Lecture 16 - Linear Equivalent Circuits - Outline

- Announcements
  Handout - Lecture Foils

- Review - CMOS inverters
  Noise Margins: when \(|\frac{dv_{\text{OUT}}}{dv_{\text{IN}}}| > 1\), disturbance grows
  Gate delay: often defined as average, i.e. \((\tau_{\text{LO-HI}} + \tau_{\text{HI-LO}})/2\) (We leave the 2 out)
  Power density: \(PD = P_{\text{ave}} @ f_{\text{max}} / \text{Area}\) (Power density is what limits IC performance)
  Using \(P_{\text{ave}} = f C_L V_{\text{DD}}^2\), and noting \(f_{\text{max}} \propto 1/GD\), and Area \(\propto W_n L_{\text{min}}\):
  \(PD \propto \mu e \varepsilon_{\text{ox}} V_{\text{DD}}(V_{\text{DD}} - V_T)^2/t_{\text{ox}} L_{\text{min}}^2\) (Must reduce dimensions and voltage - Lec 25)

- Small signal models; linear equivalent circuits
  pn diodes: linearizing the exponential diode
  incorporating the charge stores

  BJTs: linearizing the Ebers-Moll model
  incorporating the charge stores
  adding the Early effect and possible parasitics

  MOSFETs: linearizing the Gradual-Channel model
  incorporating the charge stores
  adding the Early effect and possible parasitics
CMOS Gate Delay and Power Dissipation

our focus was on the switching transients, HI to LO and LO to HI

Charge: \( v_{IN} \): HI to LO
\( Q_n \) off, \( Q_p \) on

\( v_{OUT} \): LO to HI

\[ \tau_{Charge} = \frac{6nL_{min}^2 V_{DD}}{\mu_e [V_{DD} - V_{Tn}]^2} \]

\[ E_{Stored \ in \ C_L} = \frac{1}{2} C_L V_{DD}^2 \]

\[ E_{Dissipated \ in \ Q_p} = \frac{1}{2} C_L V_{DD}^2 \]

Discharge: \( v_{IN} \): LO to HI
\( Q_n \) on, \( Q_p \) off

\( v_{OUT} \): HI to LO

\[ \tau_{Discharge} = \frac{6nL_{min}^2 V_{DD}}{\mu_e [V_{DD} - V_{Tn}]^2} \]

\[ E_{Stored \ in \ C_L} = 0 \]

\[ E_{Dissipated \ in \ Q_n} = \frac{1}{2} C_L V_{DD}^2 \]
CMOS Gate Delay and Power Dissipation, cont.

**challen.ge:** high speed without melting silicon

**Maximum data rate**

Proportional to $1/\tau_{\text{Min. Cycle}}$

$$\tau_{\text{Min.Cycle}} = \tau_{\text{Charge}} + \tau_{\text{Discharge}} = \frac{12nL_{\text{min}}^2 V_{\text{DD}}}{\mu e [V_{\text{DD}} - V_{\text{Tn}}]^2}$$

Implies we should **decrease dimensions** and **increase voltages**.

**Average power density at maximum data rate**

Assumes that the area per inverter is proportional to $W_{\text{min}} L_{\text{min}}$

$$PD_{\text{ave @ Max.f}} \propto \frac{P_{\text{ave @ Max.f}}}{W_{\text{min}} L_{\text{min}}} = \frac{\mu e \varepsilon_{\text{ox}} V_{\text{DD}} [V_{\text{DD}} - V_{\text{Tn}}]^2}{t_{\text{ox}} L_{\text{min}}^2}$$

**The Problem:** Pick a number $s > 1$ (i.e., 1.5, 2, 3, etc.)

Suppose we decrease dimensions by $1/s$, and increase voltages by $s$,

The maximum data rate increases as $s^3$, but

the power density also increases and as $s^6$!!

How do we make $f_{\text{max}}$ larger without melting the silicon?

By following CMOS scaling rules - the topic of Lecture 25
CMOS: transfer characteristic calculation revisited

Our calculation said that the transfer characteristic is vertical in Region III. We know it must have some slope, but what is it?

