類比電路設計(3349) - 2004

Stability and Frequency Compensation

Ching-Yuan Yang

National Chung-Hsing University
Department of Electrical Engineering

Overview

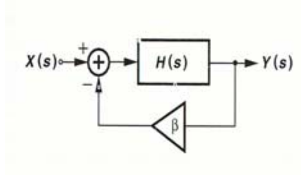
- Reading
  B. Razavi Chapter 10.

- Introduction
  In this lecture, we deal with the stability and frequency compensation of linear feedback systems to the extent necessary to understand design issues of analog feedback circuits. Beginning with a review of stability criteria and the concept of phase margin, we study frequency compensation, introducing various techniques suited to different op amp topologies. We also analyze the impact of frequency compensation on the slew rate of two-stage op amps.

Basic negative-feedback system

- General considerations

  - Close-loop transfer function: $\frac{Y}{X}(s) = \frac{H(s)}{1+\beta H(s)}$
    if $\beta H(s = j\omega 1) = -1$, the gain goes to infinity, and the circuit can amplify is own noise until it eventually begins to oscillate at frequency $\omega 1$.



  - Barkhausen's Criteria:

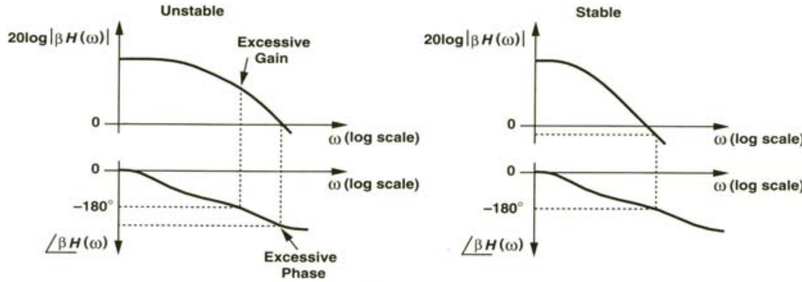
$$|\beta H(s = j\omega 1)| = 1$$
$$\angle \beta H(s = j\omega 1) = -180^\circ .$$

    The total phase shift around the loop at ω_1 is 360° because *negative* feedback itself introduces 180° of phase shift. The 360° phase shift is necessary for oscillation since the feedback must add *in phase* to the original noise to allow oscillation buildup. By the same token, a loop gain of unity (or greater) is also required to enable growth of the oscillation amplitude.

Bode diagram of loop gain

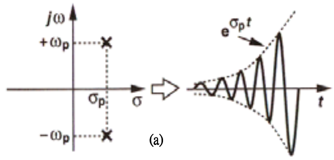
- A negative feedback system may oscillate at $\omega 1$ if
  (1) the phase shift around the loop at this frequency is so much that the feedback becomes *positive*.
  (2) the loop gain is still enough to allow signal buildup.

Unstable system *Stable system*



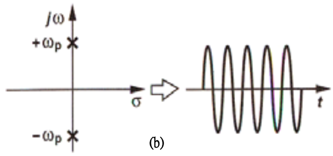
Time-domain response

Root locus:

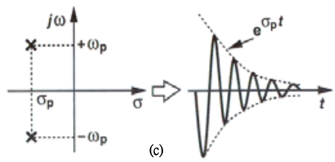


Pole frequency $s_P = j\omega_P + \sigma_P$

$\sigma_P > 0$, unstable with growing amplitude



$\sigma_P = 0$, unstable with constant-amplitude oscillation



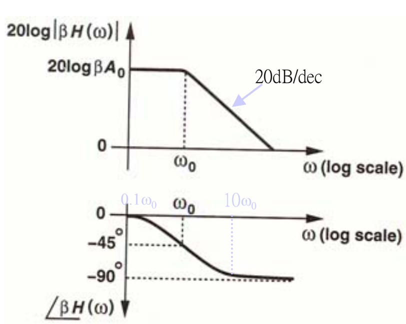
$\sigma_P < 0$, stable

Bode plots of one-pole system

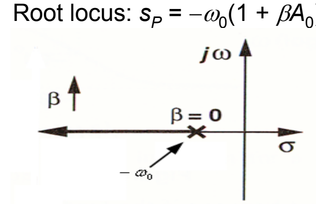
- Assuming $H(s) = A_0/(1 + s/\omega_0)$, β is less than or equal to unity and does not depend on the frequency, we have

$$\frac{Y}{X}(s) = \frac{H(s)}{1+\beta H(s)} = \frac{\frac{A_0}{1+\beta A_0}}{1+\frac{s}{\omega_0(1+\beta A_0)}}$$

- Bode plots of loop gain:



Root locus: $s_P = -\omega_0(1 + \beta A_0)$



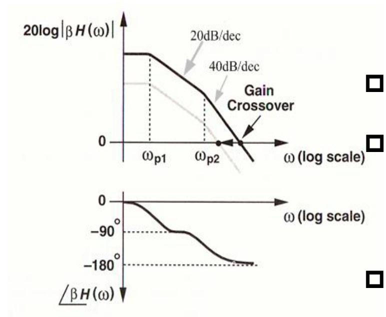
A single pole cannot contribute a phase shift greater than 90o and the system is unconditionally stable for all non-negative valuesof β.

