EC410 Experiment #4

One Week Experiment

I-V CHARACTERISTICS OF THE METAL-OXIDE-SILICON FIELD-EFFECT TRANSISTOR

Topics Relevant to this Lab Experiment:

Field-Effect Transistors

5.2.1 Physical Structure of the N-Channel Enhancement-Mode MOSFET
5.2.2 Summary of V-I Equations of the N-Channel Enhancement-Mode MOSFET
5.2.4 Nonzero Source-to Substrate Voltage (The Body Effect)
5.4.1 Upward Slope of FET V-I Characteristics

Goal: To analyze current-voltage relationships of the metal-oxide-silicon field-effect transistor, or MOSFET. The MOSFET is by far the most widely used active element in digital integrated circuits. It is the essential component of practically all random-access memory (RAM) chips used in computers.

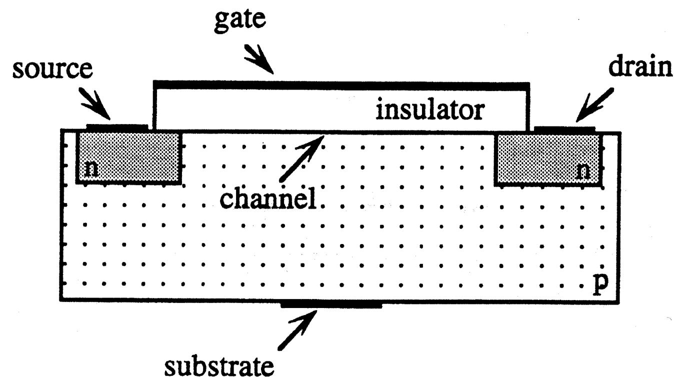
Background:

Three-terminal circuit elements offer a capability not available with the two-terminal ones studied so far; they are active elements, which are able to deliver more power to a load connected to one pair of their terminals than they receive from a control source connected to a second pair. The additional energy is supplied from a separate source connected to the third pair. These active elements may be visualized as valves that require a small amount of energy to open, but which, once open, allow the flow of a large amount of energy. This process underlies many kinds of voltage and current amplifier, which in one form or another are found in almost every electronic system.

Practically all silicon integrated circuits contain either MOSFETs or bipolar junction transistors (BJTs) as the active elements. Although most integrated circuits use only one of the two elements, newer fabrication technology allows the manufacture of chips that contain both. Although they have superficially similar i-v characteristics, they operate by entirely different mechanisms. In the BJT, to be studied in Experiment 11, the essential factor is the forward-biased base-emitter junction, whose current controls a larger collector current. The MOSFET is a purely voltage-controlled element, in which the potential drop across an insulating layer on the silicon surface controls current flow in the silicon. Moreover, the BJT (both the npn and the pnp form) requires the flow of both electrons and holes, whereas the MOSFET requires only one type of carrier.

An n-channel MOSFET, or NMOS, is shown below. There is also a p-channel, or PMOS form, in which the doping in each of the three regions of the silicon is reversed There are therefore structural similarities between the NMOS and npn transistors and between the PMOS and pnp transistors. The insulator is usually a film of silicon dioxide less than 50 nm (500 Ã) thick. The MOSFET is actually a four-terminal element, since the substrate also plays a part in its operation.

The substrate of the NMOS is always held at the most negative potential in the circuit (most positive for the PMOS) so that the source-substrate and drain-substrate junctions are never forward biased. Be sure that this is true whenever you build a MOSFET circuit.



To simplify the description of the operation of an NMOS, consider the source and substrate to be held at ground potential. When the gate potential is zero, and the drain potential is either positive or negative, no current flows through the channel between source and drain because one of the pn junctions is reverse biased. If the gate potential is now made positive, Gauss' Law requires that the positive charge on the top surface of the insulator be balanced by an equal negative charge on the bottom surface, which is in the silicon channel. For low gate potential this is accomplished by forcing positive holes away from the insulator-silicon interface, leaving behind the negatively charged ionized acceptor centers. At higher gate potential, this depletion layer reaches a maximum size; the additional negative charge is supplied by free electrons that enter the channel from the n-type source and drain. An inversion layer is now formed, which acts as an n-type region because the free carriers are electrons. The junctions at the source and drain effectively disappear, and current can now flow from drain to source if the drain is at a positive potential.

