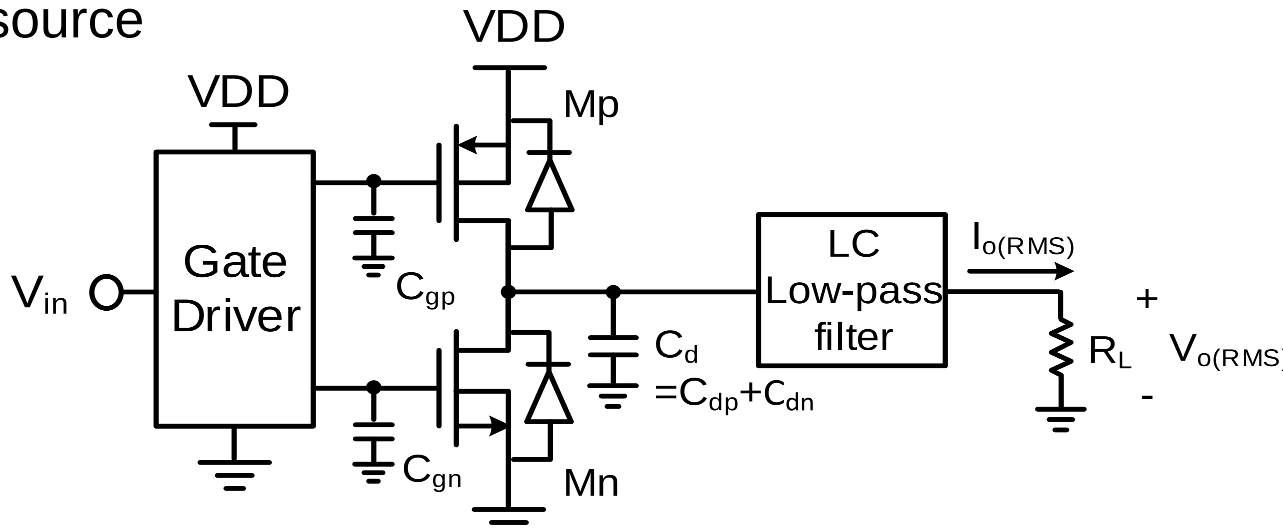
  - PWM

    Triangular wave: V_{pp}=2V, Freq.=12kHz
    Input sine wave: V_{pp}=2V, Freq.=1kHz



Efficiency of Class-D Amplifier (Cont.)

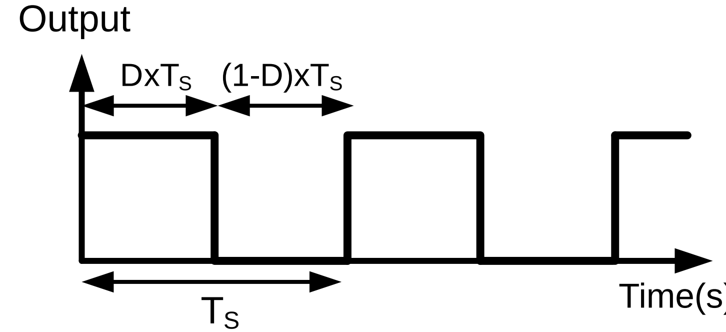
- Power loss source



  - Switching loss (P_{sw}): $P_{sw}=(C_{gp}+C_{dp}+C_{gn}+C_{dn})\times VDD^2\times f_{sw}$
  - Conduction loss (P_{con}): $P_{con}=[D\times R_{on}(M_p)+(1\text{-}D)\times R_{on}(M_n)]\times I_{o(RMS)}^{\;2}$

- Efficiency estimation

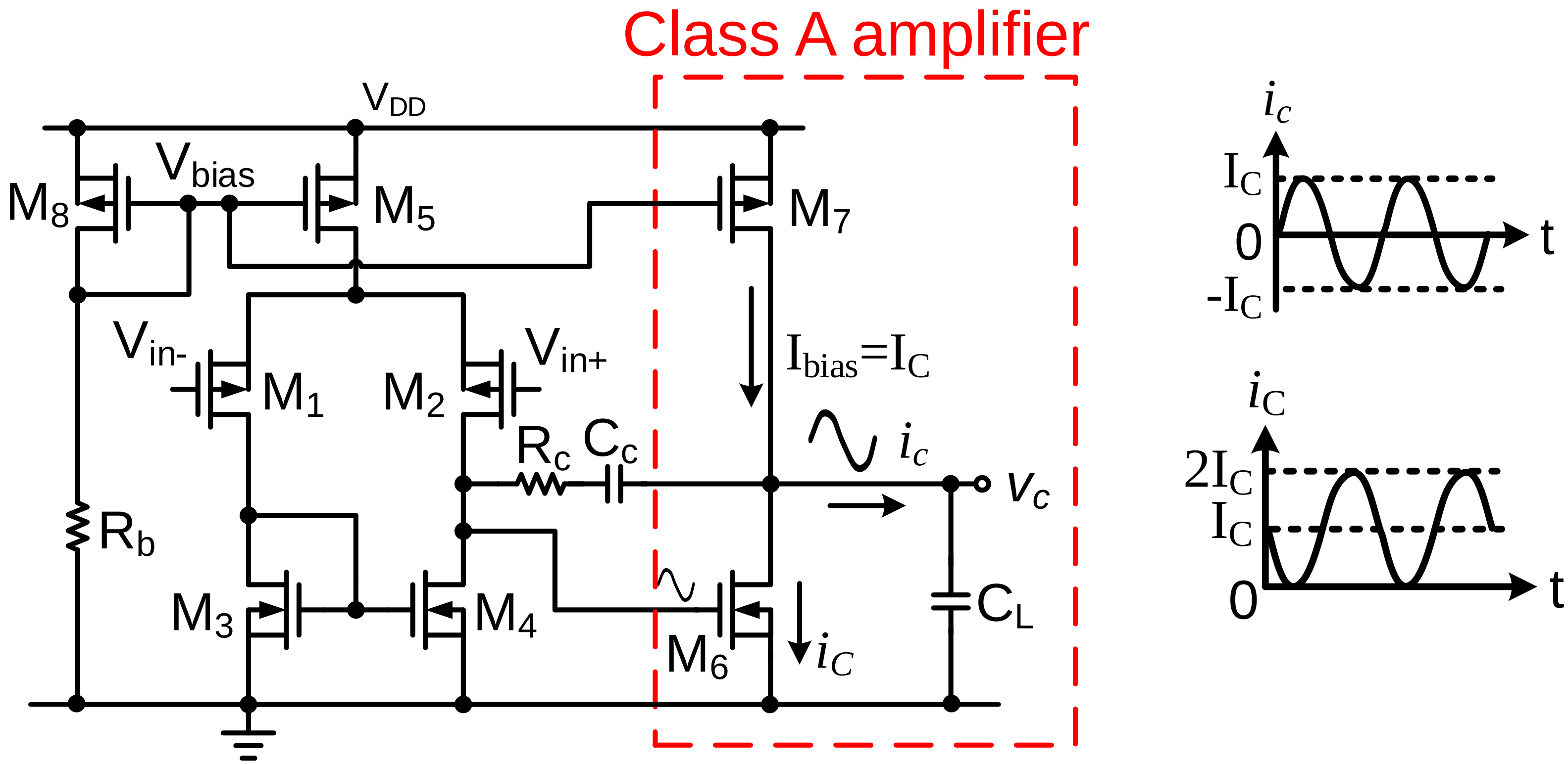
$$\eta=\frac{P_o}{P_i}\times 100\%=\frac{I_{o(RMS)}^{\;2}\times R_L}{I_{o(RMS)}^{\;2}\times R_L+P_{sw}+P_{con}}\times 100\%$$



Where D is the duty of PMOS on

Output Stage of Two-Stage OPAMP

- Current waveforms



◆ For $-I_C \leq i_c(t) \leq I_C$

➢ M_6 conducts for entire cycle of the input signal

Cascade and Cascode CMOS OPAMPs

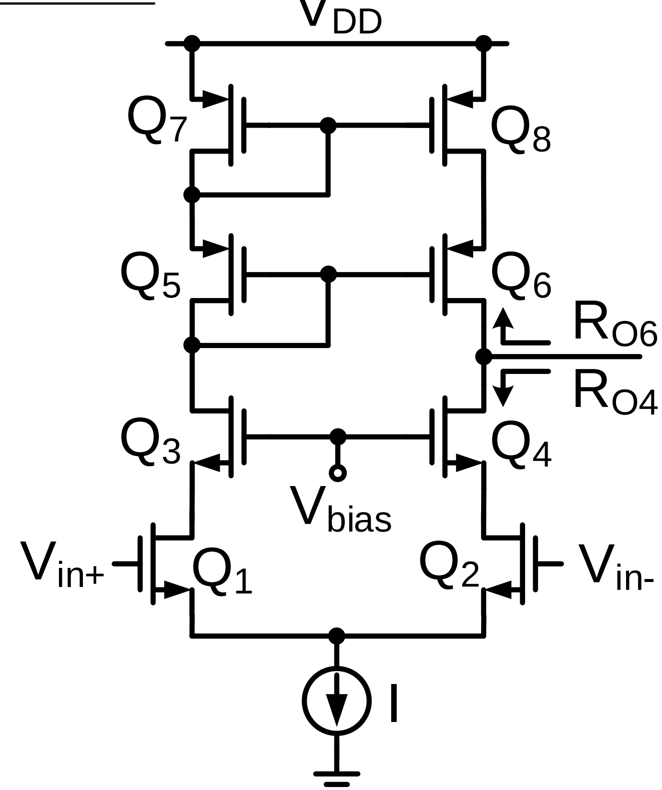
- Cascade two-stage CMOS OPAMP
  - Most popular and works well with low capacitive load
  - Problems
    - Limited slew rate due to large C_c
    - Limited bandwidth with large C_L
    - PSRR is reduced by pole-splitting

- Condition
  - Low output resistance is not required
  - High open-loop gain is required
  - Large phase margin can be maintained with large C_L

  → Cascode configuration can provide attractive solutions for the above problems.

- Cascode CMOS OPAMP
  - Gain of two-stage OPAMP can be increased by adding gain stage in cascade.
    - phase shift is increased (i.e. PM↓)
  - Cascode configurations can be used to increase gain in the existing stage.

Cascode CMOS OPAMP



- Output resistance (Ro) is increased
  - Voltage gain $A = -g_{m1}R_O$
    - Gain is increased

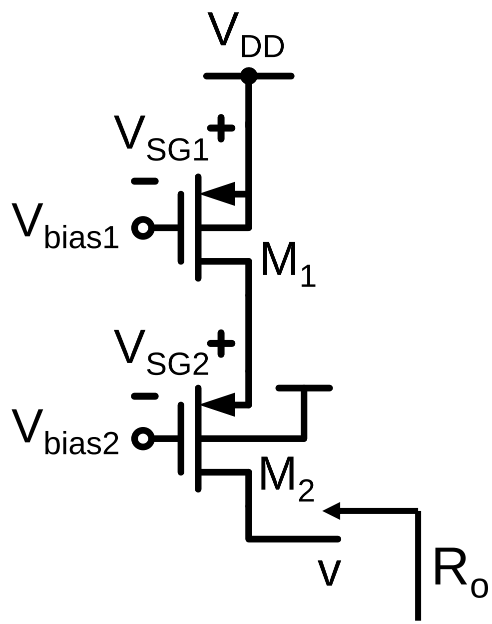
$$R_{O4} \approx (g_{m4}r_{ds4})r_{ds2}$$
$$R_{O6} \approx (g_{m6}r_{ds6})r_{ds8}$$
$$R_O = R_{O4} \parallel R_{O6}$$

- Common-mode range is lowered and more transistors are stacked between the two power supplies.
  - Folded-cascode has larger common-mode range
- Cascode and folded-cascode OPAMPs are also named as “transconductance OPAMP” or “operational transconductance amplifier (OTA)”

Cascode Circuit as a load

- Use cascode → Reduce # of stages in the design of high-gain OPAMP

- $R_o=(g_{m2}r_{ds2})r_{ds1}$

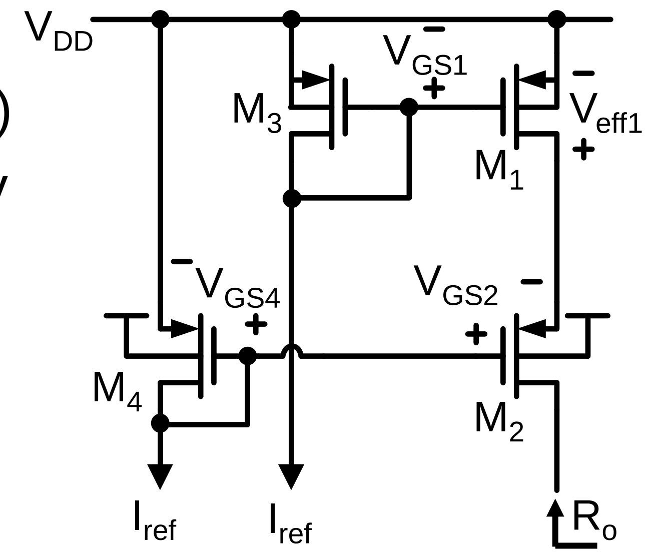


- Voltage drop on the two MOSFETs must be minimized to increase voltage swing
  - ◆ M_1 works at the voltage of $V_{DS(min)}= V_{eff1}$

Cascode Circuit as a load (Cont.)

- $(W/L)_{M1}= (W/L)_{M2}= (W/L)_{M3}= 4(W/L)_{M4}$
  - $(V_{GS3}-V_{tp})=V_{eff3}=V_{eff1}=(V_{GS4}-V_{tp})/2=(V_{GS2}-V_{tp})$
  - Make $(V_{DS1}$ & $V_{DS2}) > (V_{eff1}$ & $V_{eff2})$ for safety
  - $(W/L)_{M4}< (1/4)(W/L)_{M1}$
  - Output impedance

