Design of Analog Integrated Circuits

Operational Amplifiers
Outline

- Model of Operational Amplifiers
- Two Stage CMOS Op Amp
- Telescopic Op Amp
- Folded-Cascode Op Amp

Reference Books:
- Razavi: Chap. 9; Chap. 3
- Gray: Chap. 6
A Simplified Model of Real Amplifier

- An differential amplifier linearized at a DC point
The Basic Opamp Circuit

- One of the most widely used circuit structure
- It is also called Miller Op Amp (due to Miller compensation cap Cc)
Circuit Decomposition

- Two gain stages
  - Each stage provides voltage gain so as to approach the very high gain of an ideal Op Amp
- First input stage formed by the differential pair Q1-Q2, with current mirror load Q3-Q4
  - As discussed in Chapter IV, the input differential pair provides voltage gain and common-mode rejection ratio (CMRR)
  - It also performs differential-to-single-ended conversion
- Second output stage formed by common-source (CS) amplifier Q6 and its current-source load Q7
  - It provides additional voltage gain
  - Its gain is also used in Miller-multiplied cap (Cc) to reduce the compensation cap size, while achieving good stability
Biasing

- Current mirror formed by Q8, Q5 and Q7 provides the bias currents
  - Q5 for the input differential pair
  - Q7 for the output stage
- This current mirror is fed by a reference current $I_{REF}$ through Q8
  - Generally $I_{REF}$ is chosen to be constant-gm current so as to maintain consistent Opamp gain over process, voltage supply, temperature variations (PVT)
Biasing: Systematic DC Offset

- Input dc offset
  - Random offset: due to device mismatches
  - Systematic offset: due to poor biasing
- Minimum systematic offset biasing design
  - If no systematic offset, then Q6 and Q4 have the same $V_{GS}$
    \[ I_6 = \frac{(W/L)_6}{(W/L)_4} (I_{1/2}) \]
  - And this current must be exactly equal to the current supplied by Q7
    \[ I_7 = \frac{(W/L)_7}{(W/L)_5} I \]
  - So in order to minimize the systematic offset, the following condition must be met
    \[ \frac{(W/L)_6}{(W/L)_4} = 2 \frac{(W/L)_7}{(W/L)_5} \]
Input Common-Mode Range and Output Swing

- **Input Common-Mode Range**
  - Lowest $V_{ICM}$ has to ensure Q1 and Q2 in active (saturation) region
  - Highest $V_{ICM}$ has to ensure Q5 in active (saturation) region

  $$-V_{SS} + |V_{OV3}| + |V_{tn}| - |V_{tp}| \leq V_{ICM} \leq V_{dd} - |V_{OV5}| - |V_{tp}| - |V_{OV1}|$$

- **Output Swing**
  - Lowest $v_o$ has to ensure Q6 in active (saturation) region
  - Highest $v_o$ has to ensure Q7 in active (saturation) region

  $$-V_{SS} + |V_{OV6}| \leq v_o \leq V_{dd} - |V_{OV7}|$$
Equivalent Circuit

Two-Port

Small-Signal

\[ V_{id} \]
\[ G_{m1} V_{id} \]
\[ G_{m2} V_{i2} \]
\[ R_1 \]
\[ R_2 \]
\[ C_1 \]
\[ C_2 \]
\[ V_o \]
\[ V_{OUT} \]
\[ a_v \cdot V_{id} \]
\[ R_{OUT} \]
Voltage Gain

- Refer to the small signal equivalent circuit

\[ R_{in} = \infty \]

1st stage transconductance: \[ G_{m1} = g_{m1} = g_{m2} \]
1st stage output impedance: \[ R_1 = r_{o2} \parallel r_{o4} \]
1st stage dc gain: \[ A_1 = -G_{m1}R_1 = -g_{m1}(r_{o2} \parallel r_{o4}) \]

2nd stage transconductance: \[ G_{m2} = g_{m6} \]
2nd stage output impedance: \[ R_2 = r_{o6} \parallel r_{o7} \]
2nd stage dc gain: \[ A_2 = -G_{m2}R_2 = -g_{m6}(r_{o6} \parallel r_{o7}) \]