The minimum potential difference between the top and bottom of the gate insulator needed to form an inversion layer is called the threshold voltage V_{TR}. Since the source is at the lowest potential, this means that the minimum condition for current to flow is

$$v_{GS} > V_{TR}.$$

In normal operation, the drain is more positive than the source. At any point along the channel we have, by KVL,

$$v_{GS} = v_{G-channel} + v_{channel-S}$$

At the drain, which is the most positive point in the channel, this becomes

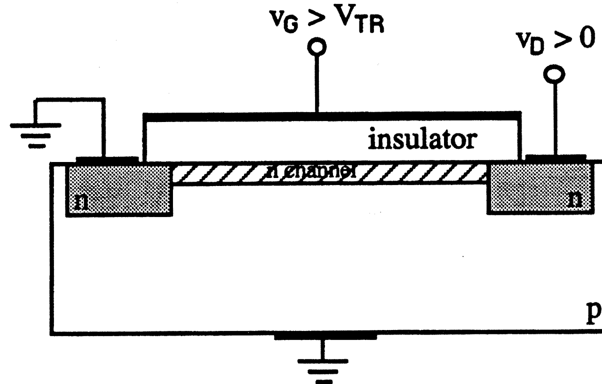
$$v_{GS} = v_{GD} + v_{DS} \quad \text{or} \quad v_{DS} = v_{GS} - v_{GD}$$

For fixed v_{GS}, as long as v_D is low enough so that the gate-to-drain potential difference v_{GD} exceeds

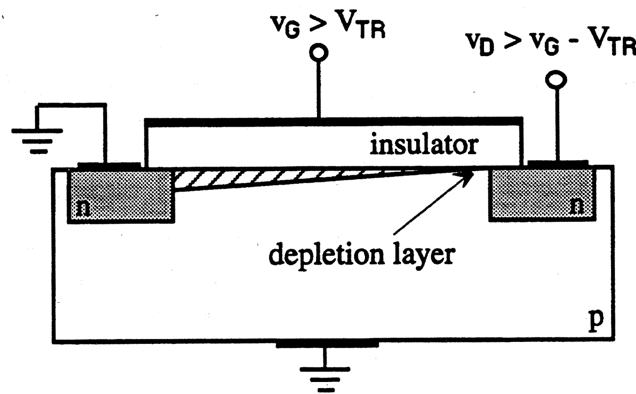
V_{TR}, the inversion layer extends across the entire channel, and the drain current is given by

$$i_D = K\left[2(v_{GS} - V_{TR})v_{DS} - v_{DS}^2\right], \quad \text{where} \quad v_{DS} < v_{GS} - V_{TR}$$

This is the <u>triode region</u> of operation



When v_{DS} is increased, with v_G unchanged, v_D eventually becomes large enough so that the the voltage difference v_G - v_D is now less than V_{TR}. The voltage drop across the insulator near the drain is no longer large enough to sustain an inversion layer (Fig. 9-3). There is now a depletion layer at that location,