To see, we calculated the small signal gain about the bias point, $V_{IN} = V_{OUT} = V_{DD}/2$, starting with the small signal model:

\[
V_{DD} - V_{TP} = \frac{V_{DD}}{2} + v_{IN}
\]

\[
V_{DD} - V_{TP} = \frac{V_{DD}}{2} + v_{OUT}
\]

\[
K_p \lambda_p \frac{V_{DD}}{2} + v_{IN} = v_{gsn} = v_{IN}
\]

\[
g_p + g_m v_{gsp} = g_m v_{gsn} = v_{gsn}
\]

\[
K_n \lambda_n \frac{V_{DD}}{2} + v_{OUT} = v_{gsn} = v_{IN}
\]

\[
g_n + g_m v_{gsp} = g_m v_{gsn} = v_{gsn}
\]
**CMOS:** transfer characteristic calculation, cont.

Redrawing the circuit we get

\[
\begin{align*}
\text{from which we see immediately that the gain and slope are} \\
A_v & \equiv \frac{\partial v_{OUT}}{\partial v_{IN}} \bigg|_Q = \frac{v_{out}}{v_{in}} = -\frac{\left[ g_{mn} + g_{mp} \right]}{g_{on} + g_{op}} \\
\end{align*}
\]

Today we will return to the origins of our incremental models and see how to relate the g's to the bias point and device characteristics.
Adding charge stores to the large signal models:

**p-n diode:** $I_{BS} \quad q_{AB}$  
$q_{AB}$: Excess carriers on p-side plus excess carriers on n-side plus junction depletion charge.

**BJT:** npn (in F.A.R.)

$q_{BC}$: C-B junction depletion charge
$q_{BE}$: Excess carriers in base plus E-B junction depletion charge
$q_{DB}$: D-B junction depletion charge

**MOSFET:** n-channel

$q_{G}$: Gate charge; a function of $v_{GS}$, $v_{DS}$, and $v_{BS}$.
$q_{DB}$: D-B junction depletion charge
$q_{SB}$: S-B junction depletion charge
• Linear equivalent circuit (LEC) for the p-n junction diode:

We begin with the static model for the terminal characteristics:

$$i_D(v_{AB}) = I_S \left[ e^{qv_{AB}/kT} - 1 \right]$$

Linearizing $i_D$ about $V_{AB}$, which we will denote by $Q$ (for quiescent bias point):

$$i_D(v_{AB}) \approx i_D(V_{AB}) + \frac{\partial i_D}{\partial v_{AB}} \bigg|_Q \left[ v_{AB} - V_{AB} \right]$$

We define the equivalent incremental conductance of the diode, $g_d$, as:

$$g_d \equiv \frac{\partial i_D}{\partial v_{AB}} \bigg|_Q = \frac{q}{kT} I_S e^{qv_{AB}/kT} \approx \frac{qI_D}{kT}$$

and we use our notation to write:

$$I_D = i_D(V_{AB}), \quad i_d = [i_D - I_D], \quad v_{ab} = [v_{AB} - V_{AB}]$$

ending up with

$$i_d = g_d v_{ab}$$

The corresponding LEC is shown at right:

$$g_d \approx \frac{qI_D}{kT}$$
LEC for the p-n junction diode, cont.:

At high frequencies we must include the charge store, \( q_{AB} \), and linearize its two components:

\[
q_{AB} = q_{DP} + q_{QNR,p-side} \quad C_d = C_{dp} + C_{df}
\]

**Depletion layer charge store,** \( q_{DP} \), and its linear equivalent capacitance, \( C_{dp} \):

\[
q_{DP}(v_{AB}) = -A q N_{Ap} x_p (v_{AB}) \approx -A \sqrt{2 q \varepsilon_{Si} N_{Ap} (\phi_b - v_{AB})}
\]

\[
C_{dp}(V_{AB}) \equiv \left. \frac{\partial q_{DP}}{\partial v_{AB}} \right|_Q = A \sqrt{\frac{q \varepsilon_{Si} N_{Ap}}{2 (\phi_b - V_{AB})}}
\]