Bodes plots of two-pole system

- Assuming the open-loop transfer function

$$H(s)=\frac{A_0}{\left(1+\frac{s}{\omega_{p1}}\right)\left(1+\frac{s}{\omega_{p2}}\right)}$$

- Bode plots of loop gain:



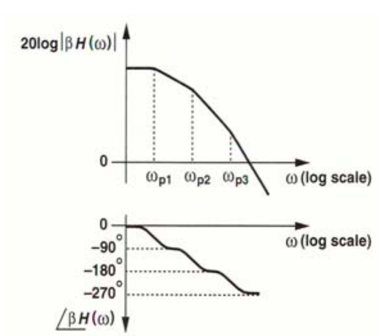
- The system is stable because $|\beta H|$ drops to below unity at a frequency for which $\angle\beta H < -180°$.
- To reduce the amount of feedback, we decrease β, obtaining the gray magnitude plot in the figure. For a logarithmic vertical axis, a change in β translates the magnitude plot vertically. Note that the phase plot does not change.
- The stability is obtained at the cost of weaker feedback.

Bodes plots of three-pole system

- Assuming the open-loop transfer function

$$H(s)=\frac{A_0}{\left(1+\frac{s}{\omega_{p1}}\right)\left(1+\frac{s}{\omega_{p2}}\right)\left(1+\frac{s}{\omega_{p3}}\right)}$$

- Bode plots of loop gain:

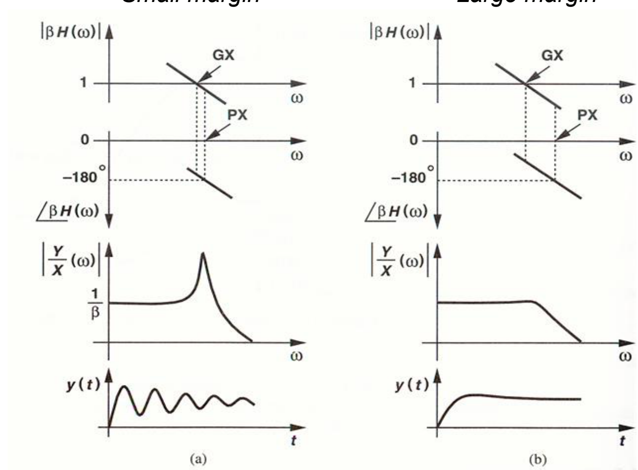


If the feedback factor β decreases, the circuit becomes more stable because the gain crossover moves toward the origin while the phase crossover remains constant.

Phase margin

- Close-loop frequency and time response

Small margin *Large margin*



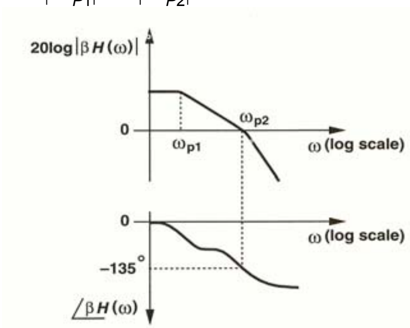
Phase margin (cont'd)

- *Phase margin* (PM) is defined as

$$PM = 180^{o} + \angle\beta H(\omega = \omega_1)$$

  where ω_1 is the gain crossover frequency.

- Example
  A two-pole feedback system is designed such that $|\beta H(\omega = \omega_{P2})| = 1$ and $|\omega_{P1}| << |\omega_{P2}|$.



Since $\angle\beta H$ reaches -135^{o} at $\omega = \omega_{P2}$, the phase margin is equal to 45^{o}.

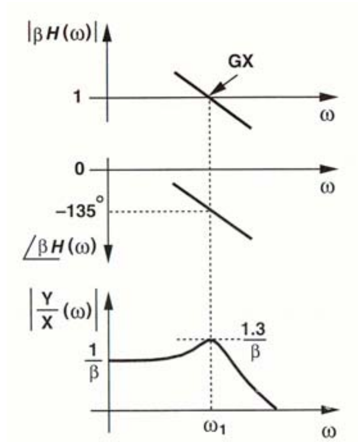
How much phase margin is adequate?