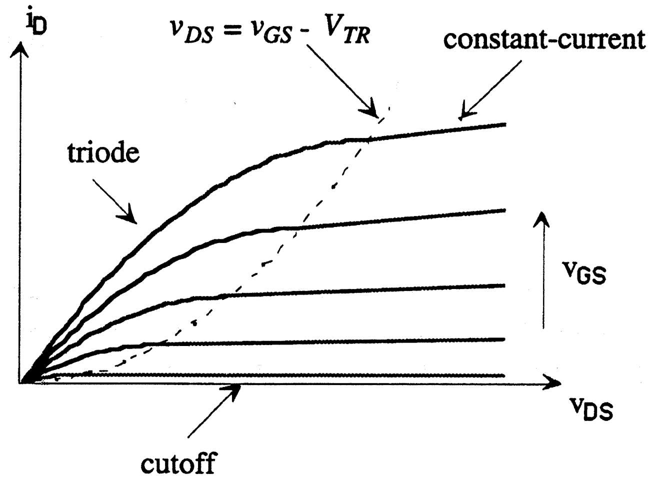
whose built-in field acts in the same way as the collector in the BJT (Experiment 5). Electrons that reach it from the channel are swept into the drain and into the external circuit. The current no longer depends upon v_{DS}; it is now given by

$$i_D = K(v_{GS} - V_{TR})^2, \quad \text{where} \quad v_{DS} > v_{GS} - V_{TR}$$

This is the <u>constant-current region</u>, sometimes called the saturation region. The latter term can be confusing because it misleadingly suggests a parallel with the saturation region of the BJT. The transition between the constant-current region and the triode region occurs where

$$v_{DS} = v_{GS} - V_{TR} \qquad (9\text{-}6)$$

Figure 9-4 shows the i-v characteristic of an NMOS. As in the case of the BJT, the PMOS characteristics have the same shape, but lie in the third quadrant, with all signs reversed.



The cutoff region corresponds to the condition $v_{GS} \leq V_{TR}$. The locus of the boundary points between the triode and constant-current regions is shown by the dashed curve. As in the case of the BJT, there is a slight dependence of i_D on v_{DS} in the constant-current region; the lines converge to a point called the Early voltage. Each curve corresponds to a different value of v_{GS}. Note that they are not equally spaced for equal increments of the controlling variable, as they are in the BJT, because of the square-law relationship.

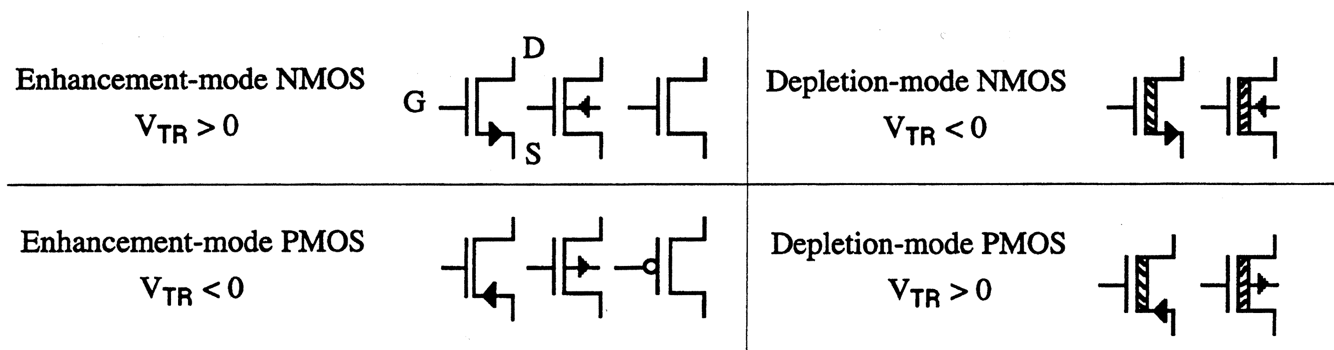
When the source-to-substrate junction is reverse biased, the depletion layer between the channel and the substrate becomes wider, causing the channel to become "pinched off" from the bottom, reducing its conductance. A larger v_{GS} is needed to return the current to its zero-bias value. The reverse bias has the effect of increasing the effective threshold voltage of the device. This body effect is described by equation (9-7).

$$V_{TR} = V_{TR0} \pm \gamma\left[\sqrt{|v_{SB}| + |2\phi_F|} - \sqrt{|2\phi_F|}\right] \qquad (9\text{-}7)$$

The parameters γ and ϕ_F are determined by the structure and dopant concentrations of the MOSFET. V_{TR0} is the nominal threshold voltage in the absence of the body effect, and v_{SB} is the source-to-body (substrate) potential. The positive sign applies to the NMOS and the negative sign to the PMOS.