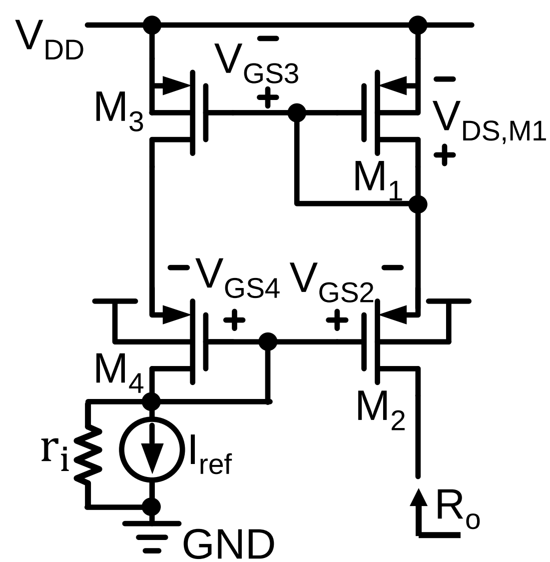
$$R_o \approx gm_2 r_{ds2} r_{ds1}$$



- One of other examples is self-biased
  - Output impedance

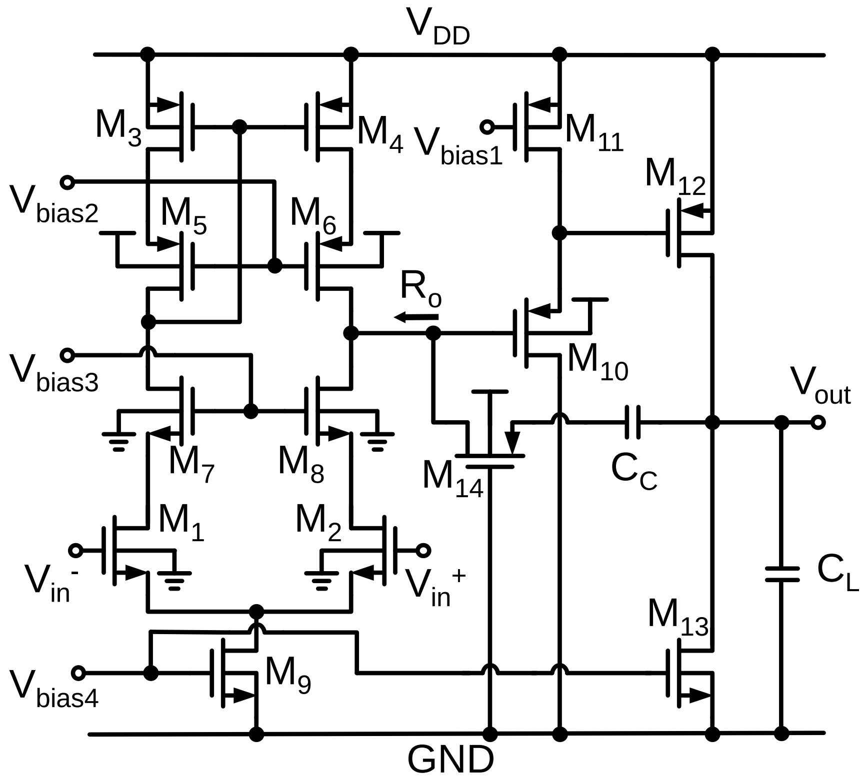
$$R_o \approx (g_{m2} r_{ds2}) \frac{g_{m3}}{g_{m1}} (r_{ds3} // r_i)$$

where r_i is the output resistance of I_{ref}



CMOS OPAMP Using Cascode Load

- High gain stage
- Example $R_0 = \dfrac{1}{\left(\dfrac{1}{g_{m6}r_{d6}r_{d4}}\right)+\left(\dfrac{1}{g_{m8}r_{d8}r_{d2}}\right)}$
- Structure
  - Differential pair with cascode load
  - Level shifter
    - Usually a source follower
  - Common source amplifier
  - Miller capacitor (C_C)→ Pole-splitting
  - M_{14} → Eliminate right-plane zero



- Problems
  - Large C_L → The bandwidth will be limited by non-dominant pole
  - Large C_c → The slew rate will be limited
  - PSRR is reduced by pole-splitting

→ All problems can be eliminated by using folded-cascode configuration

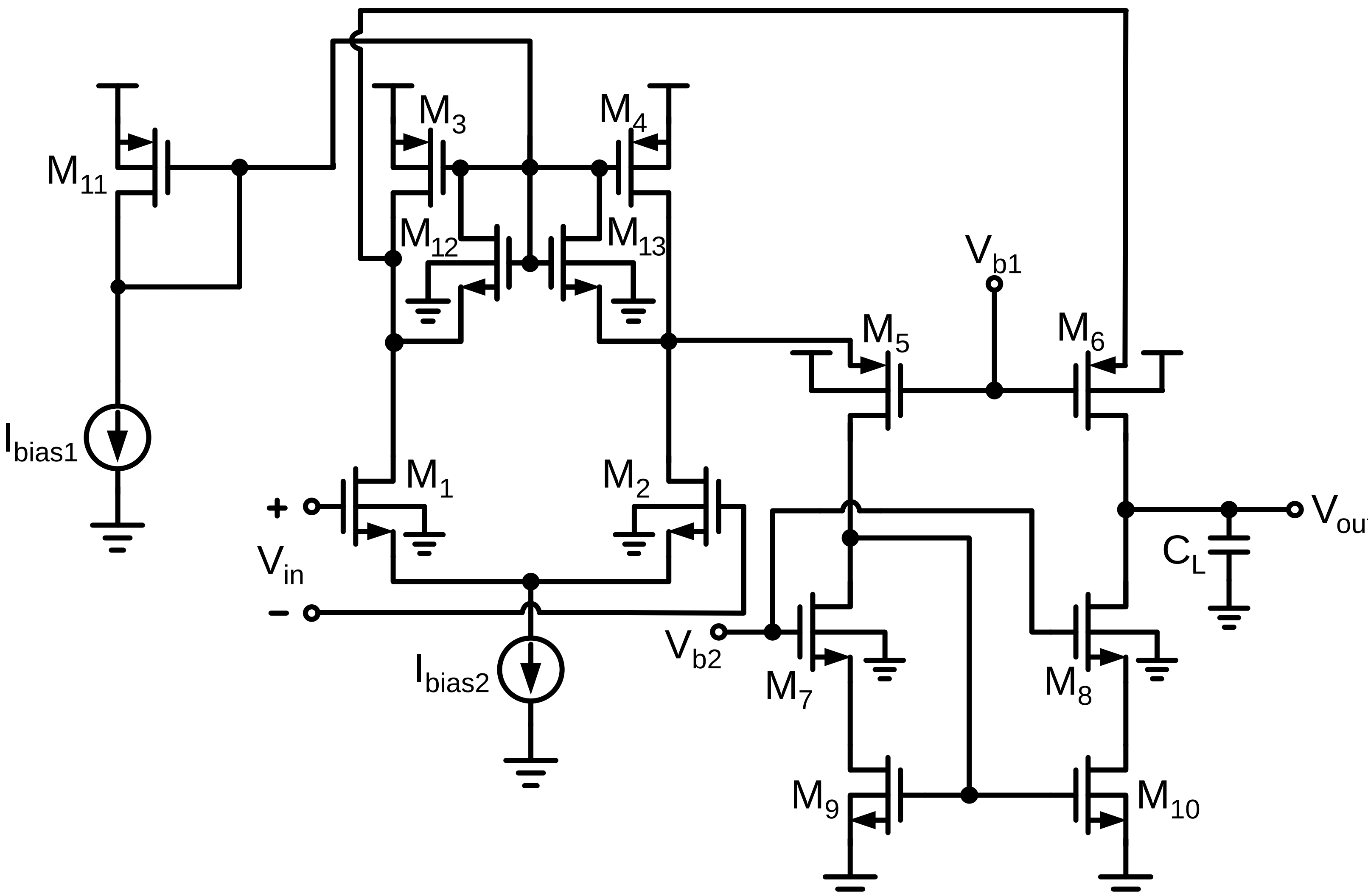
Folded-Cascode OPAMP

- Transconductance value is one of the most important parameters of these OPAMPs → Also called Operational Transconductance Amplifiers (OTAs)

- Many modern CMOS OPAMPs are designed to drive only capacitive load
  - Using a voltage buffer to obtain low output impedance is not necessary
  - Realizing OPAMPs having higher speed and larger signal swings than those that must also drive resistance loads is possible
    - Only a single high-impedance node at the output of OPAMP that drives only capacitive loads
    - The impedance seen at all other internal nodes of OPAMP is relatively low impedance (~$1/g_m$)
      - The OPAMP speed is maximized
  - The compensation is usually achieved by the load capacitance
    - As the load capacitance gets larger, the OPAMP usually becomes more stable but also slower.

Folded-Cascode OPAMP (Cont.)

- A folded-cascode OPAMP



Folded-Cascode OPAMP (Cont.)

- Current mirrors are all wide-swing cascode
  - High output impedance
  - High DC gain

- Two extra transistors, M_{12} and M_{13}, serve two purposes
  - Increase slew rate
  - Allow OPAMP to recover more quickly following a slew rate condition
    - Because M_{12} and M_{13} prevent the drain voltages of M_1 and M_2 from having large transients
    - However, if the 2nd pole is located at the source of M_5, its bandwidth is reduced

- The compensation is realized by the load capacitor (C_L) and realizes dominant pole compensation. In applications where the load capacitance is very small, it is necessary to add additional compensation capacitance in parallel with the load to guarantee stability.

Folded-Cascode OPAMP (Cont.)

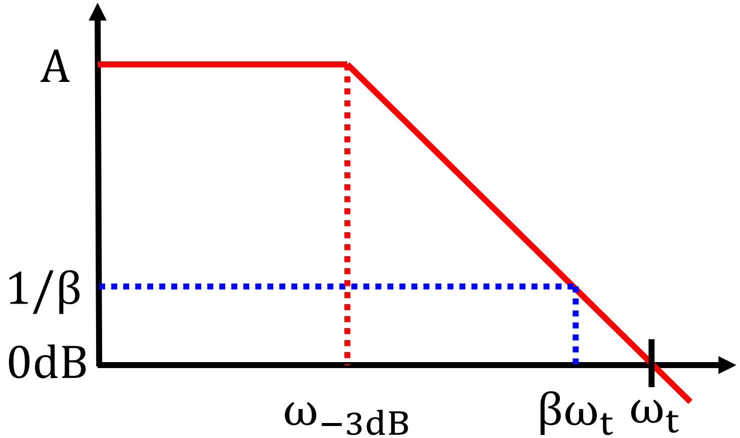
- Small-signal analysis
  - $$A_V = \frac{V_{out}(s)}{V_{in}(s)} = g_{m1} Z_L(s) = g_{m1}(r_{out} // C_L) = \frac{g_{m1} r_{out}}{1 + s r_{out} C_L}$$
  - Unity-gain frequency $\omega_t = \dfrac{g_{m1}}{C_L}$
  - 2nd pole is usually generated at the nodes of M_1(or M_3) drain & M_2(or M_4) drain

$$P_2 \approx \frac{g_{m6}}{C_{total}(\text{at } M_1 \text{ drain})}$$

    - In BiCMOS, M_5(M_6) is usually replaced by a BJT to push P_2 to higher frequency. (In BiCMOS, ω_t can therefore be maximized.)
  - Slew rate $SR = \dfrac{I_{D4}}{C_L}$
    - M_{12} and M_{13} are included to increase SR. (These two transistors are also used to clamp the drain voltage of M_{12} and M_{13}.)

Linear Settling Time

- Time constant for linear settling is approximately equal to $\frac{1}{\omega_{-3dB}}$ if nondominant poles are larger than $\omega_t (= \omega_u)$



For closed-loop OPAMP,

$$\omega_{-3dB} = \beta\omega_t$$

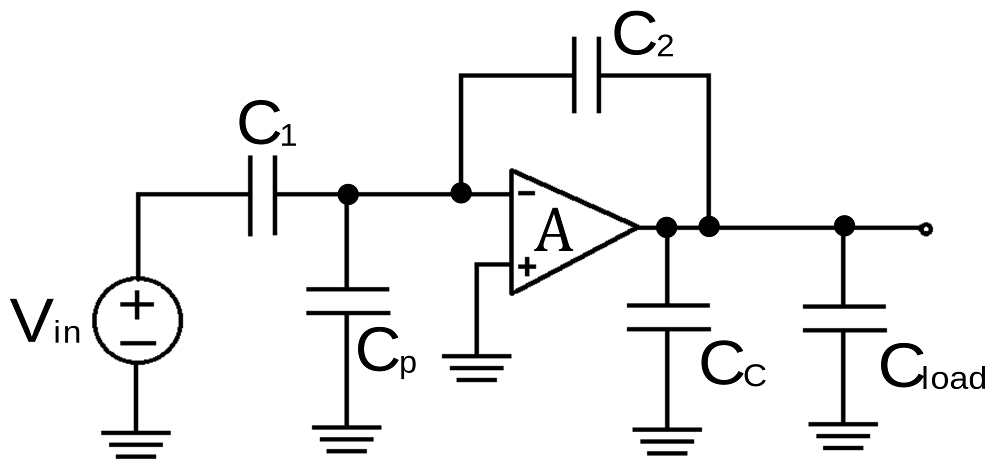
→ Time constant $\tau = \frac{1}{\beta\omega_t}$

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

$$\omega_t = \frac{g_{m1}}{C_L}$$

Linear Settling Time (Cont.)

- -3dB frequency of a closed-loop cascoded OPAMP
  - Example

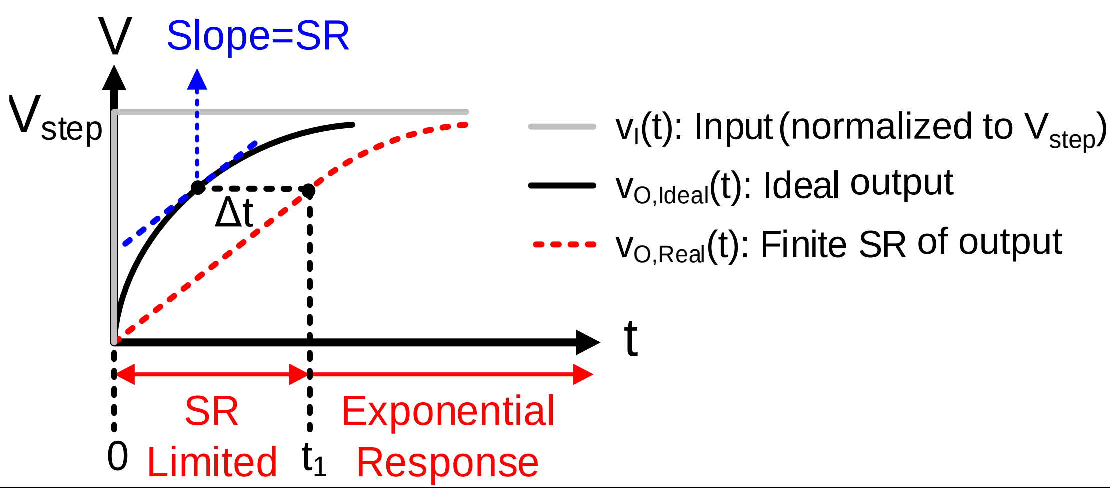


$$\beta = \frac{1/[s(C_1 + C_p)]}{1/s(C_1 + C_p) + 1/sC_2} = \frac{C_2}{C_1 + C_p + C_2}$$; C_P is parasitic capacitance

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

Linear Settling Time (Cont.)

- Step Response
  - $V_{out}(t) = V_{step}(1 - e^{-t/\tau})$
  - $\frac{d}{dt} V_{out}(t)|_{t=0} = \frac{V_{step}}{\tau}$
    - If the OPAMP slew rate is larger than this value, no slew-rate limitation would occur
  - Ideal case
    - If 1% accuracy is required, settling time (T_S) is 4.6τ
    - If 0.1% (i.e. 10-bit) accuracy is required, T_S is 7τ

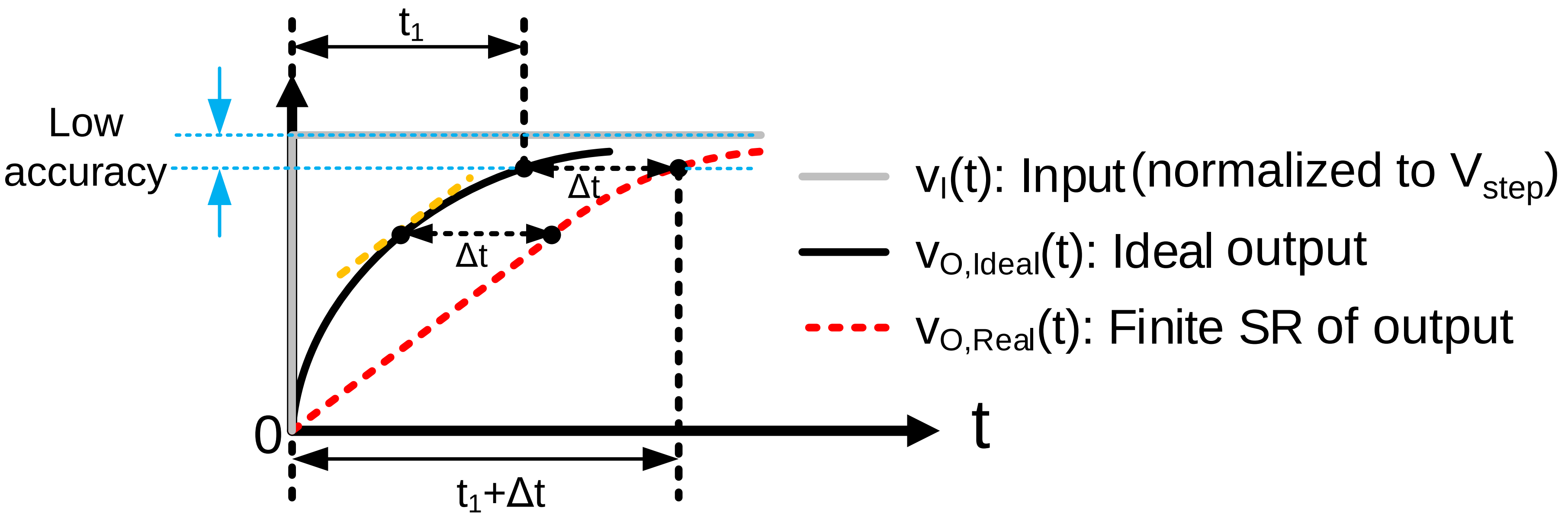


Linear Settling Time (Cont.)