- For PM = 45°, at the gain crossover frequency $\angle\beta H(\omega_1) = -135°$ and $|\beta H(\omega_1)| = 1$, yielding $\frac{Y}{X} = \frac{H(j\omega_1)}{1+1\cdot\exp(-j135°)} = \frac{H(j\omega_1)}{0.29-0.71j}$

  It follows that $\left|\frac{Y}{X}\right| = \frac{1}{\beta}\cdot\frac{1}{|0.29-0.71j|} \approx \frac{1.3}{\beta}$

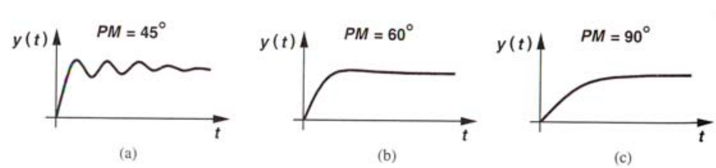
  The frequency response of the feedback system suffers from a 30% peak at $\omega = \omega_1$.

- Close-loop frequency response for 45° phase margin:



How much phase margin is adequate? (cont'd)

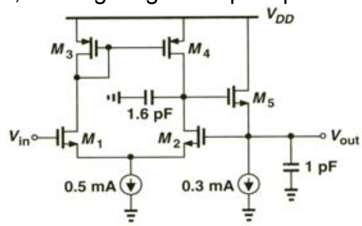
- Close-loop time response for 45°, 60°, and 90° phase margin:

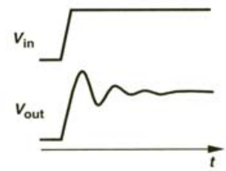


- For PM = 60°, $Y(j\omega_1)/X(j\omega_1) = 1/\beta$, suggesting a negligible frequency peaking. This typically means that the step response of the feedback system exhibits little ringing, providing a fast settling. For greater phase margins, the system is more stable but the time response slows down. Thus, PM = 60° is typically considered the optimum value.
- The concept of phase margin is well-suited to the design of circuits that process *small* signals. In practice, the large-signal step response of feedback amplifiers does not follow the illustration of the above figure. For large-signal applications, time-domain simulations of the close-loop system prove more relevant and useful than small-signal ac computations of the open-loop amplifier.

How much phase margin is adequate? (cont'd)

- Example:
  Unity-gain buffer: *PM* ≈ 65°, unity-gain frequency = 150 MHz.
  However, the large-signal step response suffers from significant ringing.





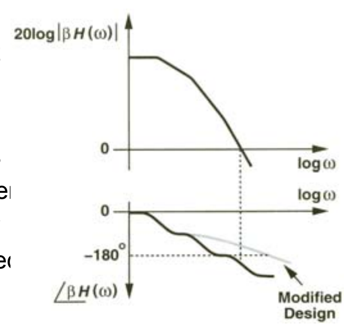
- The large-signal step response of feedback amplifiers is not only due to slewing but also because of the nonlinear behavior resulting from large excursions in the bias voltages and currents of the amplifier. Such excursions in fact cause the pole and zero frequencies to *vary* during the transient, leading to a complicated time response. Thus, for large-signal applications, time-domain simulations of the close-loop system prove more relevant and useful than small-signal ac computations of the open-loop amplifier.

Frequency compensation

- Typical op amp circuits contain many poles. For this reason, op amps must usually be "compensated," that is, the open-loop transfer function must be modified such that the closed-loop circuit is stable and the time response is well-behaved.
- Stability can be achieved by minimizing the overall phase shift, thus pushing the phase crossover *out*.

  Discussion:
  - This approach requires that we attempt to minimize the number of poles in the signal path by proper design.
  - Since each additional stage contributes at least one pole, this means the number of stages must be minimized, a remedy that yields low voltage gain and/or limited output swings.

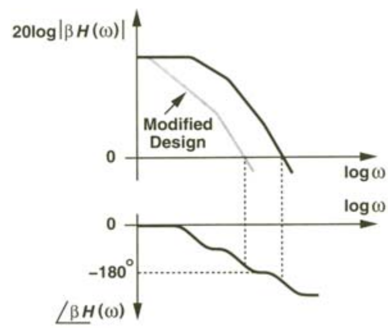
Moving PX out



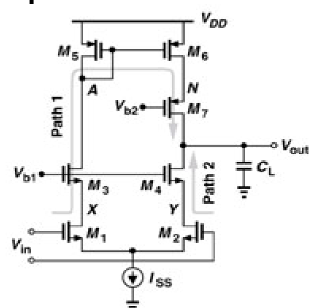
Frequency compensation (cont'd)

- Stability can be achieved by dropping the gain thereby pushing the gain crossover *in*.
  Discussion:
  - This approach retains the low frequency gain and the output swings but it reduces the bandwidth by forcing the gain to fall at lower frequencies.