Total Gain: \[ \frac{V_o}{V_{id}} = A_1A_2 = g_{m1}(r_{o2} \parallel r_{o4}) \cdot g_{m6}(r_{o7} \parallel r_{o6}) \]
Frequency Response

- Refer to the small signal equivalent circuit
  - $C_1$ is the total capacitance at the output node of 1\textsuperscript{st} stage
    \[ C_1 = C_{gd2} + C_{db2} + C_{gd4} + C_{db4} + C_{gs6} \]
  - $C_2$ represents the total capacitance at the opamp output node, including loading capacitor $C_L$
    \[ C_2 = C_{db6} + C_{db7} + C_{gd7} + C_L \]
  - Dominant pole formed by Miller cap $C_c$ ($C_1$ is negligible)
    \[ f_{P1} \approx \frac{1}{2\pi R_1 G_{m2} R_2 C_c} \]
  - 2\textsuperscript{nd} pole formed by $C_2$
    \[ f_{P2} \approx \frac{G_{m2}}{2\pi C_2} \]
  - Zero formed by $C_c$ (between input and inverting output)
    \[ f_Z \approx \frac{G_{m2}}{2\pi C_C} \]
How to Choose Pole locations

- Recall unity-gain-bandwidth: assuming one pole system

\[ f_t = \left| A_v \right| f_{p1} = \frac{G_{m1}}{2\pi C_C} \]

- We can choose the pole locations so that the opamp can be approximated as a one pole system within its unity-gain-bandwidth
  - \( f_{p2} \) and \( f_z \) have to be higher than \( f_t \)

- For \( f_{p2} > f_t \)

\[ \frac{G_{m1}}{C_C} < \frac{G_{m2}}{C_2} \]

- For \( f_z > f_t \)

\[ G_{m1} < G_{m2} \]

- Or is this good enough? We will revisit “Zero Problem” later
Phase Margin

- Miller cap \( C_c \) is utilized for pole-splitting, i.e., adding \( C_c \) lowers the dominant pole, while pushing away the 2\(^{nd}\) pole beyond \( f_t \).
- By careful design, it is possible to have about 90° phase shift at \( f_t \), leaving us a comfortable phase margin about 90°.
- However, if the 2\(^{nd}\) pole is close enough to \( f_t \), the opamp will experience much more phase shift, eating up the phase margin.
- And the zero has the same negative effect on the opamp’s phase margin:
  - the zero in on the right half plane-RHP (i.e., it is >0)
  
  \[
  f_Z \approx \frac{G_m}{2\pi C_C}
  \]
  
  - It increases phase shift and thus decreases the phase margin.
  - To make it worse, the RHF zero also increases gain and thus reduces the gain margin.
  - Generally speaking, an opamp with a RHP zero is a bad design and potentially unstable.
Bode Plot of Two Stage CMOS Opamp

RHP zero reduces gain margin
Zero: Solving Right Half Plane (RHP) Problem

- RHP zero reduces gain margin and makes opamp unstable
- Solution: adding a resistance $R$ in series with $C_c$
- Now the new zero is calculated as:

$$f_z \approx \frac{1}{2\pi C_c \left( \frac{1}{G_{m2}} - R \right)}$$
Zero: LHP or Better?

- With RC lead compensation, the new zero is
  \[ f_z \approx \frac{1}{2\pi C C \left( \frac{1}{G_{m2}} - R \right)} \]

- We can choose an R value (e.g. R approaching 1/Gm2) so that the zero is moving farther away from ft, making it negligible to phase shift at ft;
- Or we can choose R value exactly the same as 1/Gm2. The zero is at infinity;
- Or we can further increase R value, e.g., R>1/Gm2. Now the zero is at left half plane (LHP); it still increases the phase shift but it also reduces the gain (so improve the gain margin)
- Or better (?) we can choose such an R value that it moves the zero into LHP and cancel the 2\textsuperscript{nd} pole.
Zero-Pole Cancellation: Is It Better?