- If low accuracy is required, t_1 is not much longer than Δt

→ Larger slew rate is usually chosen

  - Minimize Δt to keep the response fast enough

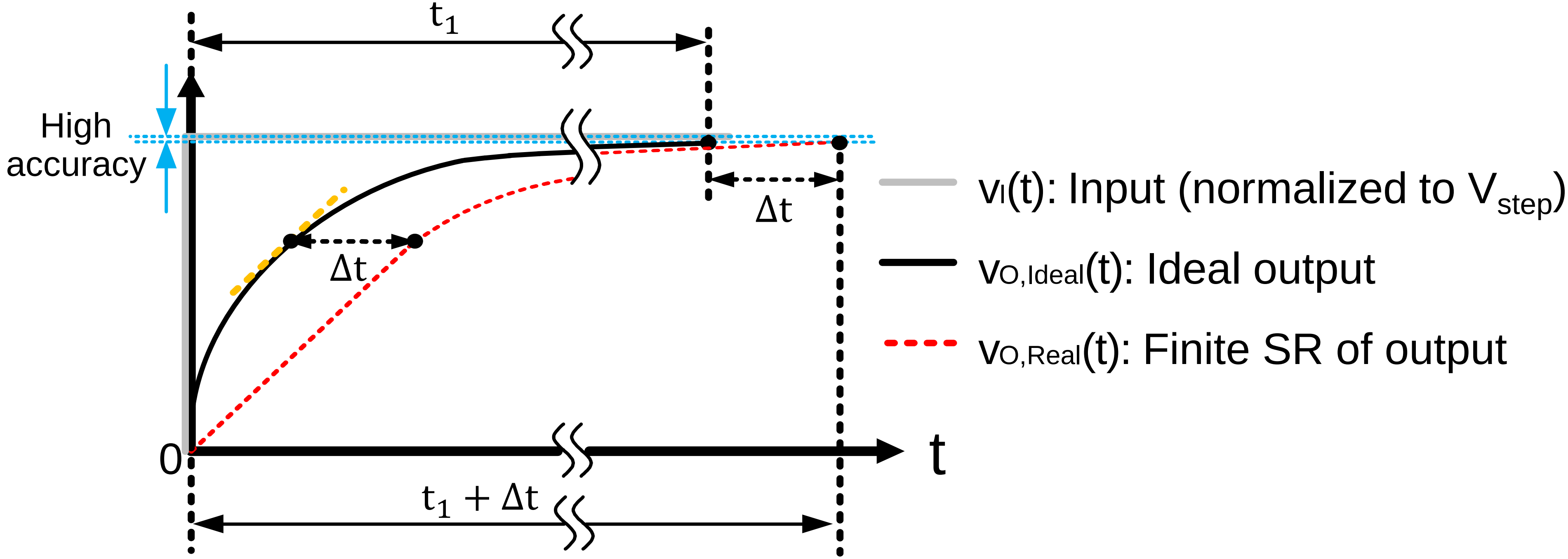


Linear Settling Time (Cont.)

◆ If high accuracy is required, t_1 is much longer than Δt

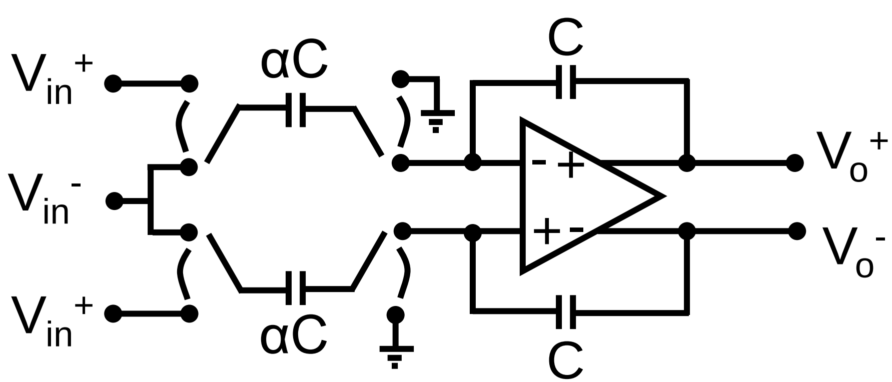
→ Smaller slew rate is usually chosen

- Increase a little bit of time in Δt, but still $t_1 >> \Delta t$
- Greatly relax the slew rate requirement
- Greatly reduce current and thus save power



Fully Differential CMOS Switched-Capacitor Circuit

- Power supply rejection is high
- Larger chip area compared with single-ended output
- Output swing is doubled
  - DR is 6dB greater than single-ended OPAMPs
- The effect of clock feedthrough noise is minimized by the differential configuration since it will appear as a common-mode signal.

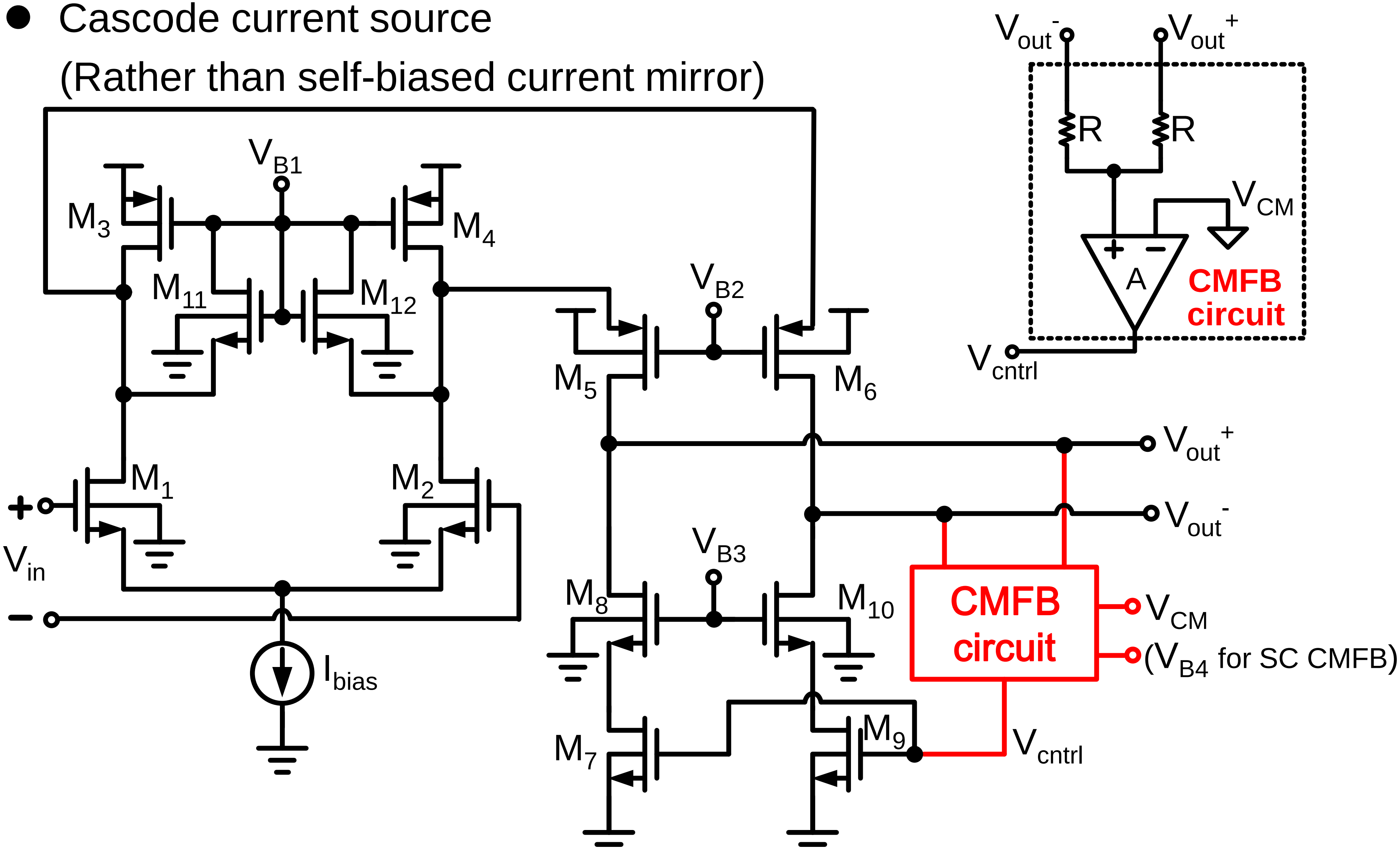


Fully Differential OPAMPs

- Fully differential signal paths
  - Differential input and differential output
  - Used in most modern high-performance analog ICs
- Help reject noise from the substrate as well as from switches turning off in switched-capacitor applications.
  - Ideally, noise affects both signal paths identically and will then be rejected since only the difference between signals is important.
  - In reality, this rejection only partially occurs since the mechanisms introducing the noise are usually nonlinear with respect to voltage levels. For example, substrate noise will usually feed in through junction capacitances, which are nonlinear with voltage.
  - Certainly, the noise rejection of a fully differential design will be much better than that for a single-ended output design. (>20dB can be expected)
- Common-mode feedback (CMFB) circuit must be added to establish the common-mode (i.e. average) output voltage.
- Reduced slew rate in one direction (compared to single-ended design)
  - Maximum current for slewing is often limited by fixed-bias currents in the output stages.

Fully Differential Folded-Cascode OPAMP

- Cascode current source
  (Rather than self-biased current mirror)



Fully Differential Folded-Cascode OPAMP (Cont.)

- CMFB circuit forces the average of the two outputs to a predetermined value
- Maximum negative slew rate is limited by I_{D7} and I_{D9}
- Clamp transistors M_{11} and M_{12}
- Dominant pole : output node

  2nd pole : node at M_1 (or M_2) drain (usually)
  - n-channel input and p-channel for M_5 and M_6
    - High transconductance
    - High gain
  - p-channel input and n-channel for M_5 and M_6
    - Maximize 2nd pole frequency
    - Unity-gain bandwidth can be maximized.

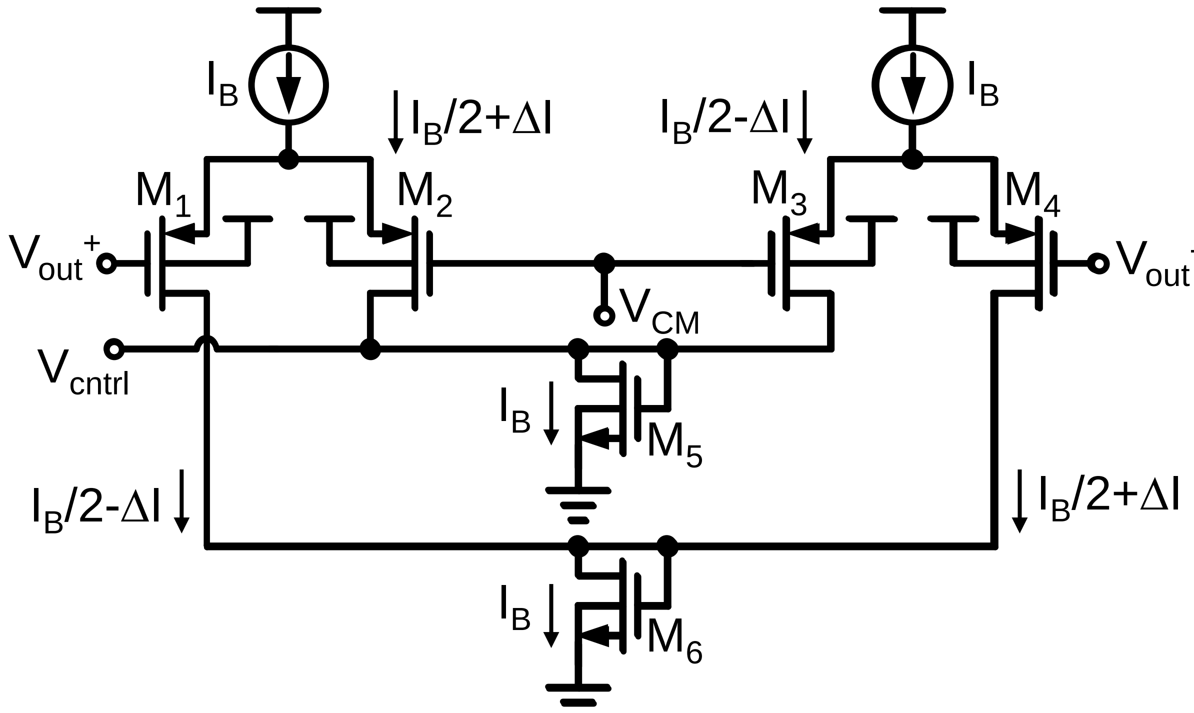
Common-Mode Feedback (CMFB) Circuits

- Force output common-mode voltage to a predetermined value
- CMFB is often the most difficult part of the OPAMP to design.
- Two typical approaches
  - Continuous-time
    - Limited signal swing
  - Switched-capacitor
    - Used in switched-capacitor circuits
    - Signal swings are not limited
    - Becomes a source of noise
    - Increases load capacitance

- By having as few nodes in the common-mode loop as is possible, compensation is simplified without having to severely limit the speed of the CMFB circuit. For this reason, the CMFB circuit is usually used to control current sources in the output stage of the OPAMP.