Moving GX in

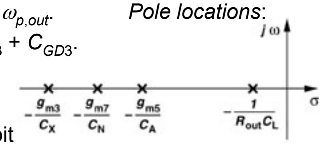


Telescopic op amp with single-ended output



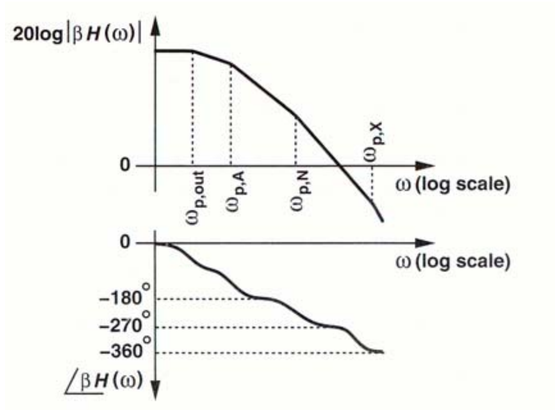
- Determine the poles of the circuit:
  We identify a number of poles in the signal paths: path 1 contains a high-frequency pole at the source of M_3, a mirror pole at node A, and another high-frequency pole at the source of M_7, whereas path 2 contains a high-frequency pole at the source of M_4. The two paths share a pole at the output.
  - Dominant pole: the closest to the origin.
    $\omega_{p,out} = 1/(R_{out}C_L)$, usually sets the open-loop 3-dB bandwidth.
  - Nondominant poles:
    $\omega_{p,A} = g_{m5}/C_A$, the closest pole to the origin after $\omega_{p,out}$.
    (mirror pole) where $C_A = C_{GS5} + C_{GS6} + C_{DB5} + 2C_{GD6} + C_{DB3} + C_{GD3}$.
    $\omega_{p,N} = g_{m7}/C_N$, $\omega_{p,X} = g_{m3}/C_X = g_{m4}/C_Y = \omega_{p,Y}$.
    Since $g_m = 2I_D/|V_{GS} - V_{TH}|$, if M_4 and M_7 are designed to have the same overdrive, they exhibit the same transconductance. From square-law characteristics, we have $W_4/W_7 = \mu_p/\mu_n \approx 1/3$. Thus, nodes N and $X(Y)$ see roughly equal small-signal resistances but node N suffers from much more capacitance.

Pole locations:



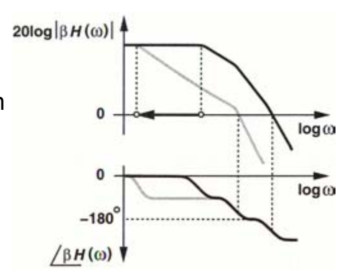
Telescopic op amp with single-ended output (cont'd)

- Bode plots of loop gain for op amp: using $\beta = 1$ for the worst case.
  - The mirror pole $\omega_{p,A}$ typically limits the phase margin because its phase contribution occurs at lower frequencies than that other nondominant poles.



Telescopic op amp with single-ended output (cont'd)

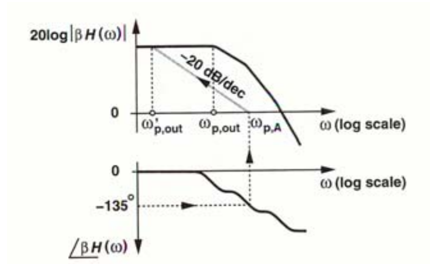
- Translating the dominant pole toward origin to compensate the op amp:
  - Assuming $\omega_{p,A} > 10\omega_{p,out}$, we must force the loop gain crossover point moves toward the origin. We can simply lower the frequency of the dominant pole ($\omega_{p,out}$) by increasing the load capacitance.
  - The key point is that the phase contribution of the dominant pole in the vicinity of the gain or phase crossover points is close to 90° and relatively independent of the location of the pole. That is, translating the dominant pole toward the origin affects the magnitude plot but not the critical part of the phase plot.

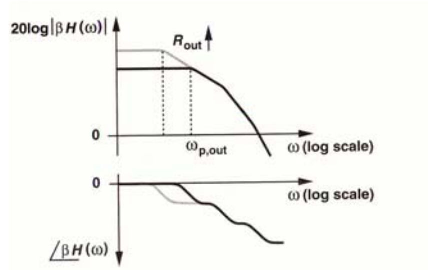


Telescopic op amp with single-ended output (cont'd)

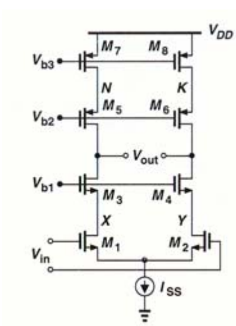
- How much the dominant pole must be shift down?

