Note 1: The Small Signal Model for Transistors

Rohit J. Braganza

Abstract In this notesheet we will discover how to deal with the complexity of transistors through the creation of a linear model. In the process we will also discover how transistors provide gain.

1 Creating a Linear Circuit Model for a Black Box

Transistors are most definitely *not* linear devices. However, as an analog designer we must be able to handle the extraordinary complexity of these devices (not to mention the circuits containing hundreds of transistors). To do so we attempt to *linearize* a non-linear behavior.

1.1 The Three Terminal Tomato Model[1]

We begin by considering a generic three terminal black box (**Fig. 1**), with one terminal grounded. Assume it has an output current given by some a function f of the other two terminal voltages. Note that we make no assumption whatsoever about the linearity of whatever is inside the box. To linearize the system, we can take a Taylor expansion (trust me, this little bit of math will become immensely useful).

$$I_{out}(V_{in}+\delta V_{in}, V_{out}+\delta V_{out}) \simeq f(V_{in}, V_{out}) + \left.\frac{\partial I_{out}}{\partial V_{in}}\right|_{V_{out}} \delta \mathrm{V_{in}} + \left.\frac{\partial I_{out}}{\partial V_{out}}\right|_{V_{in}} \delta \mathrm{V_{out}} \qquad (1)$$

Rohit Julian Braganza
email: rbraganza@berkeley.edu

[1] This is inspired by Prof. Kristofer Pister's *Potato Model of a Circuit*

Let's look carefully at the derivatives in (1). The first derivative gives a *change in current due to a small change in the input voltage*, with the output voltage held constant. We define this as the *transconductance*, G_m.

$$G_m \equiv \left.\frac{\partial I_{out}}{\partial V_{in}}\right|_{Vout} \tag{2}$$

The other derivative in our two-variable taylor expansion relates the *change in output current due to a change in the output voltage*, with the input held constant. This is simply the *output conductance* of the box, G_o, by definition!

$$G_o \equiv \left.\frac{\partial I_{out}}{\partial V_{out}}\right|_{V_{in}} = \frac{1}{R_o} \tag{3}$$

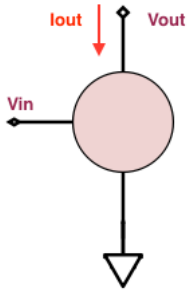

Fig. 1 The three terminal black box (or tomato), with with one terminal grounded. Note the direction of Iout.

Note that by convention instead of writing δV_{in} and δV_{out} we use the lower case v_{in} and v_{out} instead. Remember however that v_{in} and v_{out} represent *small signal values, around a fixed operating point*, V_{in} *and* V_{out}.

Assume we have an ideal voltage probe connected to the output; Note that the constants (I_{out}, equal to $f(V_{in}, V_{out})$) cancel on both sides of equation (1). Thus, substituting into (1):

$$\begin{aligned} G_m v_{in} + G_o v_{out} &= 0 \\ \implies \frac{v_{out}}{v_{in}} &= -\frac{G_m}{G_o} = -G_m R_o \end{aligned} \tag{4}$$

The critical points to take away from this:

- $-G_m R_o v_{in}$ is equal to v_{out}, the small output signal voltage for *any* three terminal box (with one grounded): the box can be an op-amp, a carbon nanotube, a potato, a tomato, or Schrödinger's cat, it does not matter. This will hold so long as the magnitude of v_{in} and v_{out} are sufficiently small for the Taylor approximation to hold.

- Note that we do not assume f to be linear. Rather we choose an *operating point*, V_{in}, V_{out}, around which we linearize. Setting this operating point is referred to as DC biasing.
- $-G_m R_o$ can therefore be thought of as the "small signal gain or transfer function," albeit an inverting one. This is also called the AC gain, since the small signal of interest is often a cosine of small amplitude[2].
- A small detail: Note that this derivation assumes a unilateral function, meaning the output does not affect the input (an example of this would be feedback). This assumption turns out to be fine except at very high frequencies. We will look at feedback later.