**Diffusion charge store,** \( q_{QNR,p-side} \), and its linear equivalent capacitance, \( C_{df} \):

\[
q_{QNR,p-side}(v_{AB}) = i_D \left[ w_p - x_p \right]^2 / 2 D_e
\]

\[
C_{df}(V_{AB}) \equiv \left. \frac{\partial q_{QNR,p-side}}{\partial v_{AB}} \right|_Q = \frac{q I_D}{kT} \frac{[w_p - x_p]^2}{2 D_e} = g_d \tau_d \quad \text{with} \quad \tau_d \equiv \frac{[w_p - x_p]^2}{2 D_e}
\]

(Note: All of this is for an n\(^{+}\)-p diode)
• Creating a linear equivalent circuit, LEC:

Suppose we have a device with three terminals, X, Y, and Z, and that we have expressions for the currents into terminals X and Y in terms of the voltages \( v_{XZ} \) and \( v_{YZ} \):

\[
i_X(v_{XZ}, v_{YZ}) \quad \text{and} \quad i_Y(v_{XZ}, v_{YZ})
\]

Suppose we also have expressions for the charge stores associated with terminals X and Y:

\[
q_X(v_{XZ}, v_{YZ}) \quad \text{and} \quad q_Y(v_{XZ}, v_{YZ})
\]

We begin with the static model for the terminal characteristics, and linearize them about an bias point, Q, defined as a specific set of \( v_{XZ} \) and \( v_{YZ} \) that we write, using our notation, as \( V_{XZ} \) and \( V_{YZ} \)

For example, for the current into terminal X we have:

\[
i_X(v_{XZ}, v_{YZ}) = i_X(V_{XZ}, V_{YZ}) + \frac{\partial i_X}{\partial v_{XZ}} \bigg|_Q (v_{XZ} - V_{XZ}) + \frac{\partial i_X}{\partial v_{YZ}} \bigg|_Q (v_{YZ} - V_{YZ}) + \text{higher order terms}
\]

For sufficiently small \( (v_{XZ} - V_{XZ}) \) and \( (v_{YZ} - V_{YZ}) \), we have:

\[
i_X(v_{XZ}, v_{YZ}) \approx i_X(V_{XZ}, V_{YZ}) + \frac{\partial i_X}{\partial v_{XZ}} \bigg|_Q (v_{XZ} - V_{XZ}) + \frac{\partial i_X}{\partial v_{YZ}} \bigg|_Q (v_{YZ} - V_{YZ})
\]

continued on the next page
• Creating a linear equivalent circuit, LEC, cont.:

Using our notation, we recognize that:

\[ I_X \equiv i_X (V_{XZ}, V_{YZ}), \quad i_x \equiv [i_X - I_X], \quad v_{xz} \equiv [v_{XZ} - V_{XZ}], \quad v_{yz} \equiv [v_{YZ} - V_{YZ}] \]

We identify the partial derivatives as conductances, and name them as:

\[
\left. \frac{\partial i_X}{\partial v_{xz}} \right|_Q \equiv g_i \quad \left. \frac{\partial i_X}{\partial v_{yz}} \right|_Q \equiv g_r
\]

Applying these to our earlier result we have:

\[ i_x (v_{XZ}, v_{YZ}) \approx I_X + g_i v_{xz} + g_r v_{yz} \quad \text{and thus} \quad i_x \approx g_i v_{xz} + g_r v_{yz} \]

Doing the same for \( i_y \), we arrive at

\[ i_y \approx g_f v_{xz} + g_o v_{yz} \quad \text{where} \quad g_f \equiv \left. \frac{\partial i_y}{\partial v_{xz}} \right|_Q \quad g_o \equiv \left. \frac{\partial i_y}{\partial v_{yz}} \right|_Q \]

A circuit matching these relationships is shown below:
• Creating a linear equivalent circuit, LEC, cont.:

Thus far our linear equivalent circuit is only good at low frequencies:

Next, to handle high frequency signals, we linearize the charge stores' dependencies on voltage. Their LECs, which are linear capacitors:

\[
\left. \frac{\partial q_X}{\partial v_{xz}} \right|_Q \equiv C_{xz} \quad \left. \frac{\partial q_Y}{\partial v_{yz}} \right|_Q \equiv C_{yz} \quad \left. \frac{\partial q_X}{\partial v_{xy}} \right|_Q \equiv C_{xy} \left( \frac{\partial q_Y}{\partial v_{yx}} \right|_Q \right)
\]

Adding these to the model gives us:
• Linear equivalent circuit for the BJT (static):

In the forward active region, our static model says:

\[
\begin{align*}
  i_B(v_{BE}, v_{CE}) &= I_{BS} \left[ e^{qv_{BE}/kT} - 1 \right] \\
  i_C(v_{BE}, v_{CE}) &= \beta_o [1 + \lambda v_{CE}] i_B(v_{BE}, v_{CE}) = \beta_o I_{BS} \left[ e^{qv_{BE}/kT} - 1 \right] [1 + \lambda v_{CE}]
\end{align*}
\]

We begin by linearizing \( i_C \) about \( Q \):

\[
\begin{align*}
  i_c(v_{be}, v_{ce}) &= \frac{\partial i_C}{\partial v_{BE}} \bigg|_Q v_{be} + \frac{\partial i_C}{\partial v_{CE}} \bigg|_Q v_{ce} = g_m v_{be} + g_o v_{ce}
\end{align*}
\]

We introduced the transconductance, \( g_m \), and the output conductance, \( g_o \), defined as:

\[
\begin{align*}
  g_m &= \frac{\partial i_C}{\partial v_{BE}} \bigg|_Q \\
  g_o &= \frac{\partial i_C}{\partial v_{CE}} \bigg|_Q
\end{align*}
\]

Evaluating these partial derivatives using our expression for \( i_C \), we find:

\[
\begin{align*}
  g_m &= \frac{q}{kT} \beta_o I_{BS} e^{qV_{BE}/kT} \left[ 1 + \lambda V_{CE} \right] \approx \frac{qI_C}{kT} \\
  g_o &= \beta_o I_{BS} \left[ e^{qV_{BE}/kT} + 1 \right] \lambda \approx \lambda I_C \quad \left( \text{or} \quad \approx \frac{I_C}{V_A} \right)
\end{align*}
\]

(Continued on next foil.)
LEC for the **BJT** (static), cont.:

Turning next to $i_B$, we note it only depends on $v_{BE}$ so we have:

$$i_b(v_{be}) = \frac{\partial i_B}{\partial v_{BE}} \bigg|_Q v_{be} = g_\pi v_{be}$$

The input conductance, $g_\pi$, is defined as:

$$g_\pi = \frac{\partial i_B}{\partial v_{BE}} \bigg|_Q$$

To evaluate $g_\pi$ we do not use our expression for $i_B$, but instead use $i_B = i_C/\beta_o$:

$$g_\pi = \frac{\partial i_B}{\partial v_{BE}} \bigg|_Q = \frac{1}{\beta_o} \frac{\partial i_C}{\partial v_{BE}} \bigg|_Q = \frac{g_m}{\beta_o} = \frac{qI_C}{\beta_o kT}$$

(Notice that we do not define $g_\pi$ as $qI_B/kT$)

Representing this as a circuit we have: (Notice that $v_{be}$ has been renamed $v_\pi$)
• **Linear equivalent circuit for the BJT (dynamic):**

To complete the model we next linearize and add the charge stores associated with the two junctions.

The base-collector junction is reverse biased so the charge associated with it, \( q_{BC} \), is the depletion region charge. The corresponding capacitance is labeled \( C_\mu \).

The base-emitter junction is forward biased as has the excess charge injected into the base as well as the base-emitter depletion charge store associated with it. The linear equivalent capacitance is labeled \( C_\pi \). The part of \( C_\pi \) due to the excess charge turns out to be \( q|I_C|w_B^2/2D_e kT \), which can also be written \( g_m \tau_b \) with \( \tau_b = w_B^2/2D_e \).