- Recall the zero and 2\textsuperscript{nd} pole

  \[ f_z \approx \frac{1}{2\pi C_C (\frac{1}{G_m^2} - R)} \]

  \[ f_{p2} \approx \frac{G_m^2}{2\pi C_2} \]

- Setting them equal, we get an R value:

  \[ R = \left( \frac{C_2}{C_C} + 1 \right) \frac{1}{G_m^2} \]

- Now, we move the zero into LHP, right at the 2\textsuperscript{nd} pole; Sounds like an interesting idea to cancel the 2\textsuperscript{nd} pole with the zero;

- Unfortunately, the exact value of C2 is usually unknown since it depends on the loading

- Food for thought: how we can play with R value to optimize the frequency compensation to achieve best phase margin? And independent of process variation (tracking Gm2)?
How About the Ideal Current Source

A bias current is independent of both the supply voltage and the MOSFET threshold voltage;

It is determined by a single resistor and the device dimensions.

\[ I_B = \frac{1}{2} \mu_n C_{ox} \left( \frac{W}{L} \right)_{12} \left( V_{GS12} - V_t \right)^2 \]

\[ I_B = \frac{1}{2} \mu_n C_{ox} \left( \frac{W}{L} \right)_{13} \left( V_{GS13} - V_t \right)^2 \]

\[ V_{GS13} = V_{GS12} + I_B R_R \]

\[ I_B = \frac{2}{\mu_n C_{ox} (W/L)_{12} R_B^2} \left( \frac{\sqrt{(W/L)_{12}}}{\sqrt{(W/L)_{13}}} - 1 \right)^2 \]
Constant $g_m$ Bias Circuits

**Constant $g_m$**

\[ g_{m12} = \sqrt{2\mu_n C_{ox} (W/L)_{12}} I_B = \frac{2}{R_B} \left( \frac{(W/L)_{12}}{(W/L)_{13}} - 1 \right) \]

- **n-channel device**
  \[ g_{mi} = g_{m12} \frac{I_{Di} (W/L)_i}{I_B (W/L)_{12}} \]
  \[ g_{mi} = g_{m12} \frac{\mu_p I_{Di} (W/L)_i}{\mu_n I_B (W/L)_{12}} \]

- **p-channel device**
Slew Rate (SR)-Concept

- A unity-gain follower with 1V step input
- One side of input differential pair will be turn off
Slew Rate (SR)-Two Stage Opamp Model

- The input stage can be modeled as a current source;
- one side of the differential pair is off
- The other side conducts all current, acting as a current source
- The 2\textsuperscript{nd} stage can be modeled as an integrator

\[ SR = \frac{I}{C_C} \]
Slew Rate (SR) and Unity-Gain-Bandwidth (ft)

- Recall SR and unity-gain-bandwidth ft
  \[ SR = \frac{I}{C_C} \quad f_t = |A_v| f_{P1} = \frac{G_{m1}}{2\pi C_C} \]

- Combining these two equations, we obtain:
  \[ SR = 2\pi f_t \frac{I}{G_{m1}} = 2\pi f_t V_{OV} \]

- Thus, for a given ft, the SR is determined by the overdrive voltage of the input differential pair
Outline

- Ideal Operational Amplifiers
- Two Stage CMOS Op Amp
- Telescopic Op Amp
- Folded-Cascode Op Amp
Single-Stage Cascode Op Amp

- Many opamp applications see pure capacitive loads
  - Not necessary to have low output impedance
- If we design an opamp which has only one high impedance node at the output:
  - No frequency compensation is needed; the dominant pole is at the output and the load capacitance simply helps stability;
  - High speed since all internal nodes having relatively low impedance and low capacitance
- These opamps are also called Operational Transconductance Amplifiers (OTAs) since their transconductance (output current vs. input voltage) is one of the key performance parameters
  - The Cascode Op Amp can be categorized as OTA and is one of the most popular modern CMOS opamps
Telescopic Op Amp

- The telescopic opamp is also known as the cascode opamp, where it uses cascode current mirror at its output load to achieve high output impedance
- Motivation: single-stage opamp to achieve the same gain as in two-stage opamp and to simplify the frequency compensation
  - The frequency compensation is done by adding load capacitance
- How to do it:
  - Two-stage opamp gain is about $1/2(g_m r_{ds})^2$
  - Recall cascode current mirror output impedance is about $g_m (r_{ds})^2$
The Circuit