A major difference between the BJT and the MOSFET is that the "threshold" for turning on the BJT is fixed by the properties of the pn junction at about 0.6 V, whereas the threshold voltage V_{TR} of the MOSFET can be controlled by introducing dopants into the channel. In fact, it can be made positive or negative for both kinds of MOSFET. This leads to four different kinds of

MOSFET, as shown below.



The commonly used circuit symbols are shown for each type. The arrows are placed either at the source or at the substrate. When at the source, they show the direction of current flow; when at the substrate, they point from the p region to the n region, as they do in the pn-junction diode symbol.

The i-v curves for enhancement-mode and depletion-mode devices are indistinguishable. The only difference is that the curve that corresponds to $v_{GS} = 0$ coincides with the x axis (cutoff) in the enhancement-mode case, but is above the x axis ($i_D \neq 0$) in the depletion-mode case. The characteristic for the PMOS has the same shape as for the NMOS, except that the curves lie in the third quadrant, where the signs of all voltages and currents are reversed.

The MOSFET is a voltage-controlled element, as opposed to the BJT, which is current-controlled. No significant gate current flows at the MOSFET gate terminal because it is connected to an insulator. Another difference between the two devices is that the MOSFET is truly symmetrical, as suggested by several of the circuit symbols whereas the BJT is not. The source and drain may be freely interchanged, providing that the substrate is not permanently connected to one of them.

The CD4007 integrated-circuit chip is a useful vehicle for this experiment because it contains both NMOS and PMOS transistors (although only enhancement-mode versions). The chip is mounted in a 14-pin DIP package with the terminals of six transistors accessible at the pins as shown in Fig. 9-6. The PMOS devices are in the top row of the diagram, and the NMOS in the bottom row.

Note that the substrate, or body, for each PMOS is to be connected to the most positive point in your circuit, and the substrate for each NMOS is to be connected to the most negative point. If you fail to do this, one of the pn junctions will be forward biased, and probably destroyed. The source is connected internally to the substrate in the two leftmost devices. Therefore, for

measurement of the body effect you must use one of the other devices so that you may control the potential difference v_{SB} between the source and the body.

