Now let's look at the electrical view of the MOSFET.

Its operation is determined by the voltages of its four terminals.

First we'll label the two diffusion terminals on either side of the gate terminal.

Our convention is to call the diffusion terminal with the highest voltage potential the "drain" and the other lower-potential terminal the "source".

With this labeling if any current is flowing through the MOSFET switch, it will flow from drain to source.

When the MOSFET is manufactured, it's designed to have a particular threshold voltage, V_TH, which tells us when the switch goes from non-conducting or "open" to conducting or "closed".

For the n-channel MOSFET shown here, we'd expect V_TH to be around The P+ terminal on the left of the diagram is the connection to the p-type substrate.

For the MOSFET to operate correctly, the substrate must always have a voltage less than or equal to the voltage of the source and drain.

We'll have specific rules about how to connect up this terminal.

The MOSFET is controlled by the difference between the voltage of the gate, V_G, and the voltage of the source, V_S, which, following the usual terminology for voltages, we call V_GS, a shortcut for saying V_G minus V_S.

The first picture shows the configuration of the MOSFET when V_GS is less than the MOSFET's threshold voltage.

In this configuration, the switch is open or non-conducting, i.e., there is no electrical connection between the source and drain.

When n-type and p-type materials come in physical contact, a depletion region (shown in dark red in the diagram) forms at their junction.

This is a region of substrate where the current-carrying electrical particles have migrated away from the junction.

The depletion zone serves as an insulating layer between the substrate and source/drain.

The width of this insulating layer grows as the voltage of source/drain gets larger relative to the voltage of the substrate.

And, as you can see in the diagram, that insulating layer fills the region of the substrate between the source and drain terminals, keeping them electrically isolated.

Now, as V_GS gets larger, positive charges accumulate on the gate conductor and generate an electrical field which attracts the electrons in the atoms in the substrate.

For a while that attractive force gets larger without much happening, but when it reaches the MOSFET's threshold voltage, the field is strong enough to pull the substrate electrons from the valence band into the conduction band, and the newly mobile electrons will move towards the gate conductor, collecting just under the thin oxide that serves the gate capacitor's insulator.

When enough electrons accumulate, the type of the semiconductor changes from p-type to n-type and there's now a channel of n-type material forming a conducting path between the source and drain terminals.

This layer of n-type material is called an inversion layer, since its type has been inverted from the original p-type material.

The MOSFET switch is now closed or conducting.

Current will flow from drain to source in proportion to V_DS, the difference in voltage between the drain and source terminals.

At this point the conducting inversion layer is acting like a resistor governed by Ohm's Law so I_DS = V_DS/R where R is the effective resistance of the channel.

This process is reversible: if V_GS falls below the threshold voltage, the substrate electrons drop back into the valence band, the inversion layer disappears, and the switch no longer conducts.

The story gets a bit more complicated when V_DS is larger than V_GS, as shown in the bottom figures.

A large V_DS changes the geometry of the electrical fields in the channel and the inversion layer pinches off at the end of the channel near the drain.

But with a large V_DS, the electrons will tunnel across the pinch-off point to reach the conducting inversion layer still present next to the source terminal.

How does pinch-off affect I_DS, the current flowing from drain to source?

To see, let's look at some plots of I_DS on the next slide.

Okay, this plot has a lot of information, so let's see what we can learn.

Each curve is a plot of I_DS as a function of V_DS, for a particular value of V_GS.

First, notice that I_DS is 0 when V_GS is less than or equal to the threshold voltage.

The first six curves are all plotted on top of each other along the x-axis.

Once V_GS exceeds the threshold voltage I_DS becomes non-zero, and increases as V_GS increases.

This makes sense: the larger V_GS becomes, the more substrate electrons are attracted to the bottom plate of the gate capacitor and the thicker the inversion layer becomes, allowing it to conduct more current.

When V_DS is smaller than V_GS, we said the MOSFET behaves like a resistor obeying Ohm's Law.

This is shown in the linear portions of the I_DS curves at the left side of the plots.

The slope of the linear part of the curve is essentially inversely proportional to the resistance of the conducting MOSFET channel.

As the channel gets thicker with increasing V_GS, more current flows and the slope of the line gets steeper, indicating a smaller channel resistance.

But when V_DS gets larger than V_GS, the channel pinches off at the drain end and, as we see in on the right side of the I_DS plots, the current flow no longer increases with increasing V_DS.

Instead I_DS is approximately constant and the curve becomes a horizontal line.

We say that the MOSFET has reached "saturation" where I_DS has reached some maximum value.

Notice that the saturated part of the I_DS curve isn't quite flat and I_DS continues to increase slightly as V_DS gets larger.

This effect is called channel-length modulation and reflects the fact that the increase in channel pinch-off isn't exactly matched by the increase current induced by the larger V_DS.

Whew!

MOSFET operation is complicated!

Fortunately, as designers, we'll be able to use the much simpler mental model of a switch if we obey some simple rules when designing our MOSFET circuits.

Up to now, we've been talking about MOSFETs built as shown in the diagram on the left: with n-type source/drain diffusions in a p-type substrate.

These are called n-channel MOSFETs since the inversion layer, when formed, is an n-type semiconductor.

The schematic symbol for an n-channel MOSFET is shown here, with the four terminals arranged as shown.

In our MOSFET circuits, we'll connect the bulk terminal of the MOSFET to ground, which will ensure that the voltage of the p-type substrate is always less than or equal to the voltage of the source and drain diffusions.

We can also build a MOSFET by flipping all the material types, creating p-type source/drain diffusions in a n-type substrate.

This is called a p-channel MOSFET, which also behaves as voltage-controlled switch, except that all the voltage potentials are reversed!

As we'll see, control voltages that cause an n-channel switch to be "on" will cause a p-channel switch to be "off" and vice-versa.

Using both types of MOSFETs will give us switches that behave in a complementary fashion.

Hence the name "complementary MOS", CMOS for short, for circuits that use both types of MOSFETs.

Now that we have our two types of voltage-controlled switches, our next task is to figure out how to use them to build circuits useful for manipulating information encoded as voltages.