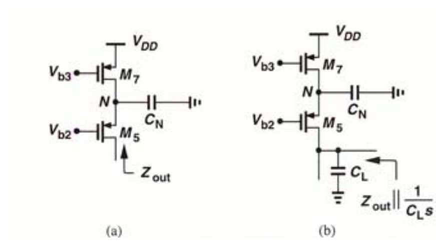
  Assume (1) the 2nd nondominant pole ($\omega_{p,N}$) is quite higher than the mirror pole so that the phase shift at $\omega = \omega_{p,A}$ is equal to -135^o, and (2) $PM = 45^o$.
  - $C_L \rightarrow (\omega_{p,out}/\omega'_{p,out})C_L$. The load capacitance must be increased by a factor of $\omega_{p,out}/\omega'_{p,out}$.
  - The unity-gain bandwidth of the compensated op amp is equal to the *frequency of the nondominant pole*.
  - To achieve a wideband in a feedback system employing an op amp, the first nondominant pole must be as far as possible.
- Although $\omega_{p,out} = (R_{out}C_L)^{-1}$, increasing R_{out} does *not* compensate the op amp. A higher R_{out} results in a greater gain, only affecting the low-frequency portion of the characteristics.





Fully differential telescopic op amp





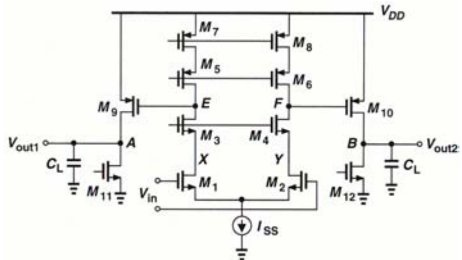
- $C_N = C_{GS5} + C_{SB5} + + C_{GD7} + C_{DB7}$

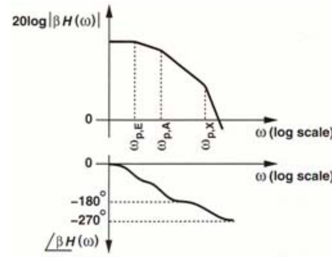
  $Z_N = r_{o7} \,\|\, (C_N\, s)^{-1}$, where body effect is neglected. We have

$$Z_{out} = (1+g_{m5}r_{o5})Z_N + r_{o5} \approx (1+g_{m5}r_{o5})\frac{r_{o7}}{r_{o7}C_N s+1} \quad \text{,and} \quad Z_{out}\left\|\frac{1}{C_L s}\right. = \frac{(1+g_{m5}r_{o5})r_{o7}}{[(1+g_{m5}r_{o5})r_{o7}C_L + r_{o7}C_N]s+1}$$

- A pole with $\tau = (1 + g_{m5}r_{o5})r_{o7}C_L + r_{o7}C_N$, where $(1 + g_{m5}r_{o5})r_{o7}C_L$ is simply due to the low-frequency output resistance of the cascode. The pole in the PMOS cascode is *merged* with the output pole, thus creating no *addition* pole. It merely lowers the dominant pole by a slight amount.
- For $g_m r_o$ >> 1 and $C_L > C_N$, then $\tau \approx g_{m5}r_{o5}r_{o7}C_L$.

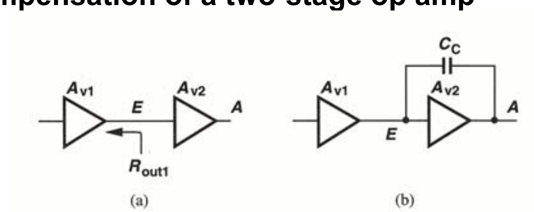
Compensation of two-stage op amps





- We identify three poles at X(or$Y$), E(or$F$) and A(or$B$).
  A pole at X(or$Y$) lies at relatively high frequencies. Since the small-signal resistance seen at E is quite high, even the capacitances of M_3, M_5 and M_9 can create a pole relatively close to the origin. At node A, the small-signal resistance is lower but the value of C_L may be quite high. Consequently, the circuit exhibits *two* dominant poles.
- One of the dominant poles must be moved toward the origin so as to place the gain crossover well below the phase crossover. If the magnitude of $\omega_{p,E}$ is to be reduced, the available bandwidth is limited to approximately $\omega_{p,A}$, a low value. Furthermore, this required dominant pole translates to a very large compensation capacitor.

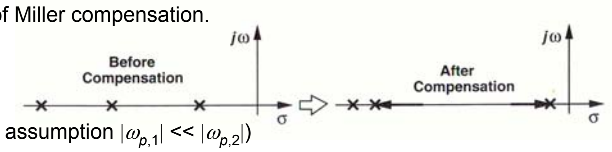
Miller compensation of a two-stage op amp



- In a two-stage amp as shown in Fig.(a), the first stage exhibits a high output impedance and the second stage provides a moderate gain, thereby providing a suitable environment for Miller multiplication of capacitors.
- In Fig.(b), we create a large capacitance at E, the pole is $\omega_{p,E} = \dfrac{1}{R_{out1}[C_E + (1 + A_{v2})C_C]}$
  As a result, a low-frequency pole can be established with a moderate capacitor value, saving considerable chip area.
- In addition to lowering the required capacitor value, Miller compensation entails a very important property: it *moves* the output pole *away* from the origin. (*pole splitting*)

Miller compensation of a two-stage op amp (cont'd)

- Pole splitting as a result of Miller compensation.