1.2 The Equivalent Tomato Model Circuit

We just found out that the Taylor expansion of our box/tomato results in $v_{out} = -G_m R_o v_{in}$. Let us try to create an equivalent "small signal circuit" in which we ignore the operating point (V_{in} and V_{out}), and concern ourselves solely with v_{in} and v_{out}. We first note that G_m represents a transconductance: $G_m v_{in}$ therefore produces a current, which is then multiplied by R_o to produce v_{out}. The resulting circuit is therefore:

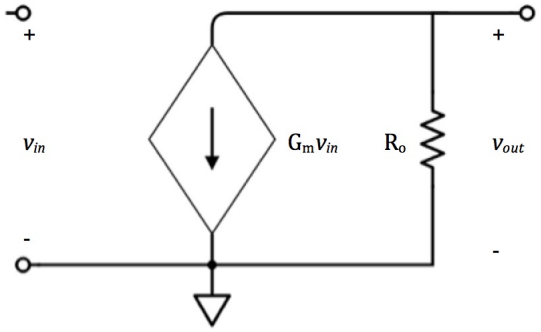

Fig. 2 The equivalent small signal (Tomato) model circuit. Note carefully the direction of the current used. Also note that so far our model does not consider any input impedance.

[2] We will explore the frequency response of circuits later.

2 Long Channel MOSFETs

We can finally tackle transistors! MOSFETs are governed by a complicated set of non-linear functions. Even worse, as FET channel lengths become shorter and shorter, quantum effects create nasty equations with no closed-form solutions. Luckily, we have our tomato model to rely on.

2.1 Small Signal Model for Long Channel NMOS Devices

Let's look at how MOSFETs provide gain through an example. First, we must decide the operating point in which we will use our N-channel MOSFET. We will have the gate be the input terminal, the drain the output terminal, and the source will be grounded. We will set the current going through the transistor with an ideal current source.

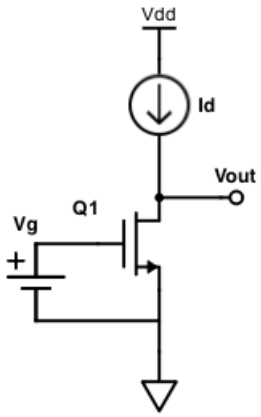

Fig. 3 Our setup: transistor Q1 is biased with a gate voltage Vg and its current is set by the current source Id.

Let's now look at a sample Id-Vd plot of a NMOS transistor. Remember we wish to maximize the small signal gain.

How to find the Transconductance and Output Resistance graphically
The transconductance is: $G_m \equiv \frac{\partial I_{out}}{\partial V_{in}}\Big|_{Vout}$. For the above NMOS transistor configuration, this is the change in I_{DS} for a small ΔV_{gs} at some fixed V_{DS}.
The Output resistance is: $R_o \equiv \frac{\partial V_{out}}{\partial I_{out}}\Big|_{V_{in}}$. This is one over the slope of the Id-Vd curve.

From **Fig. 4** we can see the regions of maximum $G_m R_o$ occur in the *saturation* operating regime. Consider the point $Id = 100$A, $Vds = 7$V. We can see a ΔVg of

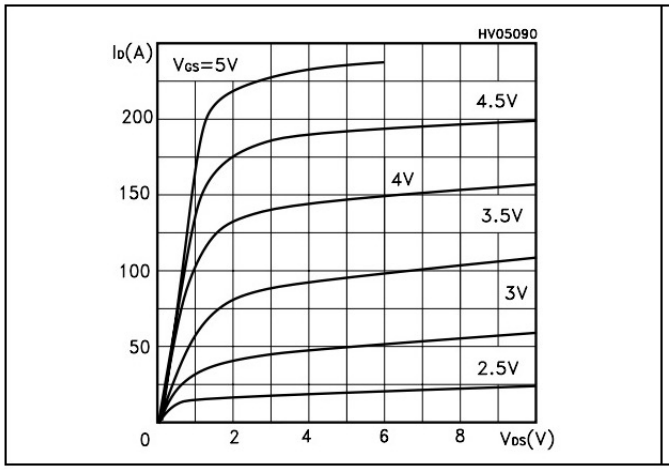

Fig. 4 The Id-Vds graph of the STD60NF3LL Power NMOS transistor.[1]

-0.5V results in a ΔId of -50A, for a G_m of 10 A/V! As for the output resistance: a ΔVd of 1V results in a ΔId of about 0.5A, for an output resistance of 2Ω. Our total small signal gain therefore is about -20.[3] Convince yourself that if we were to bias ourselves in the triode region, we would get little to no transconductance or output resistance.

To gain some intuition as to where this "gain" comes from, let's look back at what we did. We set I_d in **Fig. 3** to be 100A; we saw in **Fig. 4** that when ΔVg = -0.5, the transistor now wants to conduct *10 times as less* current. However the current source *demands* that the current through the MOSFET remain 100A! The transistor therefore must increase its Vds (the output voltage remember) *dramatically* to maintain the same current, and it must increase its Vds much, much more than the ΔVg provided (The graph runs out of room to show where it would intersect!). This is gain!