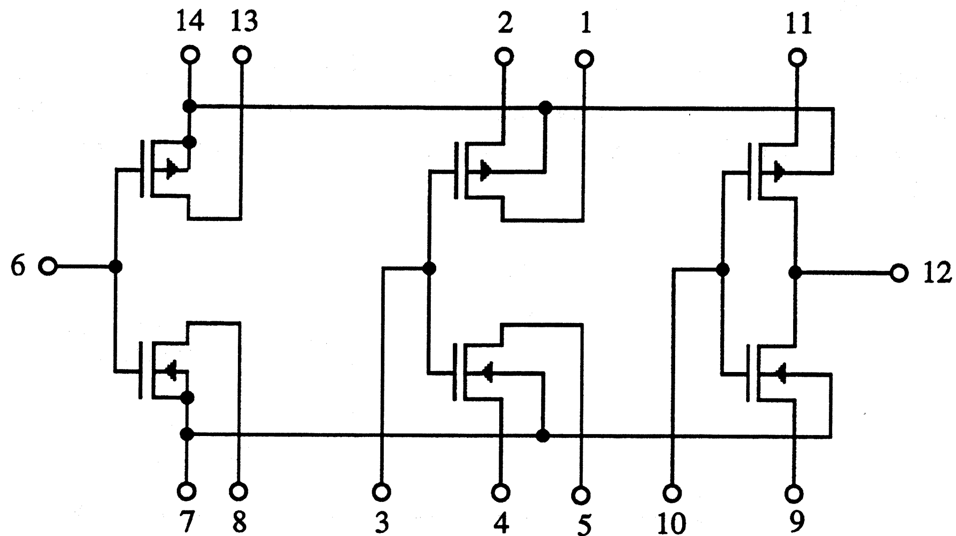


Fig. 9-6

I. **Measurement of K, V_{TR}, and γ**

Construct the circuit shown in Fig. 9-7, using one of the NMOS devices on the CD4007 chip that does not have an internal connection between source and body. If you have never used a multi-pin integrated circuit (IC) before, consult Appendix C, which gives the pin connections to the "mini-dip" package of the LM741 op amp. Be sure to double check your wiring, especially to the IC, before turning on the power in the lab.

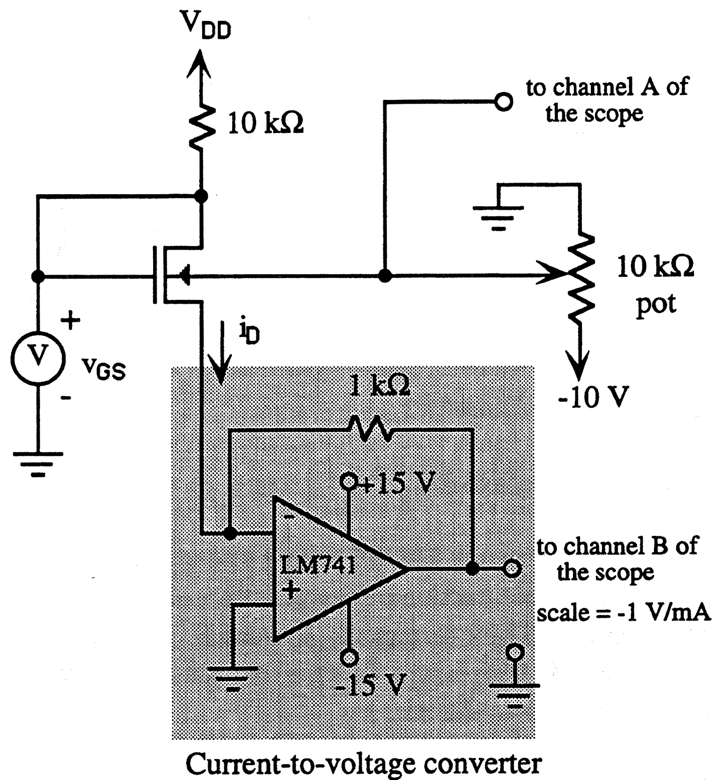
Since the NMOS has its gate and drain connnected together, it always operates in the constant-current region, because $v_{DS} > v_{GS} - V_{TR}$ when $v_{GS} = v_{DS}$ and $V_{TR} > 0$ (enhancement-mode device). The potentiometer connected to a negative voltage allows the body potential to be set anywhere between zero and -15 V. The feedback circuit effectively clamps the source potential at ground.

1) Set the body potential to zero and increase V_{DD} until a measurable drain current is reached. Measure i_D for several values of $v_{GS} = v_{DS}$ up to about 10 V.

2) Repeat the sequence of of i_D vs v_{GS} measurements for different values of v_{SB} up to about 10 V. Be sure that the substrate is never positive with respect to the source.

Level 2: Measurement of the i-v characteristic

In this part you will measure the characteristic curves of i_D vs v_{DS} for different v_{GS}. Build the circuit of Fig. 9-8, using either the NMOS transistor on the CD4007 chip that has the internal connection between source and substrate, or one of the others with an external connection between these two terminals.



Current-to-voltage converter

Fig. 9-7

1) Before connecting the signal generator to the circuit, observe its output directly on the scope. Using the dc offset capability of the generator, set the output so that it is about equal to a 15 V p-p waveform, offset by a dc voltage of about +7.5 V, so that the output range is between 0 and +15V.