**Summarizing:** \( C_\pi = g_m \tau_b + \text{B-E depletion cap.} \), \( C_\mu : \text{B-C depletion cap.} \)

Adding these C's to our model:
• Linear equivalent circuit for the **MOSFET** (static):

In saturation, our static model is: (We've said $\alpha \approx 1$)

$$
\begin{align*}
  i_G(v_{GS}, v_{DS}, v_{BS}) &= 0 \\
  i_B(v_{GS}, v_{DS}, v_{BS}) &\approx 0 \\
  i_D(v_{GS}, v_{DS}, v_{BS}) &= \frac{K}{2} \left[ v_{GS} - V_T(v_{BS}) \right]^2 [1 + \lambda v_{DS}] \\
\end{align*}
$$

with $K \equiv \frac{W}{L} \mu_e C^*_{ox}$ and $V_T = V_{To} +$ 

Note that because $i_G$ and $i_B$ are zero they are already linear, and we can focus on $i_D$. Linearizing $i_D$ about $Q$ we have:

$$
\begin{align*}
  i_d(v_{gs}, v_{ds}, v_{bs}) &= \frac{\partial i_D}{\partial v_{GS}} \bigg|_Q v_{gs} + \frac{\partial i_D}{\partial v_{DS}} \bigg|_Q v_{ds} + \frac{\partial i_D}{\partial v_{BS}} \bigg|_Q v_{bs} \\
  &= g_m v_{gs} + g_o v_{ds} + g_{mb} v_{bs}
\end{align*}
$$

We have introduced the transconductance, $g_m$, output conductance, $g_o$, and substrate transconductance, $g_{mb}$:

$$
\begin{align*}
  g_m &\equiv \frac{\partial i_D}{\partial v_{GS}} \bigg|_Q \\
  g_o &\equiv \frac{\partial i_D}{\partial v_{DS}} \bigg|_Q \\
  g_{mb} &\equiv \frac{\partial i_D}{\partial v_{BS}} \bigg|_Q
\end{align*}
$$

(Continued on next foil.)
LEC for the MOSFET (static), cont.:

Evaluating the conductances using our expression for $i_D$, we find:

\[ g_m \equiv \frac{\partial i_D}{\partial v_{GS}} \bigg|_Q = K [V_{GS} - V_T(V_{BS})][1 + \lambda V_{DS}] \approx \sqrt{2KI_D} \]

\[ g_o \equiv \frac{\partial i_D}{\partial v_{DS}} \bigg|_Q = \frac{K}{2} [V_{GS} - V_T(V_{BS})]^2 \lambda \approx \lambda I_D \]

\[ g_{mb} \equiv \frac{\partial i_D}{\partial v_{BS}} \bigg|_Q = -K [V_{GS} - V_T(V_{BS})][1 + \lambda V_{DS}] \frac{\partial V_T}{\partial v_{BS}} \bigg|_Q = \eta g_m = \eta \sqrt{2KI_D} \]

Representing this as a circuit we have:

\[ \eta \equiv -\frac{\partial V_T}{\partial v_{BS}} \bigg|_Q = \frac{1}{C_{ox}^*} \sqrt{\frac{\varepsilon_{Si}qN_A}{|q\phi_p| - V_{BS}}} \]
• Linear equivalent circuit for the **MOSFET** (dynamic):

To complete the model we next linearize and add the charge stores associated with each pair of terminals.

In saturation \( q_G \) is a function only of \( v_{GS} \) and \( v_{GB} \), so our model only accounts for \( C_{gs} \) and \( C_{gb} \). \( C_{gd} \) is a parasitic element.

We have:

\[
C_{gs} = \frac{2}{3} WL C_{ox}^* \\
C_{gd} \text{: sum of G-D fringing and overlap capacitances (all parasitics)} \\
C_{sb}, C_{gb}, C_{db} : \text{depletion capacitances}
\]

Adding these C's to our model:
• Linear equivalent circuit for transistors (dynamic):
  Collecting our results for the MOSFET and BJT biased in FAR