CMFB Circuits

- A continuous-time CMFB circuit



- ◆ The circuit can not operate correctly if the OPAMP output voltage is so large that transistors in the differential pairs turn off.
- ◆ When common-mode voltage is zero

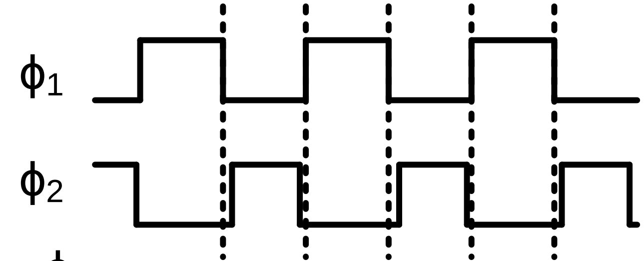
$$I_{D2} = \frac{I_B}{2} + \Delta I, \quad I_{D3} = \frac{I_B}{2} - \Delta I, \quad I_{D5} = I_B$$

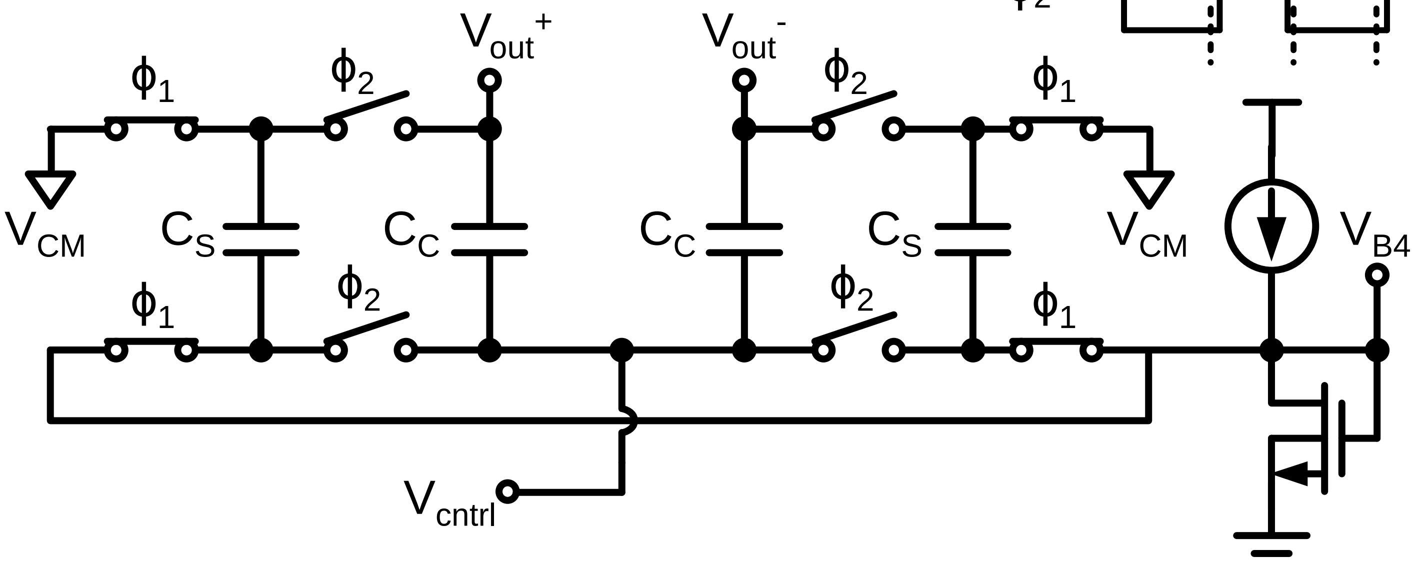
CMFB Circuits (Cont.)

- Operational principle of CMFB circuits
  - For example, when a positive common-mode signal is present → I_{M2} and I_{M3} increase → I_{M5} increase → V_{cntrl} increase
  - V_{cntrl} sets the current levels in the n-channel current sources at the output of the OPAMP, thus, bringing the common-mode voltage back to V_{CM}
  - If the common-mode loop gain is large enough, and the differential signals are not so large as to cause transistors in the differential pairs to turn off, the common-mode output voltage will be kept very close to V_{CM}.

CMFB Circuits (Cont.)

- A switched-capacitor circuit





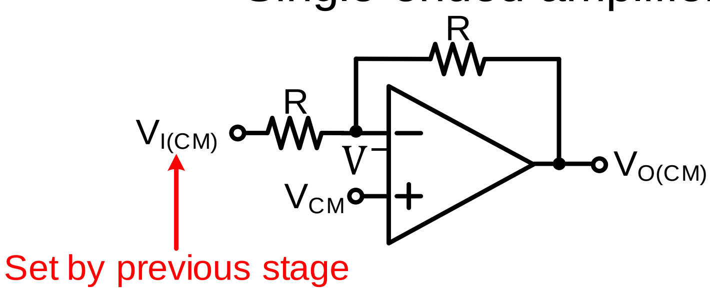
  - Using larger capacitance values overloads the OPAMP
  - Reducing the capacitors too much caused common-mode offset voltages due to charge injection of the switches.

$$\frac{V_{out}^{+} + V_{out}^{-}}{2} - V_{cntr1} \approx V_{CM} - V_{B4}$$

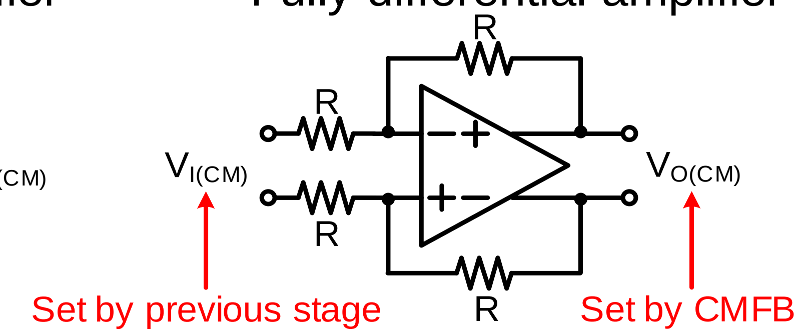
Common-Mode Voltage of OPAMP

- Take inverting amplifier (with ac gain=1) for example

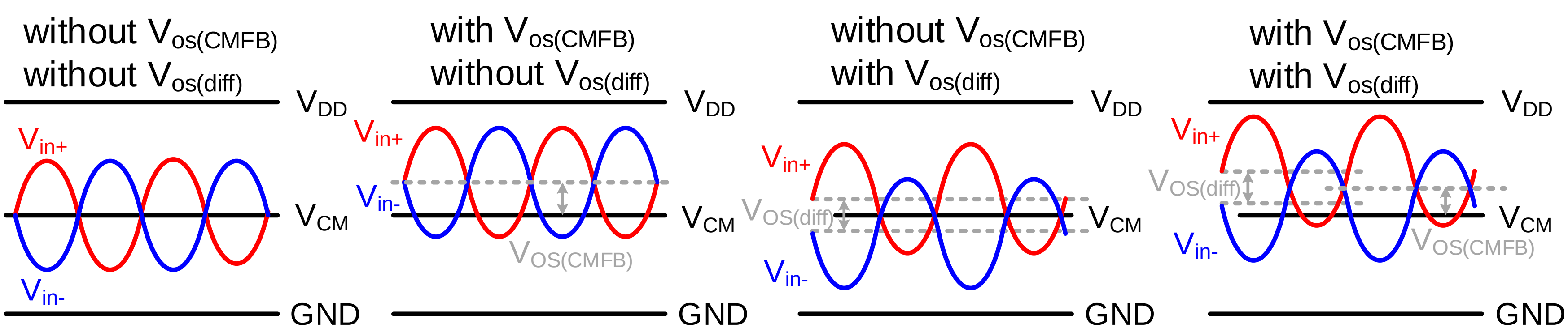
Single-ended amplifier



Fully-differential amplifier



- Single-ended amplifier
  - Input common-mode voltage at $V^- = V_{CM}$ (Virtually shorted to V^+)
  - Output common-mode voltage at $V_{O(CM)} = V_{CM} - (V_{I(CM)} - V_{CM})$
- Fully-differential amplifier
  - Input common-mode voltage at $V^+ = V^- = \frac{1}{2}(V_{I(CM)} + V_{O(CM)})$
  - Output common-mode voltage at $V_{O(CM)}$ → Set by CMFB

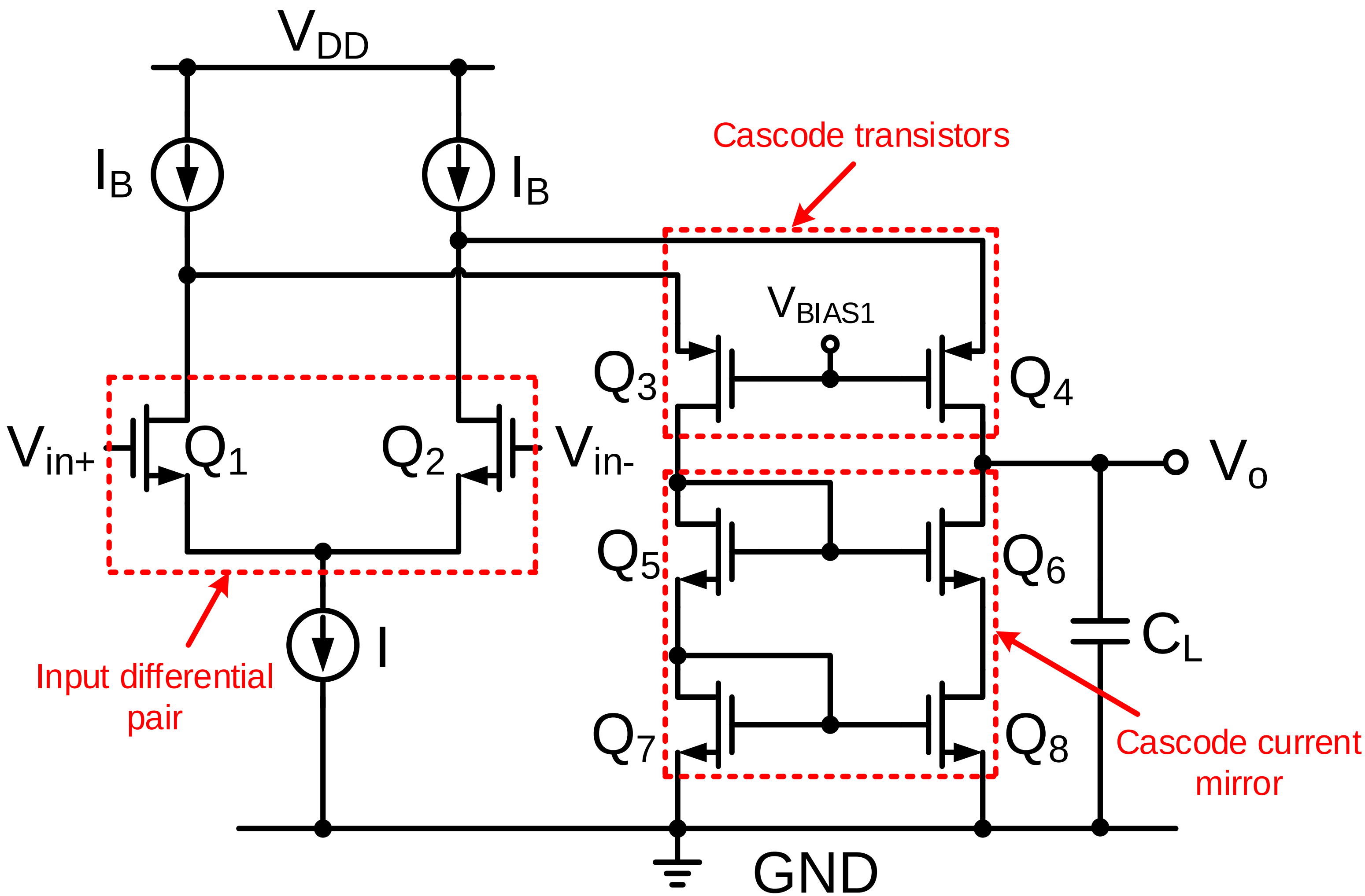


Appendix

- Folded-Cascode CMOS OPAMP
- Current mirror OPAMP
- Alternative fully differential OPAMPs
- BiCMOS amplifiers

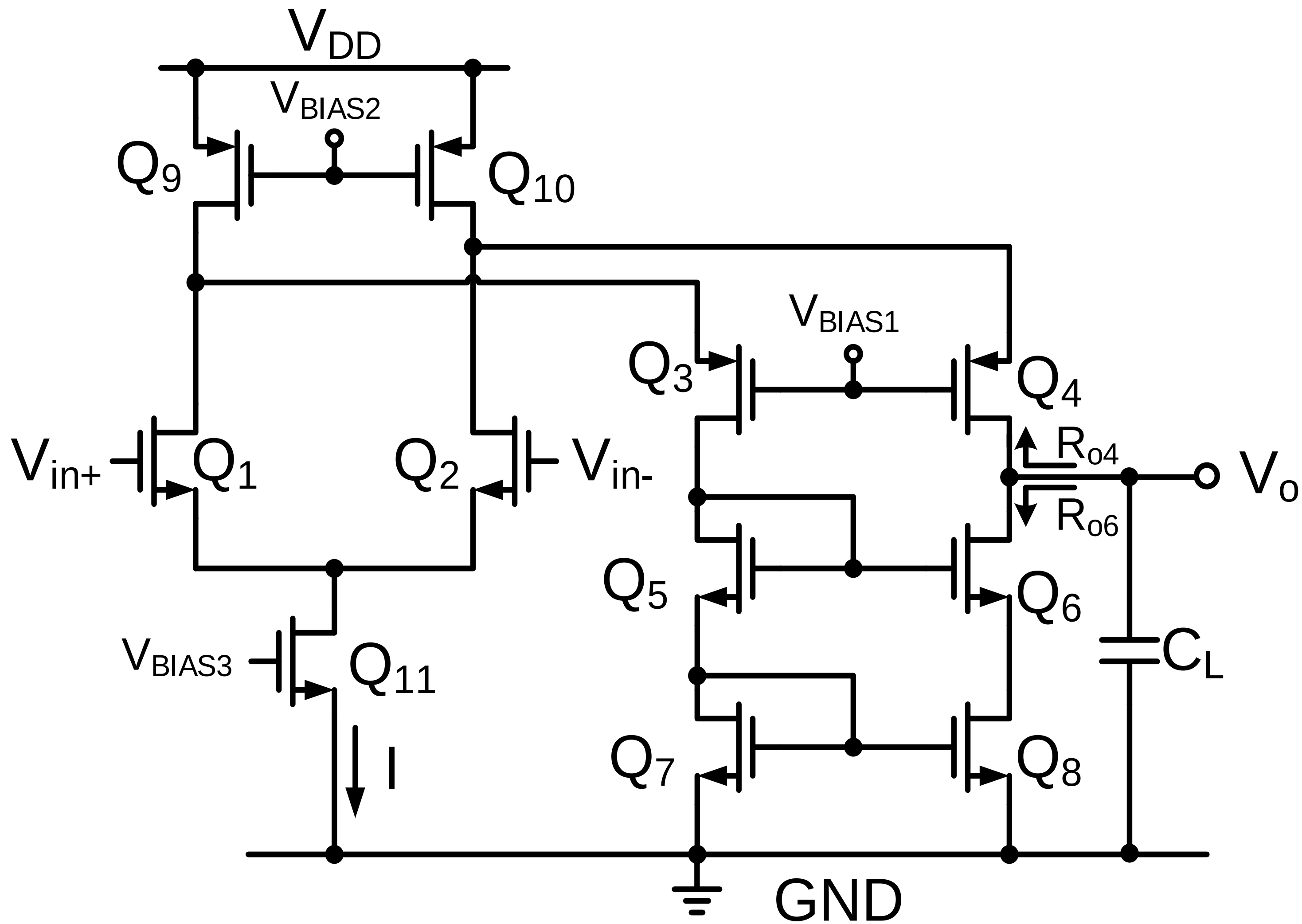
Folded-Cascode CMOS OPAMP

- $Q_3 \sim Q_8$ are folded and connected to GND



Folded-Cascode CMOS OPAMP (Cont.)

- $Q_9 \sim Q_{11}$ form externally-biased current sources
  Q_5 and Q_8 form self-biased current sources



Folded-Cascode CMOS OPAMP (Cont.)