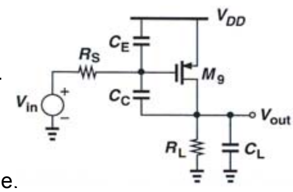


- Discussion
- Two poles: (based on the assumption $|\omega_{p,1}| << |\omega_{p,2}|$)

$$\omega_{p1} \approx \frac{1}{R_S\left[(1+g_{m9}R_L)(C_C+C_{GD9})+C_E\right]+R_L(C_C+C_{GD9}+C_L)}$$

$$\omega_{p2} \approx \frac{R_S\left[(1+g_{m9}R_L)(C_C+C_{GD9})+C_E\right]+R_L(C_C+C_{GD9}+C_L)}{R_S R_L\left[(C_C+C_{GD9})C_E+(C_C+C_{GD9})C_L+C_E C_L\right]}$$

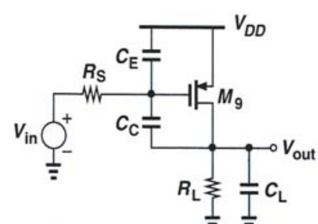
Simplified circuit of a two-stage op amp:
R_S = the output resistance of 1st stage.
$R_L = r_{o9} \,||\, r_{o11}$

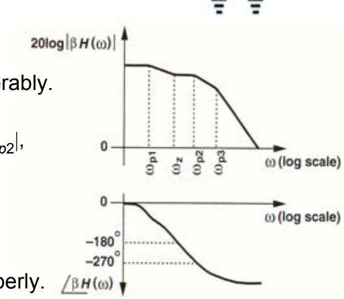


- For C_C = 0 and relatively large C_L, $\omega_{p,2} \approx 1/(R_L C_L)$.
  For $C_C \neq 0$ and $C_C + C_{GD9} >> C_E$, $\omega_{p,2} \approx g_{m9}/(C_E + C_L)$.
  Typically $C_E << C_L$, we conclude that Miller compensation increases the magnitude of the output pole ($\omega_{p,2}$) by a factor of $g_{m9}R_L$, a relatively large value.
- Miller compensation moves the interstage pole toward the origin and the output pole away from the origin, allowing a much greater bandwidth than that obtained by merely connect the compensation capacitor from one node to ground.

Miller compensation of a two-stage op amp (cont'd)

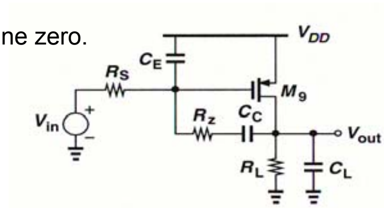
- Effect of right half plane zero
  - The circuit contains a right-half-plane zero at $\omega_z = g_{m9}/(C_C + C_{GD9})$ because $C_C + C_{GD9}$ forms a "parasitic" signal path from the input to the output.
  - A zero in the right hand plane contributes more phase shift, thus moving the phase crossover toward the origin. Furthermore, from Bode approximations, the zero slows down the drop of the magnitude, thereby pushing the gain crossover away from the origin. As a result, the stability degrades considerably.
  - For two-stage op amps, typically $|\omega_{p1}| < |\omega_z| < |\omega_{p2}|$, the zero introduces significant phase shift while preventing the gain from falling sufficiently. The right-half-plane zero is a serious issue because g_m is relatively small and C_C is chosen large enough to position the dominant pole properly.





Miller compensation of a two-stage op amp (cont'd)

- Addition of R_z to move the right hand plane zero.



  - ☐ The zero is given by .
    If $R_z \ge g_{m9}^{-1}$, then $\omega_z \le 0$.

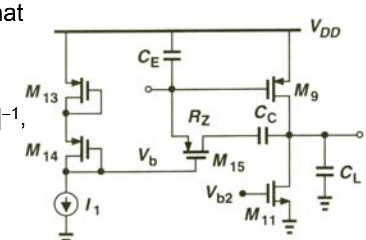
  - ☐ We may move the zero well into the left plane so as to cancel the first nondominant pole. That is
    $\frac{1}{C_C(g_{m9}^{-1}-R_z)} = \frac{-g_{m9}}{C_L+C_E} \quad \Rightarrow \quad R_z = \frac{C_L+C_E+C_C}{g_{m9}C_C} \approx \frac{C_L+C_C}{g_{m9}C_C}$, because $C_E << C_L + C_C$.

    - Drawbacks:
      1. It is difficult to guarantee the relationship of the above equation, especially if C_L is unknown or variable.
      2. The actual implementation of R_z is variable. R_z is typically realized by a MOS transistor in the triode region.

Miller compensation of a two-stage op amp (cont'd)

- Generation of V_b for proper temperature and process tracking.