BJT:

\[ g_m = \frac{q}{kT} \beta_o I_{BS} e^{qV_{BE}/kT} \left[ 1 + \lambda V_{CE} \right] \approx \frac{qI_C}{kT} \]

\[ g_\pi = \frac{g_m}{\beta_o} = \frac{qI_C}{\beta_o kT} \]

\[ g_o = \beta_o I_{BS} \left[ e^{qV_{BE}/kT} + 1 \right] \lambda \approx \lambda I_C = \frac{I_C}{V_A} \]

\[ C_\pi = g_m \tau_b + \text{B-E depletion cap. with } \tau_b = \frac{w_B^2}{2D_e}, \quad C_\mu : \text{B-C depletion cap.} \]

MOSFET:

\[ g_m = K \left[ V_{GS} - V_T(V_{BS}) \right] \left[ 1 + \lambda V_{DS} \right] \approx \sqrt{2KI_D} \]

\[ g_o = \frac{K}{2} \left[ V_{GS} - V_T(V_{BS}) \right]^2 \lambda \approx \lambda I_D = \frac{I_D}{V_A} \]

\[ g_{mb} = \eta g_m = \eta \sqrt{2KI_D} \]

with \[ \eta = -\frac{\partial V_T}{\partial V_{BS}} \bigg|_Q = \frac{1}{C_{ox}^*} \sqrt{\frac{\varepsilon_{Si}qN_A}{q\phi_p - V_{BS}}} \]

\[ C_{gs} = \frac{2}{3} W L C_{ox}^*, \quad C_{sb}, C_{gb}, C_{db} : \text{depletion capacitances} \]

\[ C_{gd} = W C_{gd}^*, \text{ where } C_{gd}^* \text{ is the G-D fringing and overlap capacitance per unit gate length (parasitic)} \]
**CMOS:** transfer characteristic calculation re-revisited

Returning to slope of the transfer in Region III, we had it expressed in terms of the g's, but now we can express it in terms of the bias point and the transistor parameters. We had:

$$A_v = \left. \frac{\partial v_{OUT}}{\partial v_{IN}} \right|_Q = -\frac{2\sqrt{2K_n}}{\left[\lambda_n + \lambda_p\right]I_{Dn}}$$

Using what we just learned about the MOSFET l.e.c. parameters we can write:

$$g_{mn} = \sqrt{2K_nI_{Dn}}, \quad g_{mp} = \sqrt{2K_pI_{Dp}} = g_{mn}, \quad g_{on} = \lambda_n I_{Dn}, \quad g_{op} = \lambda_p I_{Dp} = \lambda_p I_{Dn}$$

Incorporating these relationships gives:

$$A_v = \left. \frac{\partial v_{OUT}}{\partial v_{IN}} \right|_Q = -\frac{2\sqrt{2K_nI_{Dn}}}{\left[\lambda_n + \lambda_p\right]I_{Dn}} = -\frac{2\sqrt{2K_n}}{\left[\lambda_n + \lambda_p\right]I_{Dn}}$$
Lecture 16 - Linear Equivalent Circuits - Summary

- Analog circuit design; small signal models
  Linear amplification and processing of signals

- Linear equivalent circuits: it all depends on the bias point

**pn diodes:**
\[ g_d = q|I_D|/kT \]
\[ C_d = g_d\tau_d + C_{dp}(V_{AB}) \]

**BJTs:** (in FAR)
\[ g_m = q|I_C|/kT \]
\[ g_{\pi} = g_m/\beta_F \]
\[ g_o = |I_C/V_A| \quad [\text{or } \lambda |I_C|] \]
\[ C_{\pi} = g_m\tau_b + C_{dp,be}(V_{BE}) \]
\[ C_{\mu} = C_{dp,be}(V_{BC}) \]

**MOSFETs:** (in saturation)
\[ g_m = K(V_{GS} - V_T) = (2K|I_D|)^{1/2} \]
\[ g_{mb} = \eta g_m \]
\[ [\eta = \{\varepsilon_{Si}qN_A/2(2\phi_p - V_{BS})\}^{1/2}/C_{ox}^*] \]
\[ g_o = |I_D/V_A| \quad [\text{or } \lambda |I_D|] \]
\[ C_{gs} = (2/3) WL C_{ox}^* \]
\[ C_{gd}: \text{G-D fringing and overlap capacitance, all parasitic} \]
\[ C_{sb}, C_{gb}, C_{db}: \text{depletion capacitances} \]