It only has one high impedance node at the output, where the dominant pole is; Frequency compensation is done by adding load capacitance;

Question: where is the first non-dominant pole?
The current are balanced so,
\[ I_{DS} = I_{REF} \]
except for M13 which has,
\[ I_{DS13} = 2I_{REF} \]
Since,
\[ I_1 = I_2 = I_{REF} \]
Then,
\[ V_{GS3} = V_{GS4} = V_{GS9} \]
This also implies that,
\[ V_{DS4} = V_{DS3} \]
Similarly,
\[ V_{DS5} = V_{DS6} \]
Output Swing

Assuming differential input signal is zero:
\[ V_{id} = 0 \]
Then,
\[ V_{OUT} = V_{DD} - 2V_{TO} - 2V_{DSAT} = V_{G6} \]
The gate of M8 is also at,
\[ V_{DD} - 2V_{TO} - 2V_{DSAT} \]
Due to its connection to M11, so
\[ V_{OUT} = V_{G6} = V_{G8} \]
The swing in the positive direction will be,
\[ V_{OUT,MAX} = V_{G6} + V_T \]
But since
\[ V_{OUT} = V_{G6} \]
Thus output swing is limited to only 1 \( V_T \) in the positive direction!
Output Swing—cont.

In the negative direction,
\[ V_{OUT,MIN} = V_{G8} - V_T \]
but since
\[ V_{OUT} = V_{G8} \]
the swing is only 1 \( V_T \) again.

Thus the total output swing is only 2\( V_T \), not good for most applications.

So is there a solution for the limited output swing?
- we could use a high swing cascode configuration as was described before.
- Or “Folded-cascode” opamp
Outline

- Ideal Operational Amplifiers
- Two Stage CMOS Op Amp
- Telescopic Op Amp
- Folded-Cascode Op Amp
The Circuit
Circuit Decomposition

- CS amplifier (differential pair) with CG amplifier (cascode transistors)
- Input stage formed by the differential pair Q1-Q2
- Q3-Q4 form the CG amplifier or cascode transistors
- Cascode current mirror Q5-Q8 forms the output current source load so as to achieve the high output impedance.
- Q9/Q10//Q11 are biasing transistors
- This is also known as “folded-down” cascode op amp. Why?
Biasing

- Q11 provides the constant current I utilized for biasing the differential pair;
- Input differential pair Q1-Q2 is operating at I/2 for each transistor
- Q9/Q10 provide the constant current I_B.
- Each of cascode transistors Q3-Q4 is biased at (I_B-I/2)
- I_B has to be greater than I in order to keep Q3/Q4 in active region and for large signal operation (during SR limiting, all I will go through Q1 or Q2)
  - Usually consumes more power than telescopic cascode op amp.
- \( V_{BIAS1} / V_{BIAS2} / V_{BIAS3} \) has to be carefully selected in order to achieve optimum input common-mode range and output voltage swing.
Input Common-Mode Range and Output Swing

- **Input Common-Mode Range**
  - Highest $V_{ICM}$ has to ensure Q1 and Q2 in active (saturation) region
  - Lowest $V_{ICM}$ has to ensure Q11 in active (saturation) region

  \[-V_{SS} + |V_{OV11}| + |V_{tn}| - |V_{OV1}| \leq V_{ICM} \leq V_{dd} - |V_{OV9}| + |V_{tn}|\]

  - Note: $V_{ICM}$ can be larger than $V_{DD}$, a significant improvement over two-stage opamp.