  - ☐ If I_1 is chosen with respect to I_{D9} such that $V_{GS13} = V_{GS9}$, then $V_{GS15} = V_{GS14}$.
    Since $g_{m14} = \mu_p C_{ox}(W/L)_{14}(V_{GS14} - V_{TH14})$
    and $R_{on15} = [\mu_p C_{ox}(W/L)_{15}(V_{GS15} - V_{TH15})]^{-1}$,



    we have $R_{on15} = \frac{(W/L)_{14}}{g_{m14}(W/L)_{15}}$

    For pole-zero cancellation to occur, $\frac{(W/L)_{14}}{g_{m14}(W/L)_{15}} = \frac{C_L+C_C}{g_{m9}C_C}$

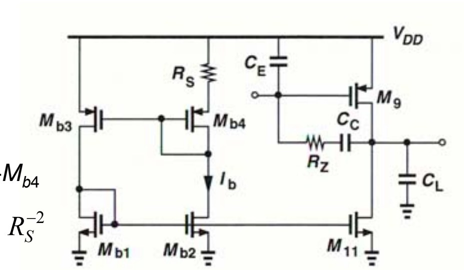
    and hence $(W/L)_{15} = \sqrt{(W/L)_{14}(W/L)_9}\sqrt{\frac{I_{D9}}{I_{D14}}\frac{C_C}{C_C+C_L}}$

    If C_L is constant, it can be established with reasonable accuracy because it contains only the *ratio* of quantities.

Miller compensation of a two-stage op amp (cont'd)

- Method of defining g_{m9} with respect to R_S.

Goal: $R_z = \dfrac{C_L + C_C}{g_{m9} C_C}$ .... *(A)*



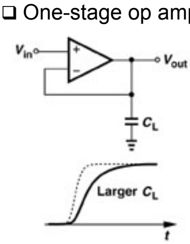
- ☐ The technique incorporates M_{b1}-M_{b4} along with R_S to generate $I_b \propto R_S^{-2}$

  Thus, $g_{m9} \propto \sqrt{I_{D9}} \propto \sqrt{I_{D11}} \propto R_S^{-1}$

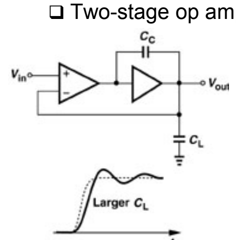
  Proper ratioing of R_Z and R_S therefore ensures (A) is valid even with temperature and process variations

Effect of increased load capacitance on step response

❑ One-stage op amps:

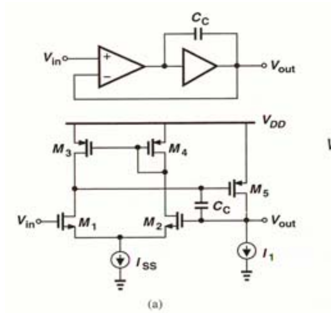


❑ Two-stage op amps:

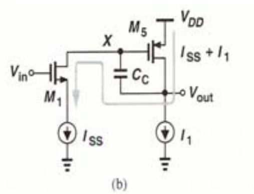


- In one-stage op amps, a higher load capacitance brings the *dominant* pole closer to the origin, *improving* the phase margin (albeit making the feedback system more overdamped).
- In two-stage op amps, since Miller compensation establishes the dominant pole at the output of the first stage, a higher load capacitance presented to the second stage moves the second pole toward the origin, *degrading* the phase margin.
- Illustrated in the figure is the step response of a unity-gain feedback amplifier, suggesting that the response approaches an oscillatory behavior if the load capacitance seen by the two-stage op amp increases.

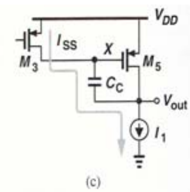
Slewing in two-stage op amps



❑ Positive slewing:

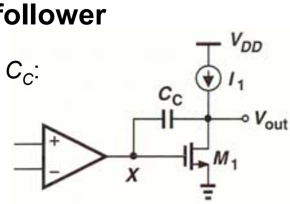


❑ negative slewing:

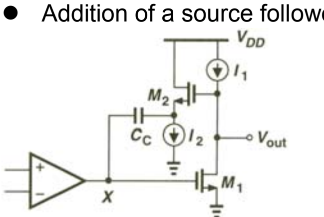


- The positive slew rate equals I_{SS}/C_C. During slewing, M_5 must provide two currents: I_{SS} and I_1. If M_5 is not wide enough to sustain $I_{SS} + I_1$ in saturation, then V_X drops significantly, possibly driving M_1 into the triode region.
- During negative slew rate, I_1 must support both I_{SS} and I_{D5}. For example, if $I_1 = I_{SS}$, then V_X rises so as to turn off M_5. If $I_1 < I_{SS}$, then M_3 enters the triode region and the slew rate is given by I_{D3}/C_C.