2) Now connect the signal generator to the circuit. Note that the signal generator functions as a varying supply voltage V_{DD}, which allows the load line imposed on the transistor to sweep across values of open-circuit voltage between 0 and 15 volts, thereby continually sweeping across various operating points of the transistor.

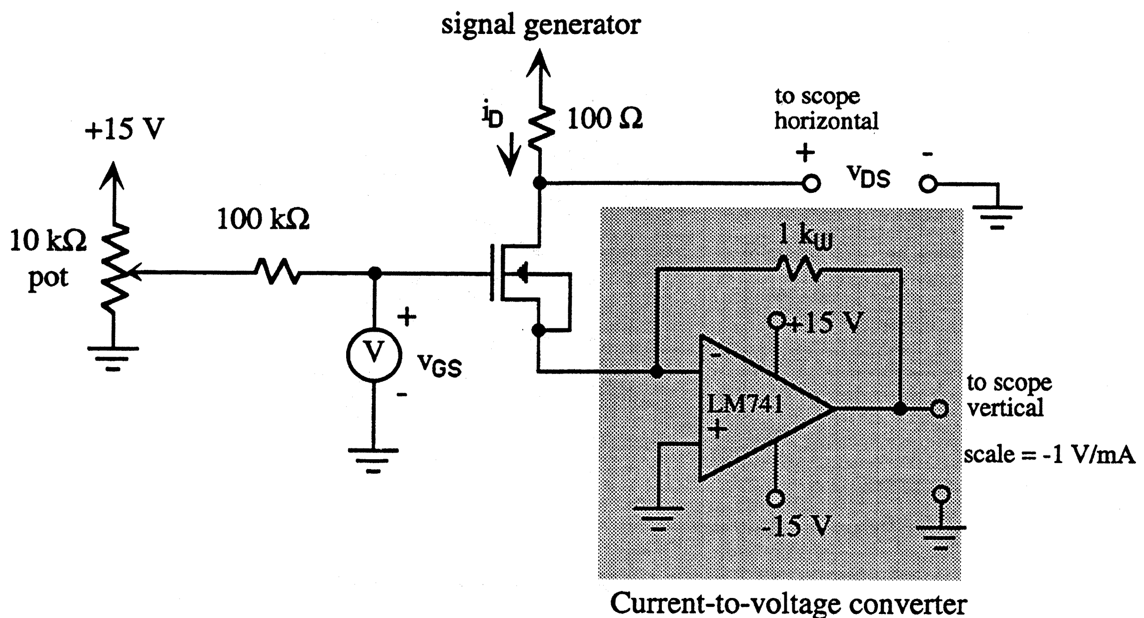


Fig. 9-8

3) Set the oscilloscope to the X-Y mode, with each channel set to 1V/div. Set the vertical channel on "invert", so that the negative voltage output of the op amp circuit will register as a positive deflection of the scope. Connect the horizontal channel between the drain lead of the NMOS and real ground (NOT the op amp input lead), and connect the vertical inverted channel of the scope between the op amp output lead and ground. In this way the scope will display the transistor output-port current vs output-port voltage. Gradually increase the setting of the potentiometer, recording the value of v_G as measured on the voltmeter. Record in your notebook the resulting transistor i-v curve for several values of v_{GS} (note that $v_S \approx 0$ because of the negative feedback).

4) Repeat steps 1 through 3 using one of the PMOS devices with its source and substrate connected together. The circuit should be modified as shown in Fig. 9-9 so that the signs of all the terminal voltages are reversed. In this case, set the output of the voltage generator connected the drain so that it is about equal to a 15 V p-p waveform, offset by a dc voltage of about -7.5 V, so that the output range is between 0 and -15V. Just as in the NMOS case, the source is held at virtual ground by the op amp feedback circuit. The potentiometer allows the gate voltage to be set negative with respect to the source. The vertical channel of the oscilloscope should again be inverted so that the i-v curves will be displayed in the third quadrant, where v_{DS}, v_{GS}, and i_D are negative.