- Input common-mode range

  Common-mode range is increased (compared with cascode OPAMPs). However, it is small compared with 2-stage OPAMPs

$$V_{OV11} + V_{OV1} + V_{tn} \le V_{ICM} \le V_{DD} - |V_{OV9}| + V_{tn}$$

- Output voltage swing

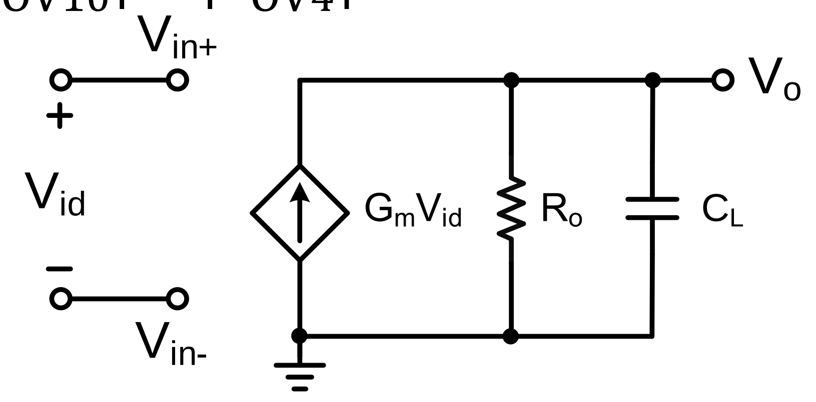
$$V_{OV7} + V_{OV5} + V_{tn} \le V_o \le V_{DD} - |V_{OV10}| - |V_{OV4}|$$

- Voltage gain

$$A = G_m R_O = g_{m1} R_O$$

$$R_O = R_{O4} \parallel R_{O6}$$

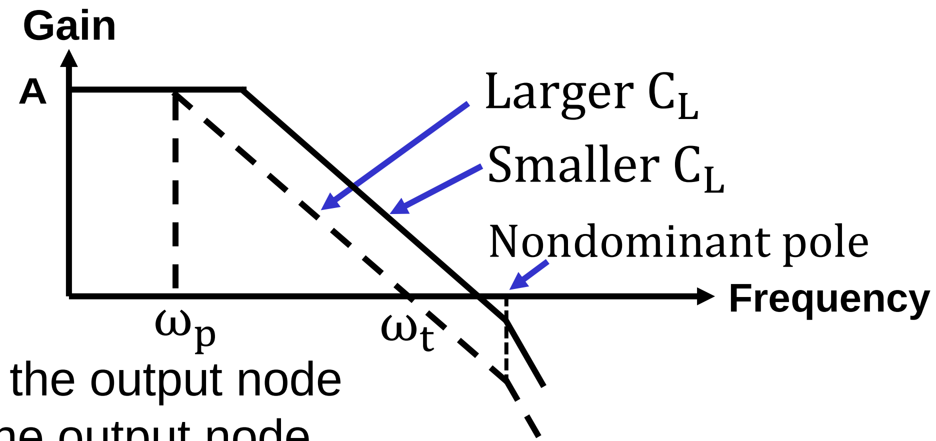
$$= [g_{m4} r_{ds4}(r_{ds2} \parallel r_{ds10})] \parallel [g_{m6} r_{ds6} r_{ds8}]$$



Folded-Cascode CMOS OPAMP (Cont.)

- Frequency response
  - Bode plot

$$\omega_p \approx \frac{1}{R_0 C_L} \qquad \omega_t \approx \frac{g_{m1}}{C_L}$$



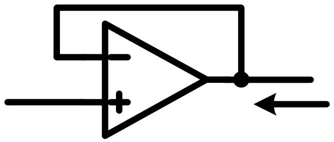
  - The only high-impedance point is the output node
    → Dominant pole is generated at the output node
  - The resistance of other nodes at level of 1/gm
    → Nondominant poles occur at other nodes
    The 2nd pole is usually at the source of Q_3 and Q_4
  - Nondominant poles are usually at frequencies beyond ω_t
    → If C_L is increased, then phase margin is increased
    → If C_L is not large enough, it can be augmented
  - No frequency compensation is required → Wide bandwidth
- Slew rate

$$SR = I/C_L = 2\pi f_t V_{OV1} = \omega_t V_{OV1}$$

Folded-Cascode CMOS OPAMP (Cont.)

- Folded-cascode OPAMPs have high open-loop output resistance It has been given the name **operational transconductance amplifier (OTA)**
- Its high output resistance (in the order of $g_m r_o^2$) is far from that for an ideal OPAMP (which has zero output resistance)
- To alleviate this concern somewhat, let us find the closed-loop output resistance R_{of} of a unity-gain follower (β = 1) formed by connecting the output terminal back to the negative input terminal

$$R_{of} = \frac{R_o}{1+A\beta} = \frac{R_o}{1+A} \approx \frac{R_o}{A} \Rightarrow R_{of} \approx \frac{1}{G_m}$$



A general result applying to any OTA with 100% voltage feedback.

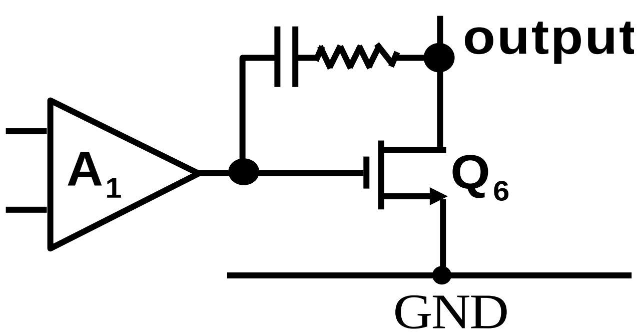
For folded-cascode OPAMPs, $G_m \approx g_{m1} \Rightarrow R_{of} \approx \dfrac{1}{g_{m1}}$

- g_{m1} is in the order of 1mA/V, and R_{of} will be of the order 1kΩ Although this is not very small, it's reasonable in view of the simplicity of the OPAMP circuit as well as the fact that this type of OPAMP (OTA) is not usually intended to drive low-valued resistive load.

Folded-Cascode CMOS OPAMP (Cont.)

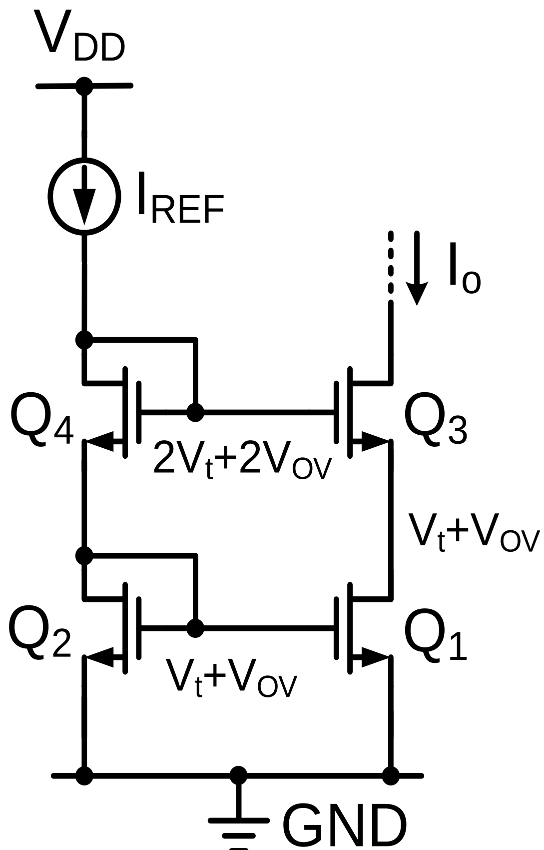
- High PSRR (to-V_{SS})
  - much less susceptible to the effect of high-frequency noise on GND
  - power supply noise may be induced from
    - logic circuit
    - switches of SC circuit
    - current switching

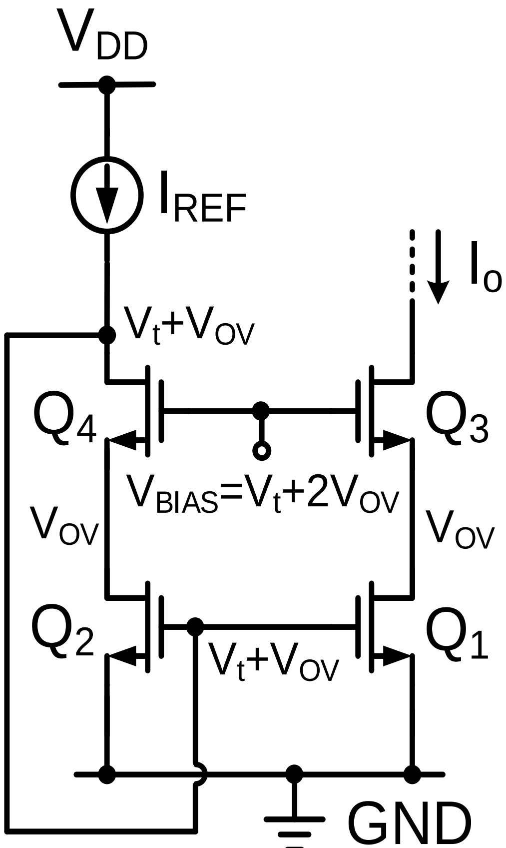
★ Low PSRR (to-V_{SS}) in cascaded 2-stage OPAMP
  - GND noise $\rightarrow$ Q_6 source $\rightarrow$ Q_6 gate $\rightarrow$ C,R $\rightarrow$ output
  - GND noise $\rightarrow$ Q_6 source $\rightarrow$ Q_6 V_{GS} $\rightarrow$ amplified and appear at output



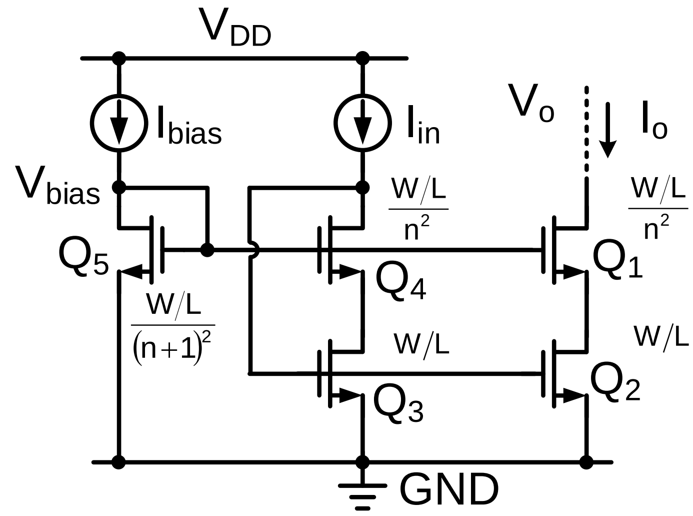
Wide-Swing Current Mirror

- Increased output voltage range
  - $v_{Omin} \geq V_{OV1} + V_{OV3} + V_{tn}$
  - $v_{Omin} \geq V_{OV1} + V_{OV3}$





Wide-Swing Current Mirror (Cont.)



- Design example

  a varying signal $I_{in} \leq I_{bias}$

$$V_{eff_2} = V_{eff_3} = \sqrt{\frac{2I_{D_2}}{\mu_n C_{ox}(W/L)}} = V_{eff}$$

$$(\because\ I_{D2} = \frac{\mu_n C_{ox}}{2}\frac{W}{L}V_{eff}^{\ 2})$$

Since $\left(\frac{W}{L}\right)_2 = \left(\frac{W}{L}\right)_3 = (n+1)^2\left(\frac{W}{L}\right)_5 = n^2\left(\frac{W}{L}\right)_1 = n^2\left(\frac{W}{L}\right)_4$

$$V_{eff_1} = V_{eff_4} = nV_{eff} \quad \text{for the target } I_{in} = I_{bias}$$

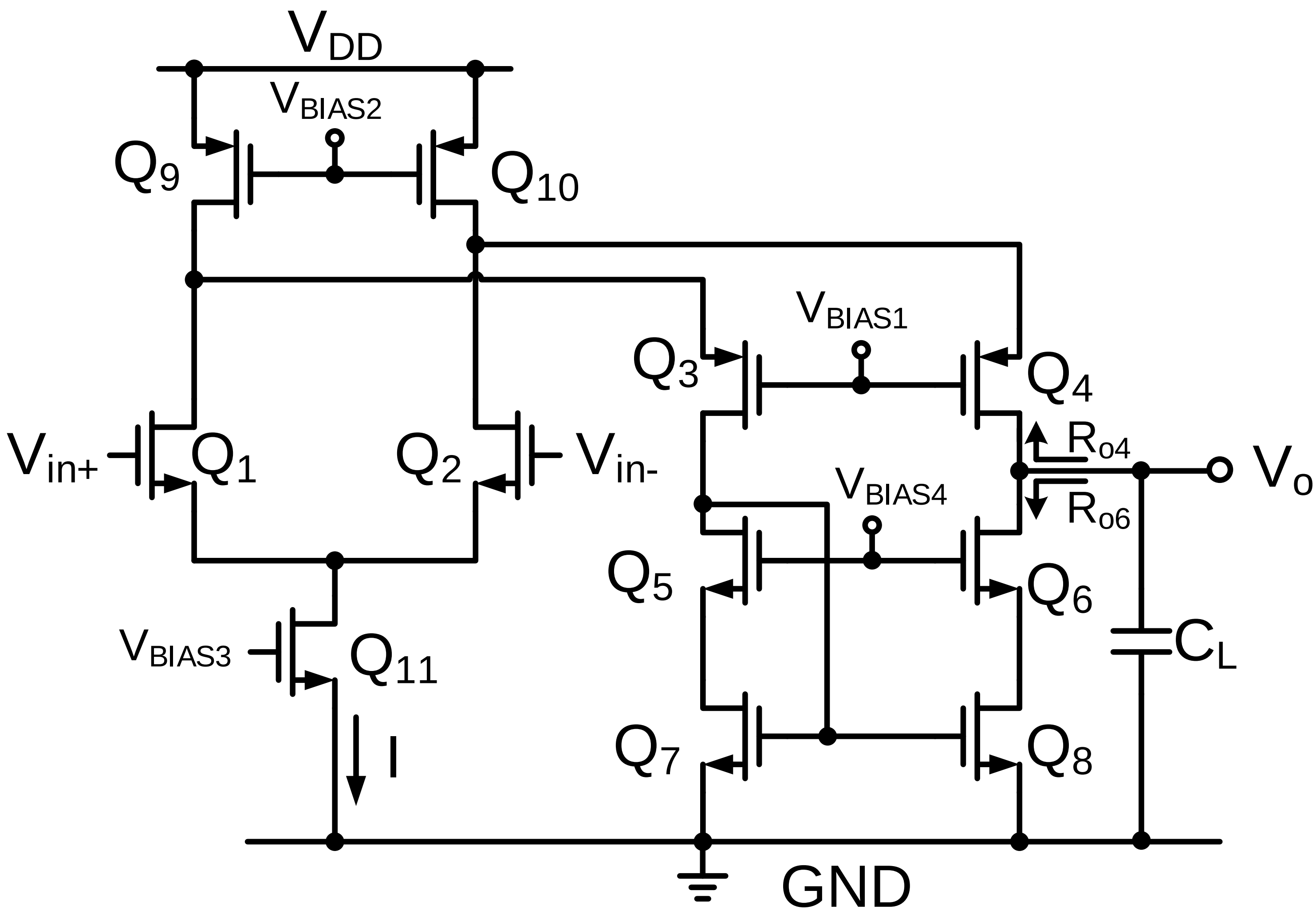
$$V_{G_5} = V_{G_4} = V_{G_1} = (n+1)V_{eff} + V_{th}$$

$$V_{DS_2} = V_{DS_3} = V_{G_5} - V_{GS_1} = V_{G_5} - (nV_{eff} + V_{th}) = V_{eff}$$