MOSFET Small Signal Equations

We have discovered that when biased in the saturation regime, MOSFETs provide small-signal gain. Now we will formulate two easy equations that will allow any analog designer to do quick, back of the envelope calculations to find the intrinsic gain of any long channel MOSFET.

[3] The reader might be a little worried for their safety at this point: 10A is enough to electrocute a human, 100A would probably fry one to a cinder. Note though that these are *power* MOSFETs. In integrated circuits, a typical MOSFET will conduct only *microamps* of current. A typical Gm for an IC is around 0.1mA/V, and Ro is tens of kΩs.

Refresher: Requirements for Saturation (Long-Channel NMOS)
We must have:

$$V_{DS} > V_{GS} - V_{Th}$$
$$V_{GS} > V_{Th}$$

For a long channel (square law) NMOS FET in saturation, its I-V characteristic is given by:

$$I_D = \frac{W}{L}\frac{\mu_n C_{ox}}{2}(V_{GS} - V_{Th})^2(1 + \lambda V_{DS})$$

Using definition (2) for G_m, and using the same setup used in Fig 3, with V_{in} at the gate, V_{out} at the drain, we get:

$$g_m \equiv \left.\frac{\partial I_{out}}{\partial V_{in}}\right|_{Vout} \implies g_m = \frac{W}{L}\mu_n C_{ox}(V_{GS} - V_{Th})(1 + \lambda V_{DS}) \tag{5}$$

Note that by convention, we use the lowercase g_m to denote the transconductance of the MOSFET itself. Uppercase G_m is used to denote the transconductance of a circuit.
For simplicity, we drop the λV_{DS} error term, to give:

$$g_m = \frac{W}{L}\mu_n C_{ox}(V_{GS} - V_{Th}) = \frac{2I_D}{V_{GS} - V_{Th}} \tag{6}$$

Similarly, using definition (3), we find the output resistance:

$$r_o = \frac{W}{L}\frac{\mu_n C_{ox}}{2}(V_{GS} - V_{Th})^2\lambda \simeq \frac{1}{\lambda I_D} \tag{7}$$

We can now almost create a complete small-signal equivalent circuit for a MOSFET- we are only missing the input resistance and parasitic capacitances. For a MOSFET, the gate is an insulating oxide, meaning (at low frequencies) it has an infinite input resistance. As for the parasitic capacitances, let's observe the structure of a MOSFET, as seen in **Fig. 5**. In addition to the substrate capacitances, in saturation the fringing electric fields result in parasitic capacitances given by:

$$C_{gs} = \frac{2}{3}WLC_{ox} + WC'_{ov} \tag{8}$$
$$C_{gd} = WC'_{ov} \tag{9}$$

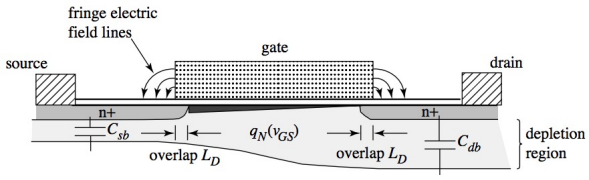

Fig. 5 An NMOSFET, in saturation. Note the fringing capacitances between the gate and the source/drain regions.[2]

We can thus make our small signal model, shown in **Fig 6**. Note that it excludes the substrate capacitances, we will consider them later.

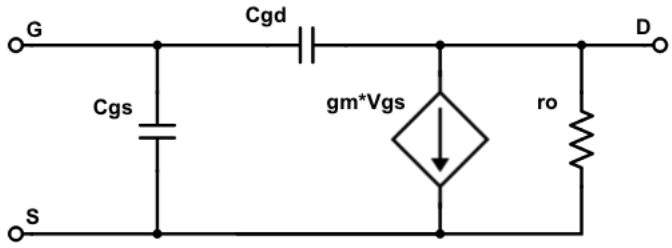

Fig. 6 A simple but effective Small-Signal Model for an NMOS transistor.

3 Bipolar Junction Transistors

Bipolar junction transistors, known as BJTs, are another type of transistor widely used in analog circuitry. We will quickly derive their equivalent circuits in the Tomato Model.