Identify the transistors that you used for these measurements so that you can use them later if necessary.

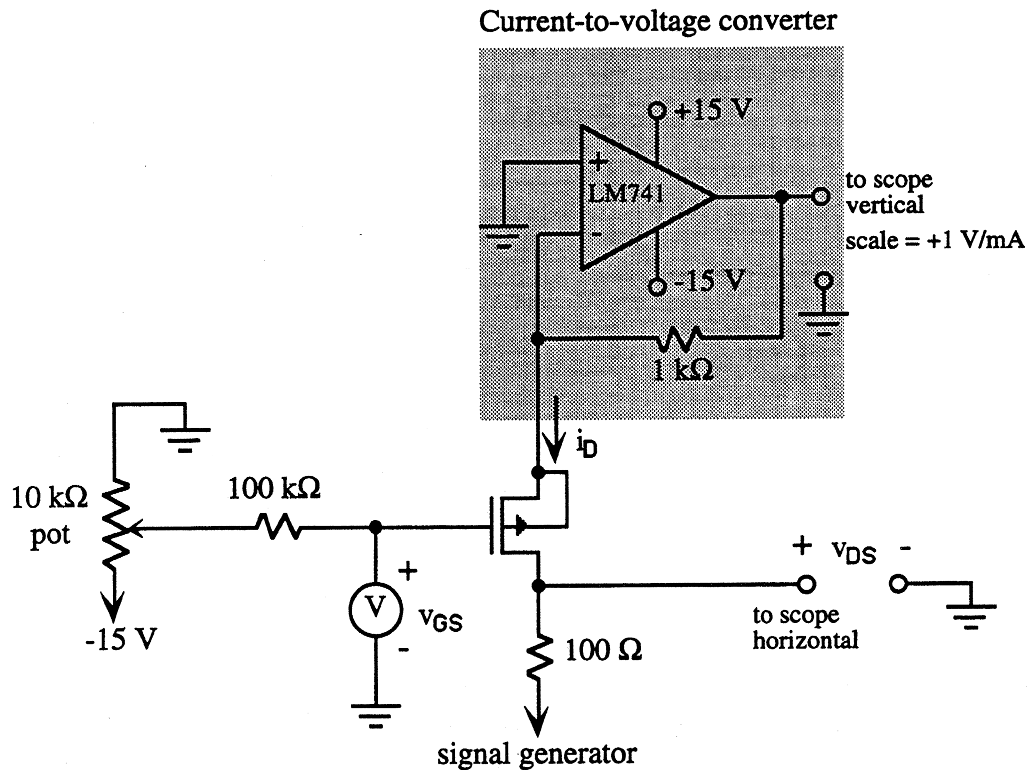


Fig. 9-9

Level 3: Data Analysis

A) Using the data fromLevelI for the NMOS, obtain values for K and V_{TR} from a plot of the square root of i_D vs v_{GS} for each value of v_{SB}. Rearranging equation (9-5), we get

$$\sqrt{i_D} = \sqrt{K}v_{GS} - \sqrt{K}V_{TR} \qquad (9\text{-}8)$$

Above the threshold voltage, the plots should be straight lines with slope = $K^{0.5}$. The intercept at $i_D = 0$ is V_{TR}.

B) Calculate a value for the NMOS body coefficient γ from (9-7), using the calculated values of V_{TR} for different v_{SB}, and 0.3 V as an appropriate value for ϕ_F.

C) For both types of transistor, estimate the value of v_{DS} where the triode and constant-current regions meet. Is it consistent with equation (9-6) for the NMOS case when you use your calculated value for V_{TR}?