$$\Rightarrow V_o > V_{eff_1} + V_{eff_2} = (n+1)V_{eff}$$

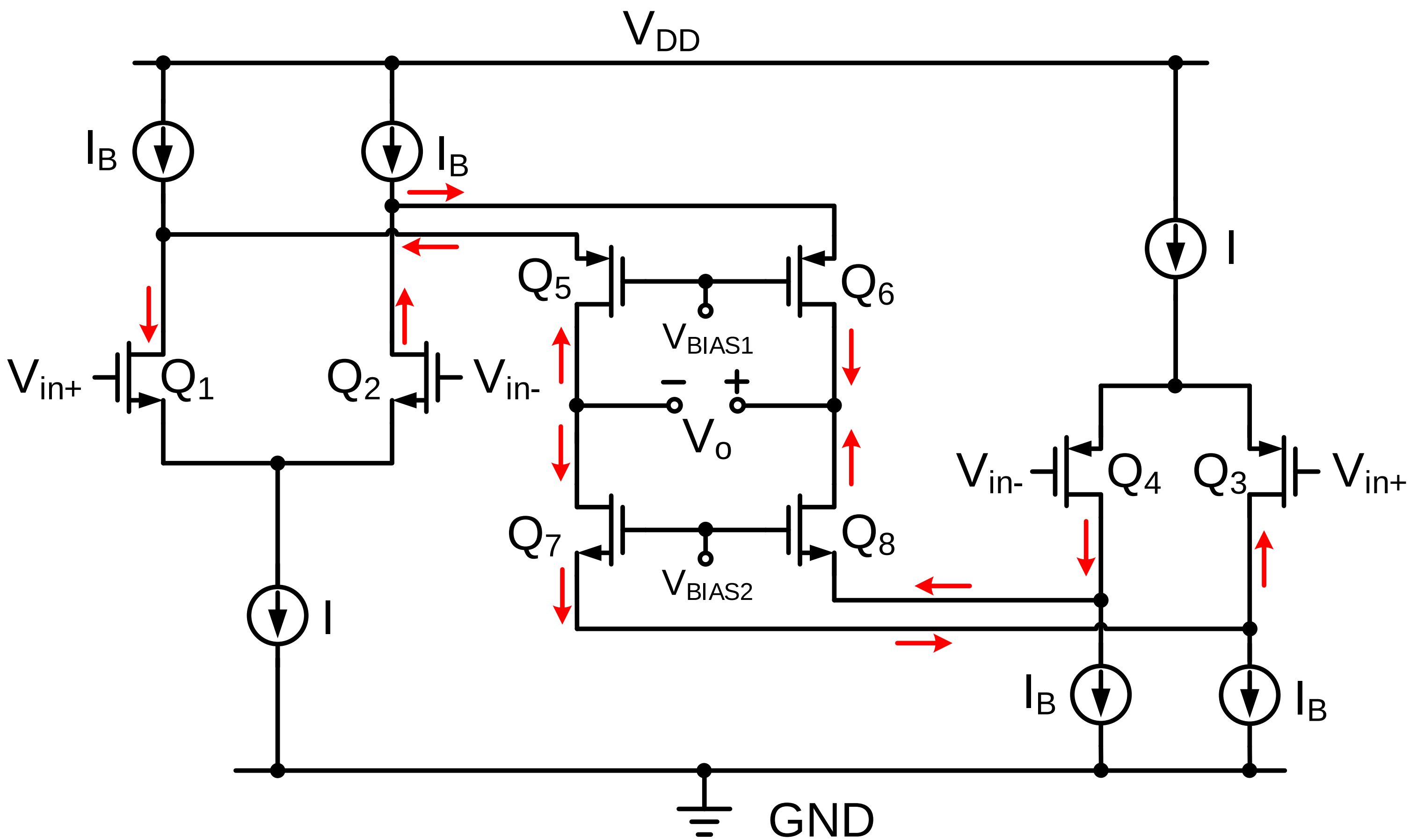
- ◆ A common choice, $n = 1,\ V_{out} > 2V_{eff}$

Folded-Cascode with Wide-Swing Current Mirror



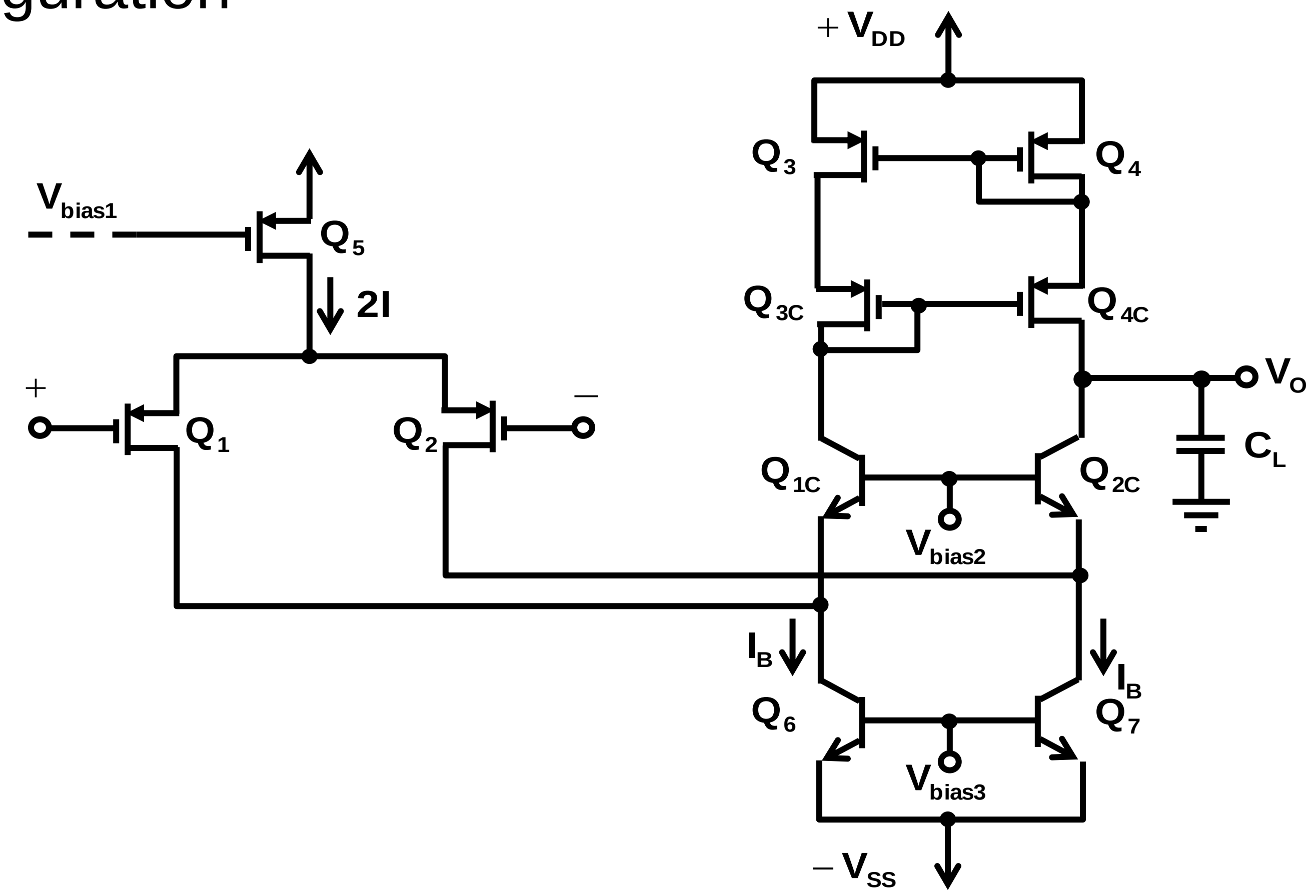
Folded-Cascode with Rail-to-Rail Input Operation

- Increased input common-mode range, rail-to-rail or even larger
- Voltage gain, if $g_{m1}=g_{m3}=G_m$
  - $A = (g_{m1}+g_{m3})R_o = 2G_mR_o$ for middle V_{ICM}
  - $A = g_{m1}R_o$ for high V_{ICM}
  - $A = g_{m3}R_o$ for low V_{ICM}



BiCMOS Folded-Cascode OPAMP

- Configuration



- When it is necessary to drive a resistive load, a low resistance output buffer is needed

BiCMOS Folded-Cascode OPAMP (Cont.)

- The largest nondominant pole is usually generated at the emitter nodes of Q_{1C} and Q_{2C}

$$\omega_{p2} \approx \frac{1}{R_{1C}C_{p1}} \approx \frac{g_{m1C}}{C_{p1}}, \text{where } R_{1C} \approx R_{elc} \parallel r_{O(Q16)} \parallel r_{O(Q1)} \approx R_{elc} = \frac{1}{g_{mlc}}$$

  - ◆ The transconductance of BJT can be much larger than that of CMOS

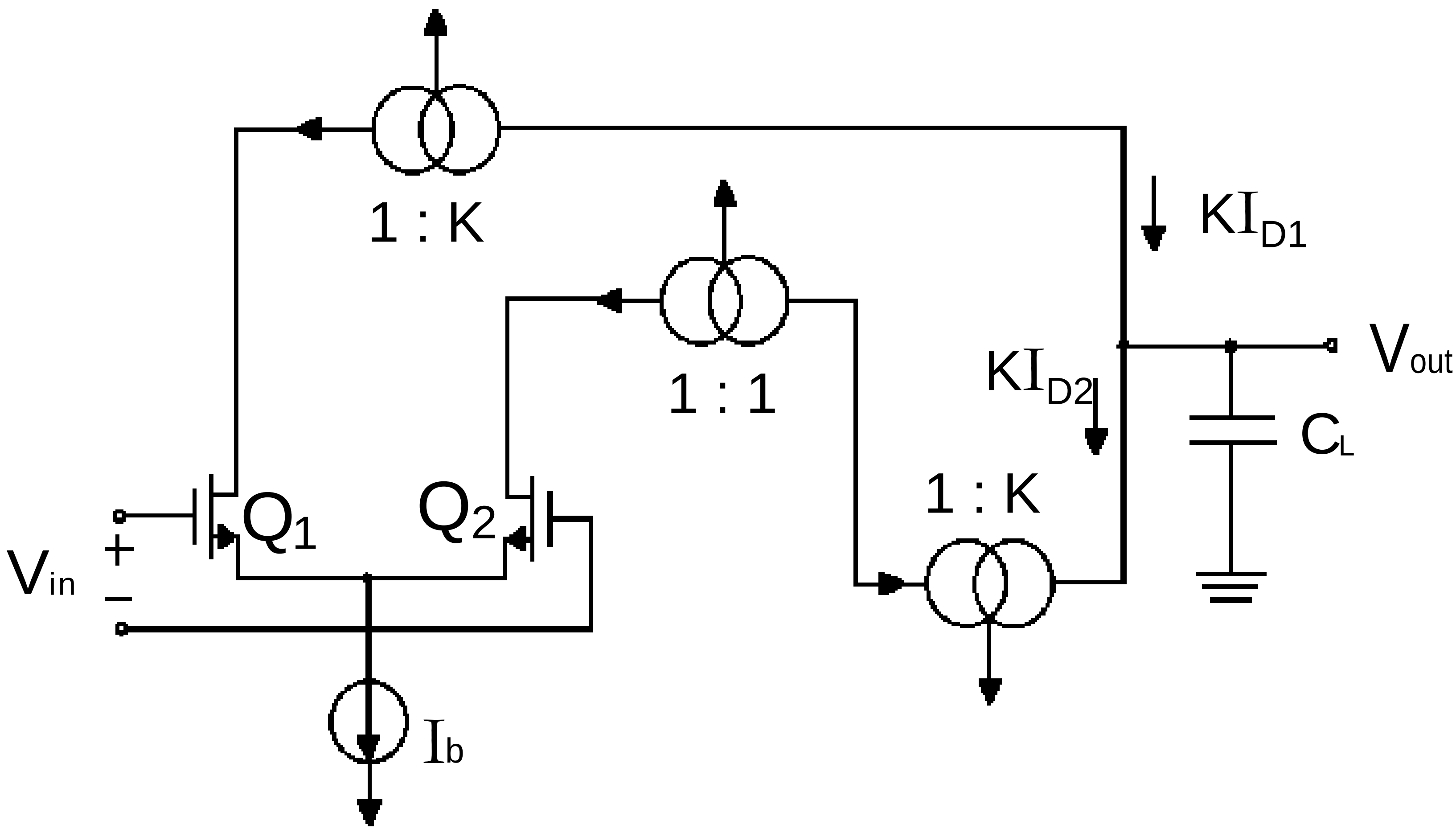
    $\Rightarrow$ ω_{P2} can be increased

    $\Rightarrow$ ω_u can be increased while enough phase margin is maintained

    $\Rightarrow$ Wider bandwidth than that of CMOS folded-cascode OPAMP

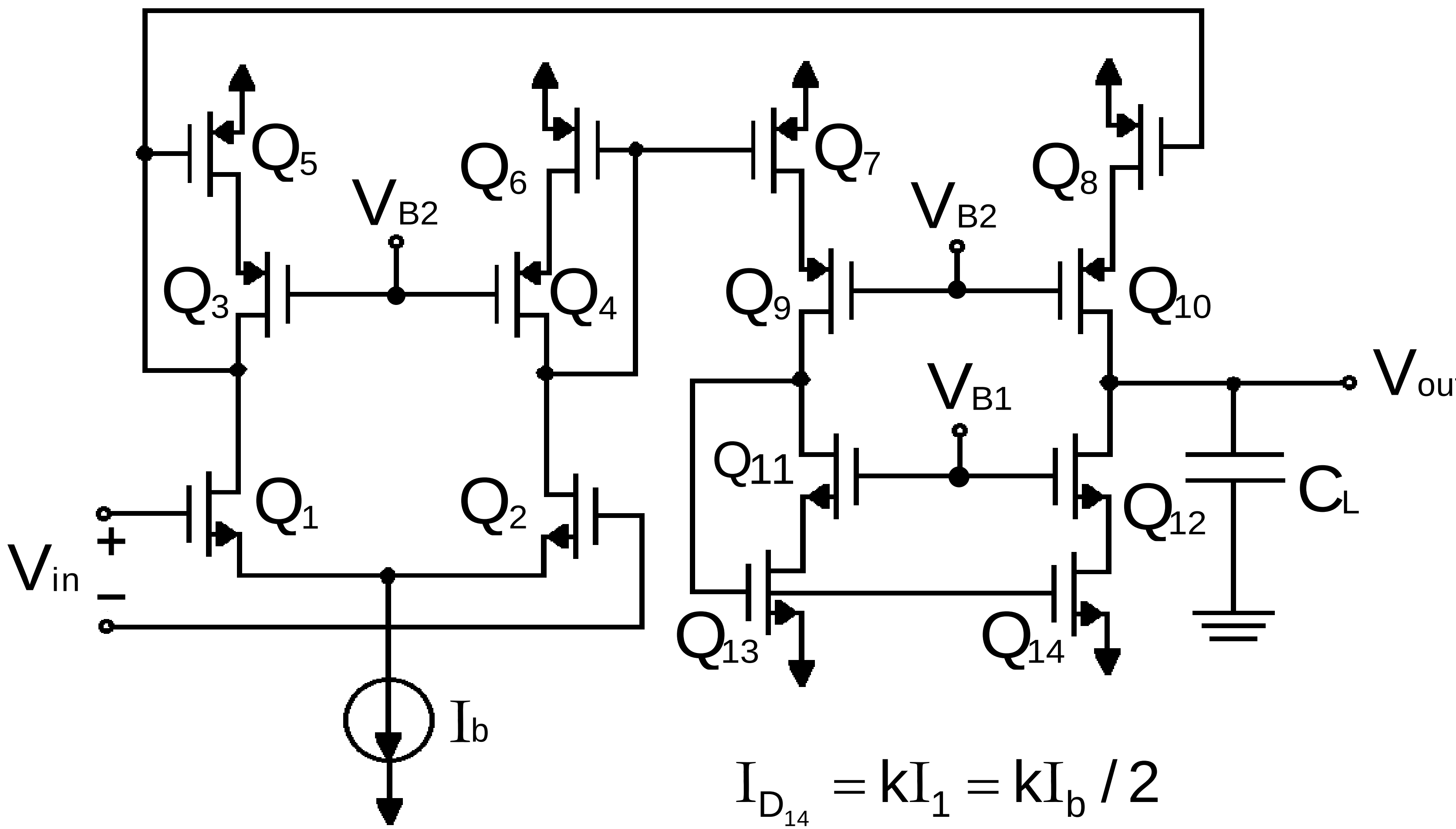
Current Mirror OPAMP

- A simplified current-mirror OPAMP



Current Mirror OPAMP (Cont.)