3.1 NPN Small Signal Model

An NPN transistor biased in Forward-Active mode has the following current characteristic equation, where V_T is the thermal voltage (26mV at room temperature):

$$I_C = I_s e^{\frac{V_{BE}}{V_T}} [1 + \frac{V_{CE}}{V_A}] \tag{10}$$

Refresher: Requirements for Forward Active Mode
The actual device-physics requirements are not particularly useful for the analog designer. Instead, use these guidelines[4]:

$$I_B > 0 \quad \& \quad I_C > 0$$
$$V_{CE} > V_{CE,sat} \simeq 200mV$$

Also remember that unlike the MOSFET, the BJT is a *current controlled device*, meaning an input bias current must be provided.

$$I_B = \frac{I_C}{\beta} \tag{11}$$

Lastly, because the BJT has an exponential collector current equation, similar to a diode, its base-emitter voltage is essentially constant across a wide current range.

$$V_{BE} \simeq 700 - 800mV \tag{12}$$

Using the definitions (2) and (3) found in the tomato model, we can find the transconductance and output impedance respectively[5].

$$g_m = \frac{I_C}{V_T} \tag{13}$$
$$r_o \simeq \frac{V_A}{I_C} \tag{14}$$

As a current controlled deivce, the BJT now has a finite input impedance. The small signal input impedance can be found by the partial derivative:

$$g_{in} \equiv \left.\frac{\partial I_B}{\partial V_{BE}}\right|_{V_{out}=0} = 1/r_{in} \tag{15}$$

This results in an input impedance, by convention denoted as r_π for a BJT, of:

[4] The device physics standpoint says: $V_C > V_B > V_E$ for forward-active operation. In reality, the collector can (and will) drop below the base and it will still work okay as long as V_{CE} is above the saturation voltage. It is therefore always important to look at the I-V curves of the devices you are given to see the regions of acceptable operation.

[5] Note that for the output resistance we have ignored the effect of V_{CE}, similar to what we did for the MOSFET.

$$r_\pi = \frac{\beta}{g_m} \tag{16}$$

The BJT contains similar parasitic capacitances as the MOSFET, between the Collector-Emitter junction and the Base-Emitter junction. Unlike the MOSFET, the dimensions of a BJT are often defined by the process used, and therefore the equations for these capacitances are not terribly relevant for the designer.[6,7]
Our small signal model is thus:

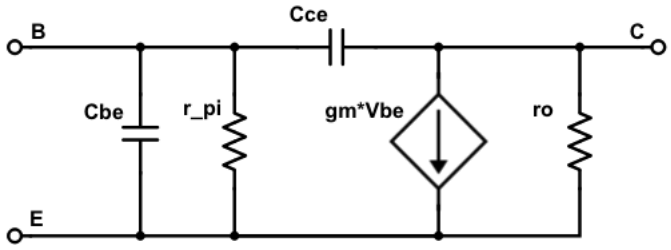

Fig. 7 The simple yet effective Small Signal Model for a NPN BJT

3.2 BJTs versus MOSFETS

What exactly are the differences

- MOSFETs are voltage-controlled devices, whereas BJTs are currrent-controlled.
- MOSFETs can be cheaper process-wise (not always true anymore).
- MOSFET lengths and widths are designer controlled, and can be made to be much smaller than BJTs. They can therefore be made with smaller parasitic capacitances[8] .
- MOSFETs are fantastic as switches, whereas BJTs are terrible.

[6] General trends are more important to remember: namely that if you are given the choice, larger BJTs will have larger capacitances.

[7] I will label these capacitances by C_{ce}, C_{be}, etc. Note though that sometimes they are referred to by other various Greek terms that are not worth memorizing.

[8] This implies that MOSFETs can be *faster*... wait a bit to learn more![9]

[9] There are actually special hetero-bi junction (HBT)-usually SiGe- BJTs that have transit frequencies of around 100 GHz, but these are fancy processes.

- BJTs provide superior intrinsic gain for the same power consumption, compared to a MOSFET. This is due to their higher transconductance and large Early voltages.

Do not worry if the differences do not make complete sense yet; it will all come together later as we continue to explore circuit design!

4 Summary

So far, the reader should understand how we have derived small signal circuits and their resulting gain for *any* type of device. In particular, we looked at MOSFETs and BJTs. In the next notesheet, we will begin to see how we can use these models in actual circuits!

Acknowledgements

All schematics were made with Digikey's free Scheme-it applet. Also thanks to prof. Ali Niknejad for his comments.

References

1. STD60NF3LL Datasheet. `https://www.pololu.com/file/0J57/STD60NF3LL.pdf`
2. Figure: Mosfet Capicitances in Saturation. `http://www.prenhall.com/howe3/microelectronics/pdf_folder/lectures/mwf/lecture12.fm5.pdf`