- **Output Swing**
  - Lowest $v_o$ has to ensure Q6 in active (saturation) region
  - Highest $v_o$ has to ensure Q10 and Q4 in active (saturation) region
    - By carefully selecting $V_{BIAS1}$ so that Q10 operates at the edge of saturation

  \[-V_{SS} + |V_{OV7}| + |V_{OV5}| + V_{tn} \leq v_o \leq V_{dd} - |V_{OV4}| - |V_{OV10}|\]

  - Note: Folded-cascode has larger output swing than telescopic.
Improving Output Swing

![Circuit Diagrams](image)

\[ -V_{ss} + |V_{OV7}| + |V_{OV5}| \leq V_o \leq V_{dd} - |V_{OV4}| - |V_{OV10}| \]
Equivalent Circuit

Small-signal equivalent circuit of the folded-cascode CMOS opamp. This circuit is in effect an OTA.
Voltage Gain

- Refer to the small signal equivalent circuit
- Transconductance \( G_m = g_m1 = g_m2 \)
- Output resistance

\[
R_o = R_{o4} \parallel R_{o6}
\]

\[
R_{o4} \approx (g_{m4}r_{o4})(r_{o2}||r_{o10})
\]

\[
R_{o6} \approx g_{m6}r_{o6}r_{o8}
\]

- The dc open-loop gain

\[
A_v = G_m R_o = g_m1 \{[g_{m4}r_{o4}(r_{o2}||r_{o10})](g_{m6}r_{o6}r_{o8})\}
\]

- Its gain is about \( \frac{1}{2} \) of that of a telescopic cascode opamp
Frequency Response

- Refer to the small signal equivalent circuit
  - $C_L$ is the total capacitance at the output node
- Dominant pole formed by $C_L$ and output resistance
  \[ f_{P1} \approx \frac{1}{2\pi R_o C_L} \]
- Unity gain frequency
  \[ f_t = |A_v| f_{P1} = \frac{G_m}{2\pi C_L} \]
- Adding $C_L$
  - Decreases opamp’s bandwidth (unity gain frequency)
  - Improves opamp’s stability
Slew Rate (SR)

- When a large input signal is applied, the input differential pair can be modeled as a current source:
  - one side of the differential pair is off
  - The other side conducts all bias current $I_b$, acting as a current source
  - One of cascode transistors will carry a current of $(I_B-I_b)$, while the other carrying a current of $I_B$
  - The output current to $C_L$ will then be $I_B-(I_B-I)=I$

- The slew rate:

\[ SR = \frac{I}{C_L} \]

- For a given $f_t$, the SR is determined by the overdrive voltage of the input differential pair

\[ SR = 2\pi f_t \frac{I}{G_m} = 2\pi f_t V_{OV} \]
Input Common-Mode Range—Rail-to-Rail
Rail-to-Rail Operation

- Two parallel complementary input stages
  - Q5-Q6 are the cascode transistors for the Q1-Q2 NMOS input pair
  - Q7-Q8 are the cascode transistors for the Q3-Q4 PMOS input pair
  - $V_{ICM}$ can be higher than $V_{DD}$ and lower than $-V_{SS}$, achieving rail-to-rail operation.

- However:
  - When $V_{ICM}$ is close to $V_{DD}$, only Q1-Q2 pair is operating and Q3-Q4 pair is off; The total gain is:
    $$A_v = G_{mn} R_o$$
  - When $V_{ICM}$ is close to $-V_{SS}$, only Q3-Q4 pair is operating and Q1-Q2 pair is off; The total gain is:
    $$A_v = G_{mp} R_o$$
  - When $V_{ICM}$ is in the mid-range, both Q1-Q2 and Q3-Q4 pairs are operating; The total gain is:
    $$A_v = (G_{mn} + G_{mp}) R_o$$

- Question: rail-to-rail operation with consistent gain?
Summary

- The IC op amp is a versatile circuit building block
  - Ideally only amplify the input difference signal
  - Ideally infinite input resistance and zero output resistance
  - Virtual Ground: potentials at two input terminals track each other when applying negative feedback
- On-chip CMOS op amps
  - Usually drive capacitive load so do not need zero output resistance
  - High gain: two stages or cascode amplifiers
- Two-stage op amps
  - Miller compensation
  - Lead compensation to move RHP zero into infinity or LHP
  - SR is determined by the first-stage bias current and Miller cap
- Cascode op amps
  - Frequency compensation is done by adding or increasing load capacitance
  - Telescopic op amps have limited input common-mode range and output swing
  - Folded-cascode op amps have better ICM and output swing at the cost of higher current consumption and lower gain (about half)
  - SR is determined by the input pair bias current and loading cap