- A current-mirror OPAMP with wide-swing cascode current mirrors



$$I_{D_{14}} = kI_1 = kI_b / 2$$

Current Mirror OPAMP (Cont.)

$$\frac{(W/L)_8}{(W/L)_5} = k, \quad \frac{(W/L)_7}{(W/L)_6} = 1, \quad \frac{(W/L)_{12}}{(W/L)_{11}} = \frac{(W/L)_{14}}{(W/L)_{13}} = k$$

$$A_V = \frac{V_{out}(s)}{V_{in}(s)} = k\,g_{m1} z_L(s) = k\,g_{m1}(r_{out} // C_L) = \frac{k\,g_{m1} r_{out}}{1 + s\,r_{out} C_L}$$

(where k is the current gain from Q_5 to Q_8)

Unity-gain freq. (ω_t): $\omega_t = \frac{k g_{m1}}{C_L} = \frac{k\sqrt{2 I_{D1} \mu_n c_{ox} (W/L)_1}}{C_L}$

Total OPAMP current I_{total} = (3 + K) I_{D1}

$$\omega_t = \frac{k\sqrt{2\left(\frac{I_{total}}{3+k}\right)\mu_n C_{ox}(W/L)_1}}{C_L} = \frac{k}{\sqrt{3+k}} \frac{\sqrt{2 I_{total} \mu_n C_{ox} (W/L)_1}}{C_L}$$

$k \uparrow \Longrightarrow \omega_t \uparrow$ for a specified power dissipation

The important nodes for determining the nondominant pole are the drain of Q_1, primarily, and the drains of Q_2 and Q_9, secondly.

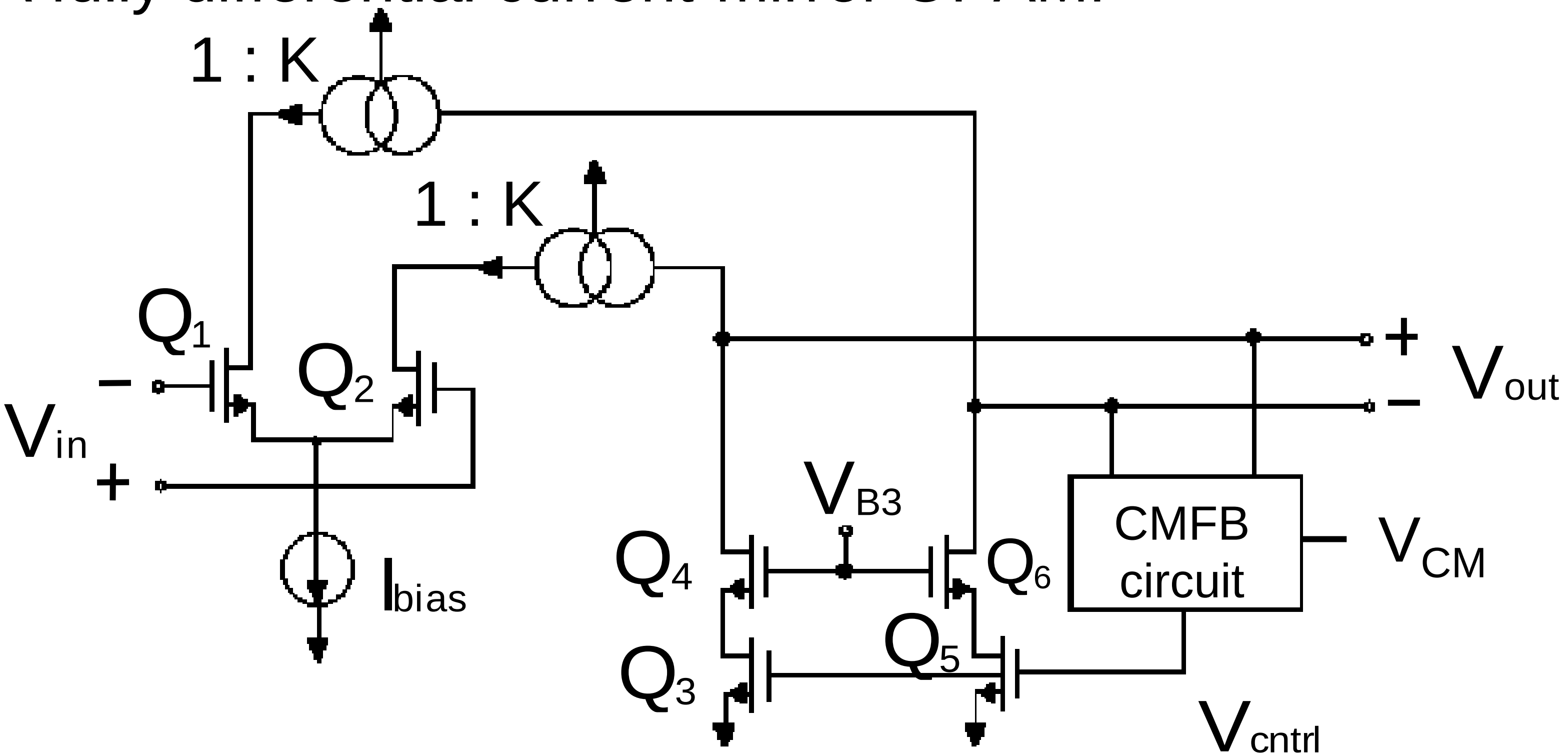
Increasing K increases the capacitances of these nodes while also increasing the equivalent resistances.

Current Mirror OPAMP (Cont.)

- As a result, the equivalent Pole-2 moves to lower frequencies. If K is increased too much, an increase in C_L will be required to keep ω_t below the frequency of the equivalent second pole to maintain stability. Thus, increasing K decreases the bandwidth when the equivalent second poles dominate.
- In the case where the load capacitance is small, the equivalent second pole will limit the unity-gain frequency of the opamp, and if it is very important that speed is maximized, K might be taken as small as one.
- From experience it has been found that a reasonable compromise for a general-purpose opamp might be to let K = 2.
- Slew rate ($SR = \frac{kI_b}{C_L}$)→ Larger compared to folded-cascode
- Due primarily to the larger bandwidth and slew rate, the current-mirror OPAMP is usually preferred over a folded-cascode OPAMP.
- However, it will suffer from larger thermal noise when compared to a folded-cascode amplifier because its input transistors are biased at a lower current level and therefore have a smaller transconductance.

Alternative Fully Differential OPAMPs

- A fully differential current-mirror OPAMP

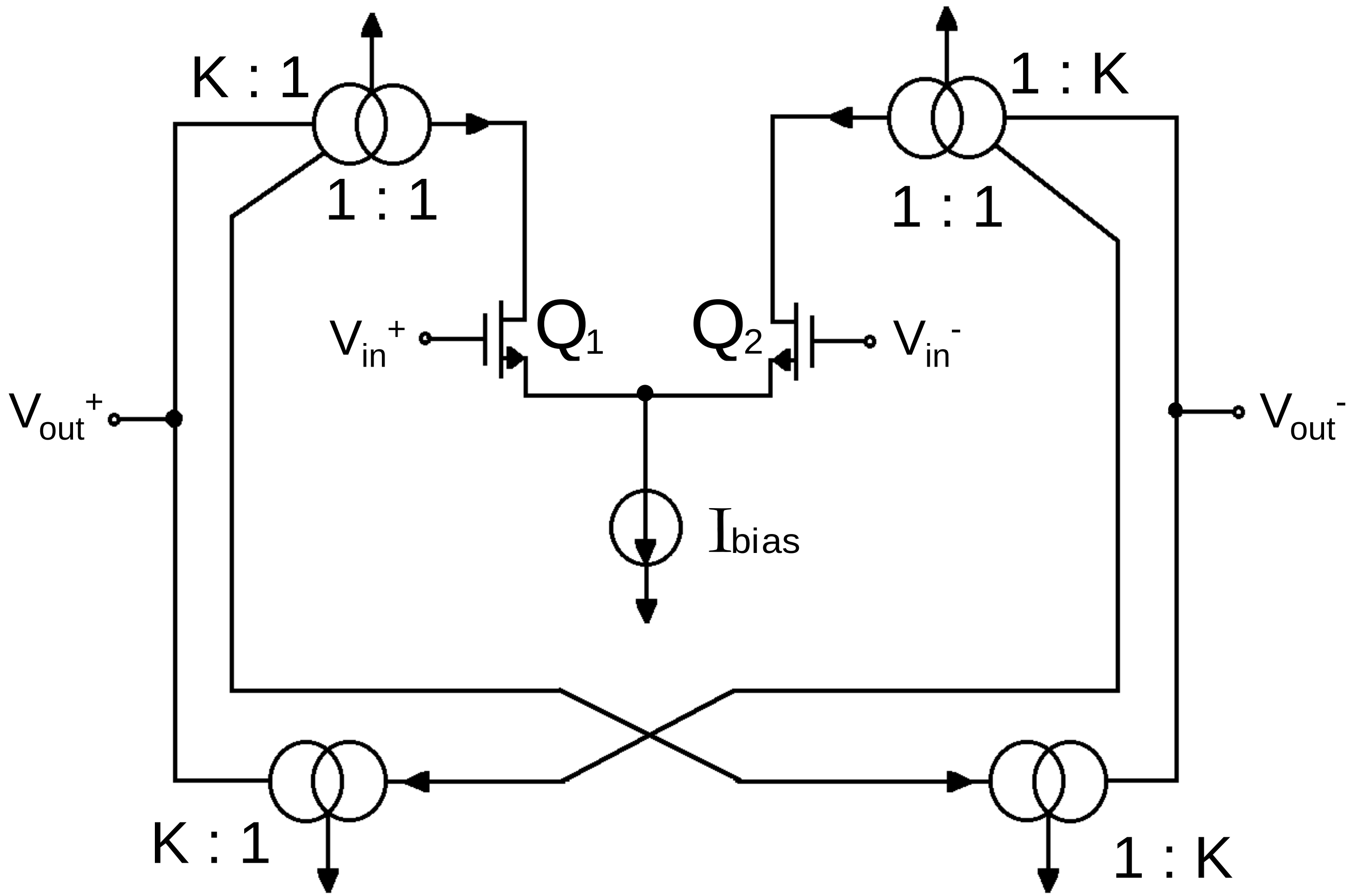


◆ n-channel input { high gain / lower thermal noise

◆ p-channel input { wide bandwidth / low 1/f noise

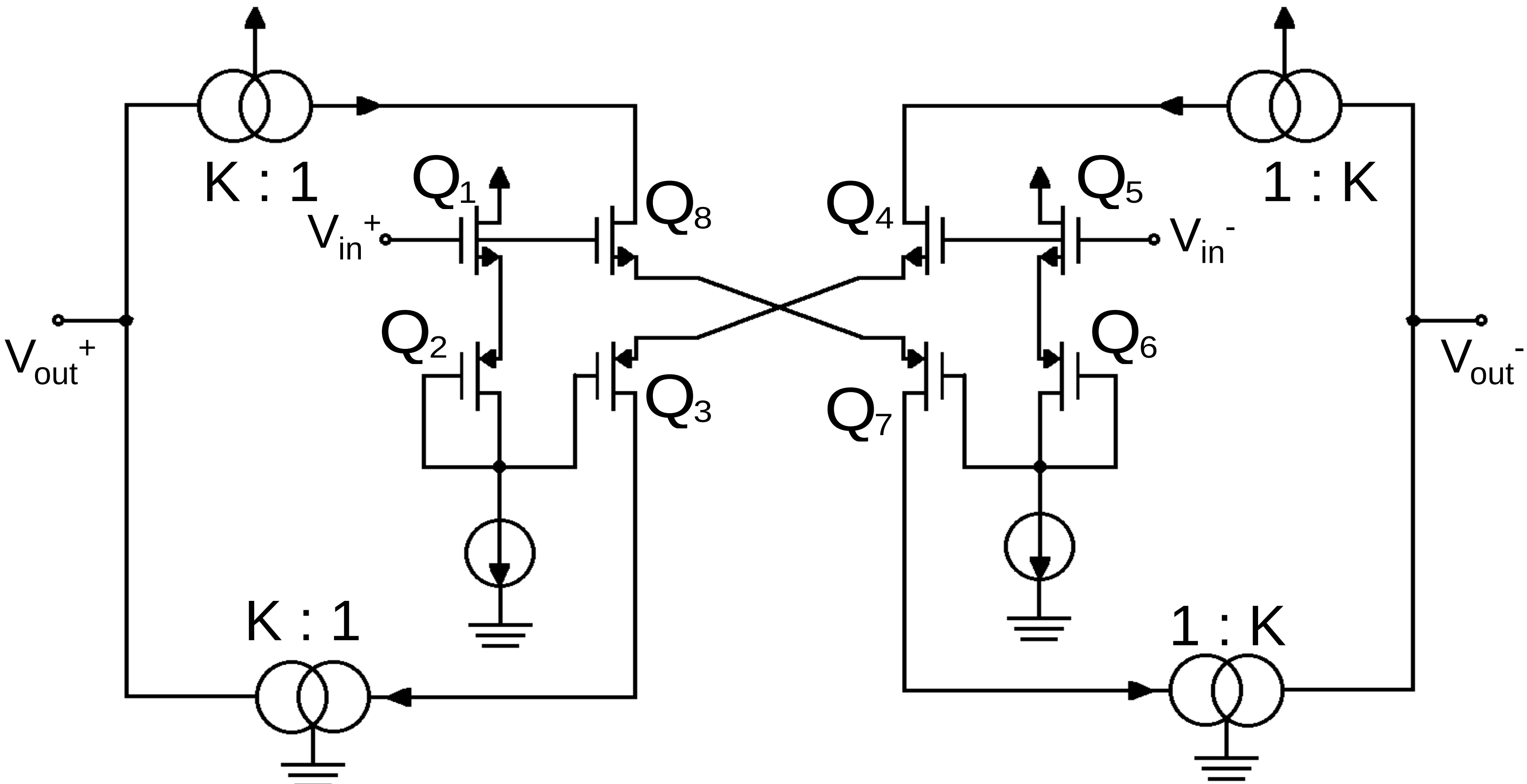
Alternative Fully Differential OPAMPs (Cont.)

- A fully differential OPAMP with bidirectional output drive



Alternative Fully Differential OPAMPs (Cont.)