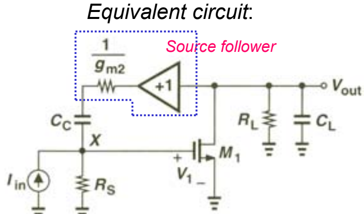
Compensation technique using a source follower



- Two-stage op amp with right half plane zero due to C_C:
- Addition of a source follower to remove zero:



Equivalent circuit:



☐ Since C_{GS} of M_2 is typically much less than C_C, we expect the right frequencies.

☐ $-g_{m1}V_1 = V_{out}\left(R_L^{-1} + C_L s\right)$ ⇨ $V_1 = \dfrac{-V_{out}}{g_{m1}R_L}(1 + R_L C_L s)$

and $\dfrac{V_{out} - V_1}{\dfrac{1}{g_{m2}} + \dfrac{1}{C_C s}} + I_{in} = \dfrac{V_1}{R_S}$, then

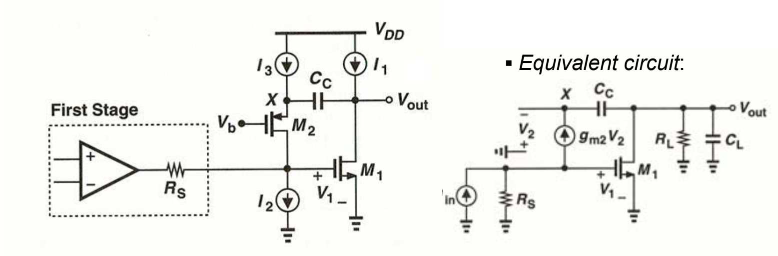
Zero in the *left* hand plane

$$\frac{V_{out}}{V_{in}} = \frac{-g_{m1}R_L R_S(g_{m2} + C_C s)}{R_L C_L C_C(1 + g_{m2}R_S)s^2 + [(1 + g_{m1}g_{m2}R_L R_S)C_C + g_{m2}R_L C_L]s + g_{m2}}$$

☐ Assume $\omega_{p1} << \omega_{p2}$, since typically $1 + g_{m2}R_S >> 1$, $(1 + g_{m1}g_{m2}R_L R_S)C_C >> g_{m2}R_L C_L$, we have

$\omega_{p1} \approx \dfrac{1}{g_{m1}R_L R_S C_C}$ and $\omega_{p2} \approx \dfrac{g_{m1}}{C_L}$ (Note ω_{p2}: $\dfrac{1}{R_L C_L} \rightarrow \dfrac{g_{m1}}{C_L}$)

Compensation technique using a CG stage



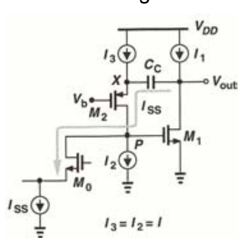
- The primary issue is that the source follower limits the lower end of the output voltage to $V_{GS2} + V_{I2}$. In the CG topology, C_C and the CG stage M_2 convert the output voltage swing to a current, returning the result to the gate of M_1.
- $V_{out} + \dfrac{g_{m2}V_2}{C_C s} = -V_2$, $g_{m1}V_1 + V_{out}\left(\dfrac{1}{R_L} + C_L s\right) = g_{m2}V_2$ and $I_{in} = \dfrac{V_1}{R_S} + g_{m2}V_2$, we obtain

$$\frac{V_{out}}{I_{in}} = \frac{-g_{m1}R_S R_L (g_{m2} + C_S s)}{R_L C_L C_C s^2 + [(1+g_{m1}R_S)g_{m2}R_L C_C + C_C + g_{m2}R_L C_L]s + g_{m2}}$$

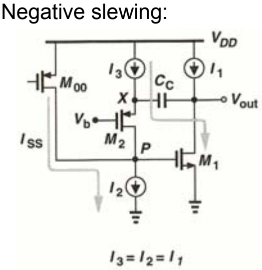
Using approximations, $\omega_{p1} \approx \dfrac{1}{g_{m1}R_L R_S C_C}$ and $\omega_{p2} \approx \dfrac{g_{m2}R_S g_{m1}}{C_L}$.

Compensation technique using a CG stage (cont'd)

Positive slewing:



Negative slewing:



- For positive slewing, M_2 and I_1 must support I_{SS}, requiring $I_1 \geq I_{SS} + I_{D1}$. If I_1 is less, then V_P drops, turning M_1 off, and if $I_1 < I_{SS}$, M_0 and its tail current source must enter the triode region, yielding a slew rate equal to I_1/C_C.
- For negative slewing, I_2 must support both I_{SS} and I_{D2}. As I_{SS} flows into node P, V_P tends to rise, increasing I_{D1}. Thus, M_1 absorbs the current produced by I_3 through C_C, tuning off M_2 and opposing the increase in V_P. We can therefore consider P a virtual ground node.
- For equal positive and negative slew rates, I_3 (and hence I_2) must be as large as I_{SS}, raising the power dissipation.

Alternative method of compensation two-stage op amps