D) For both types of transistor, try to calculate the slopes of the i-v curves in the constant-current region for several dfferent values of v_{GS}. The inverse of the slope is often designated r_0, the incremental resistance in this region. The magnitude of the slope should increase with increasing $|v_{GS}|$. If the slopes can be determined with reasonable accuracy, you should find that all the i-v

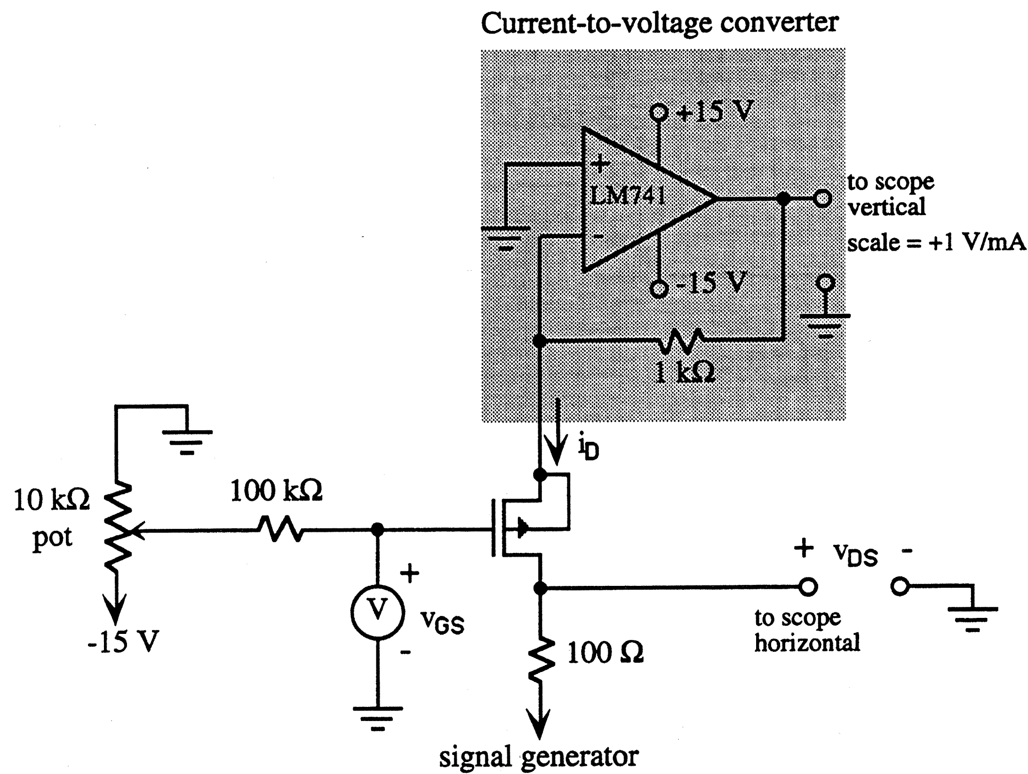


Fig. 9-9

E) Using the data from Part I for the NMOS, obtain values for K and V_{TR} from a plot of the square root of i_D vs v_{GS} for each value of v_{SB}. Rearranging equation (9-5), we get

$$\sqrt{i_D} = \sqrt{K} v_{GS} - \sqrt{K} V_{TR} \qquad (9\text{-}8)$$

Above the threshold voltage, the plots should be straight lines with slope = $K^{0.5}$. The intercept at $i_D = 0$ is V_{TR}.

F) Calculate a value for the NMOS body coefficient γ from (9-7), using the calculated values of V_{TR} for different v_{SB}, and 0.3 V as an appropriate value for ϕ_F.

For both types of transistor, estimate the value of v_{DS} where the triode and constant-current regions meet. Is it consistent with equation (9-6) for the NMOS case when you use your calculated value for V_{TR}?

G) For both types of transistor, try to calculate the slopes of the i-v curves in the constant-current region for several dfferent values of v_{GS}. The inverse of the slope is often designated r_0, the incremental resistance in this region. The magnitude of the slope should increase with increasing $|v_{GS}|$. If the slopes can be determined with reasonable accuracy, you should find that all the i-v curves in the forward-active region converge to approximately the same value of v_{DS} at $i_C = 0$. This point is called the Early voltage V_A. Try to find its value for your transistors.