- A class AB fully differential OPAMP

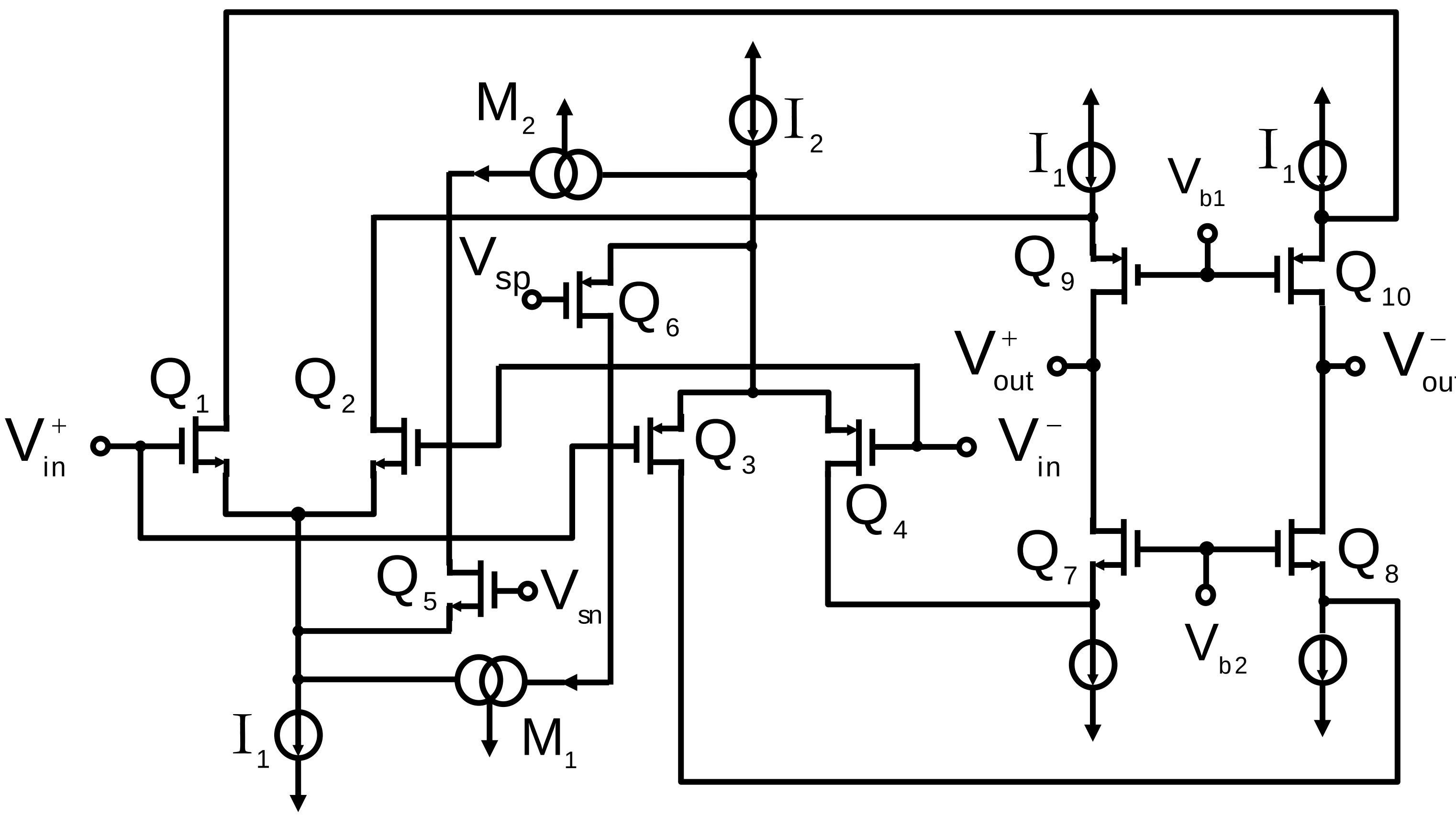


Alternative Fully Differential OPAMPs (Cont.)

- The advantage of the input stage in this OPAMP is that during slew-rate limiting, one differential pair will turn off, but the total current in the other differential pair will dynamically increase substantially.
- The disadvantage of this design is that the level-shift circuitry required at the input increases the noise and adds additional parasitics, which contribute to the equivalent second pole. In addition, the common-mode range of the input must remain at least $2V_t + 3V_{eff}$ above the lower power supply (and typically higher for the slew-rate performance to be maintained). This is a major problem when 5-V power supplies are being used, and it effectively eliminates this design from consideration for use with 3.3-V power supply voltages. However, for applications where the power-supply voltages are large, the load capacitances are large, and the slew rate is very important, this approach is quite reasonable.

Alternative Fully Differential OPAMPs (Cont.)

- A fully differential OPAMP composed of two single-ended output current-mirror OPAMPs
  - ☆ Reading Assignment p.284 ~ 286
- An OPAMP having rail-to-rail common-mode voltage range

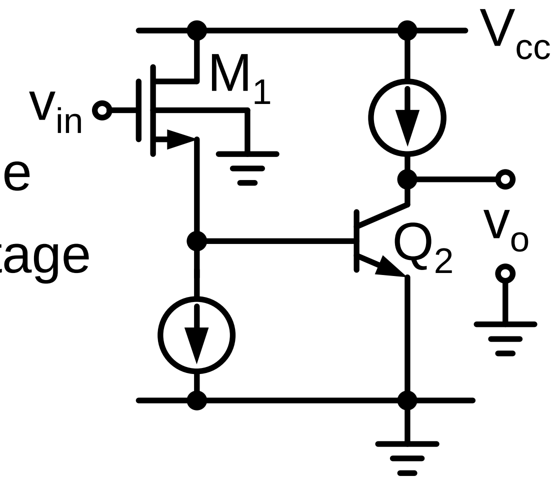


Alternative Fully Differential OPAMPs (Cont.)

- When the input common-mode voltage range is close to one of the power-supply voltages, one of the input differential pairs will turn off, but the other one will remain active.
- In an effort to keep the OPAMP gain relatively constant during this time, the bias currents of the still-active differential pair are dynamically increased. M_1, M_2, Q_5, Q_6 are added for this purpose.
- With careful design, it has been reported that the transconductance of the input stage can be held constant to within 15% of its nominal value with an input common-mode voltage range as large as the difference between the power-supply voltages.

BiCMOS Amplifiers

- Source follower–common emitter
  - $R_i = \infty$



$$A_V = \frac{r_{\pi 2}}{\frac{1}{g_m} + r_{b2} + r_{\pi 2}} \cdot \frac{V_A}{V_T}$$

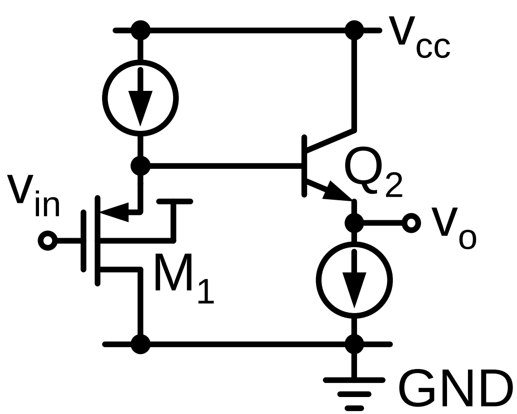
; V_A is Early voltage

; V_T is thermal voltage

  - Advantages:
    - Infinite input resistance
    - Higher gain than MOS common source AMP
  - Drawback: pole at $\omega_p = \dfrac{1}{[(\frac{1}{g_{m1}} + r_{b2}) // r_{\pi 2}]C_{\pi 2}} = \dfrac{g_{m1}}{C_{\pi 2}}$

(Assuming $r_{b2} << 1/g_{m1} < r_\pi$)

Source Follower-Emitter Follower

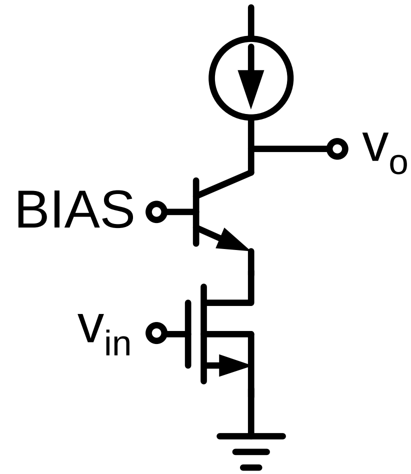


$$;\frac{V_o}{V_i}=1, R_i=\infty, R_o=\frac{1}{g_{m2}}$$

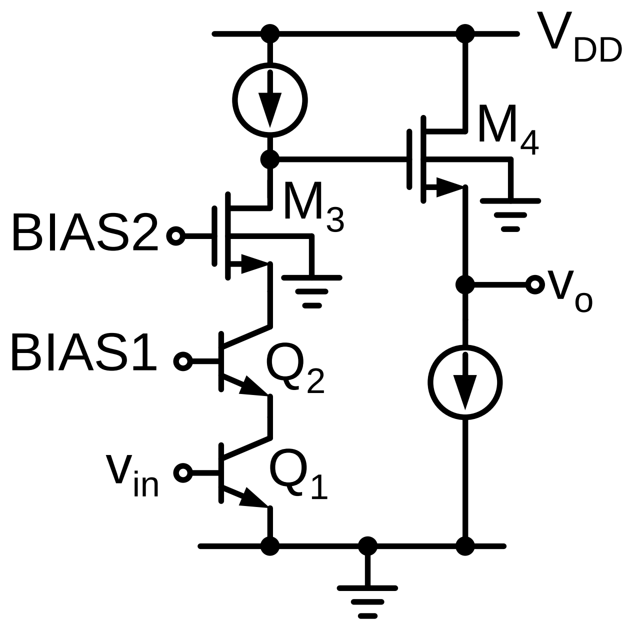
- Note: use PMOS input for better output swing no back-gate effect(N-well process).
  - Advantages: Infinite input resistance
    Low output resistance
  - Disadvantages: pole at $\frac{g_{m1}}{C_{\pi 2}}$

Cascode Amplifiers

- Cascode to increase R_o
  - $R_i = \infty$
  - $A_v = g_{m1}(\beta r_{o2})$
- Advantages:
  - Infinite input resistance
  - High gain
  - Good dynamics(2nd pole at f_T of NPN)
- The above circuit chooses BJT on MOSFET.
  - Higher R_o
  - Higher R_i
  - Wider bandwidth



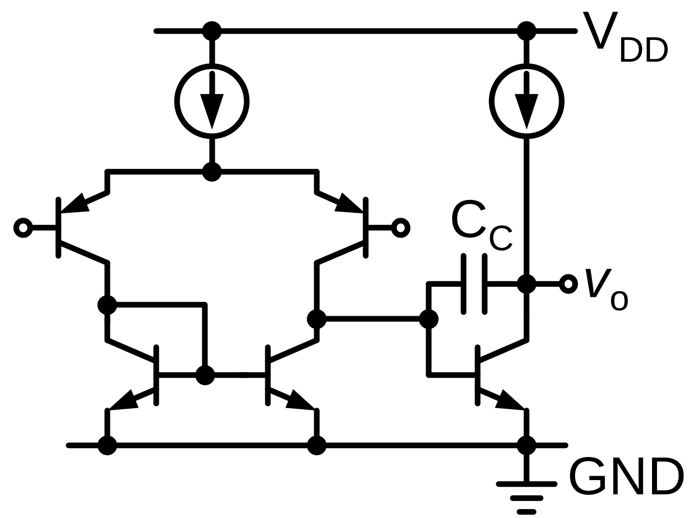
Double Cascode Amplifiers



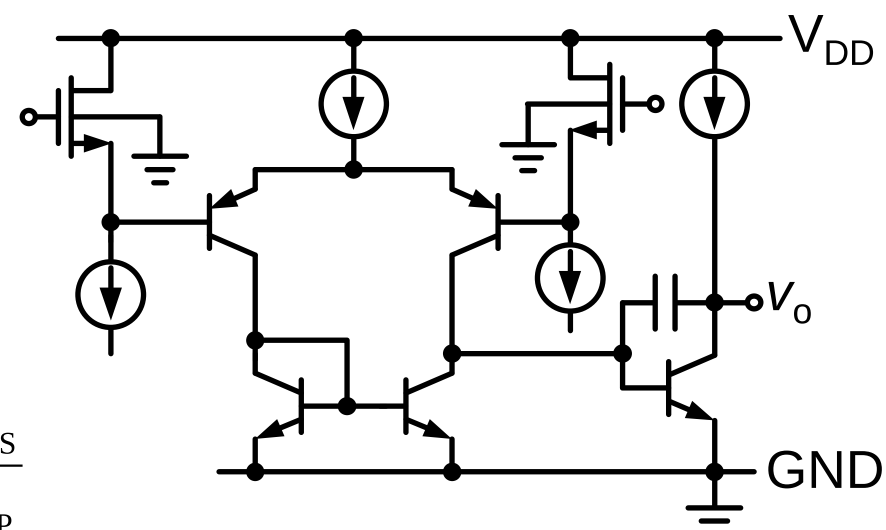
- $R_i = r_{\pi 1}$
- $A_v = g_{m1}(g_{m3}r_{o3})(\beta r_{o2})$
- Advantage: extremely high gain
  (gain of more than 10^6 achievable)
- Note: A source follower can be added
  if any resistive load is to be driven.

OPAMP Circuits

- Bipolar OPAMP



- Source follower input bipolar OPAMP



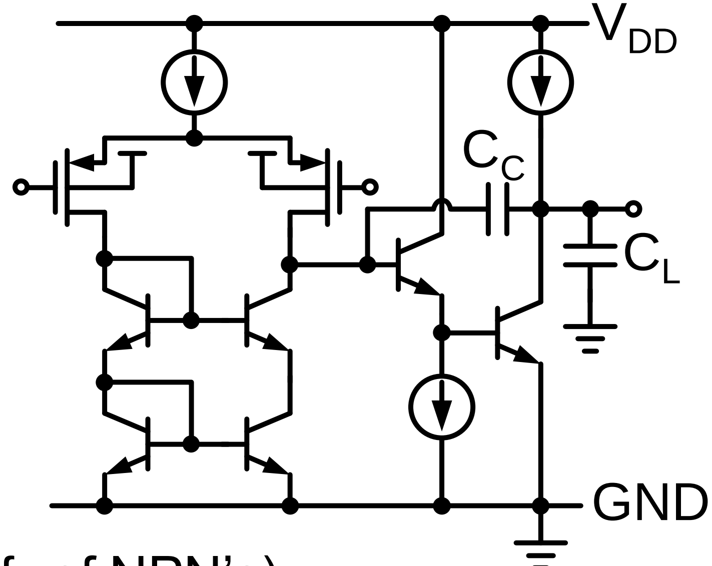
  - Drawback:
    - Additional pole at $\frac{g_{m,MOS}}{C_{\pi,PNP}}$

BiCMOS Differential Amplifier

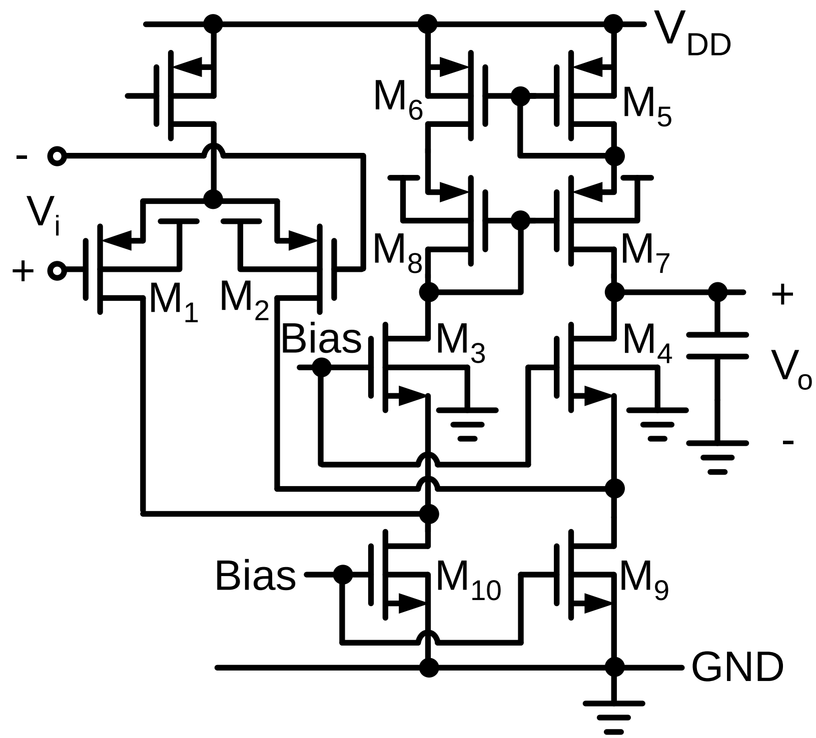
- For high input resistance and zero input bias current
  - Use MOSFET input
- For low offset
  - Use BJT input.
- Usually, the subsequent stages utilize BJT to obtain a wide bandwidth.

- BiCMOS OPAMP
  - $R_i = \infty$
  - $W_u < P_2 = \frac{G_{m2}}{C_L}$
  - Advantages: high W_u

    (higher poles at f_T of NPN's)



CMOS Folded-Cascode OPAMP



$$P_1 \approx \frac{-1}{R_0 C_L}$$

$$P_2 \approx \frac{-g_{m3}}{C_S}$$

; C_S is the total cap. at the source of the common gate transistors

$$\omega_t \approx \frac{g_{m1}}{C_L}$$