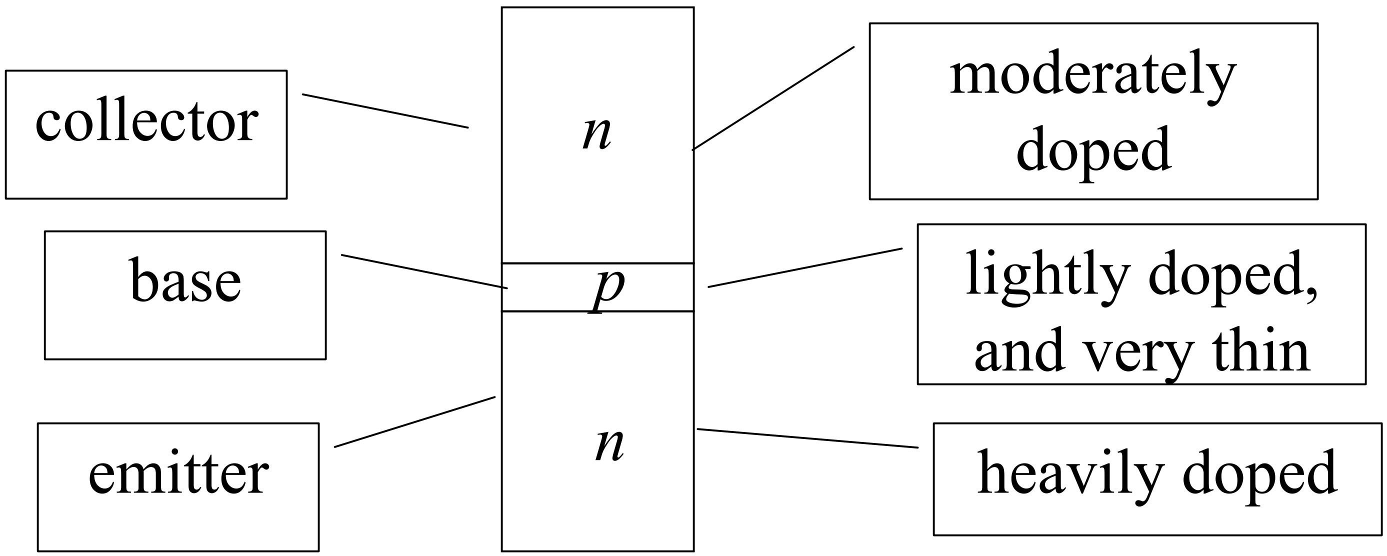
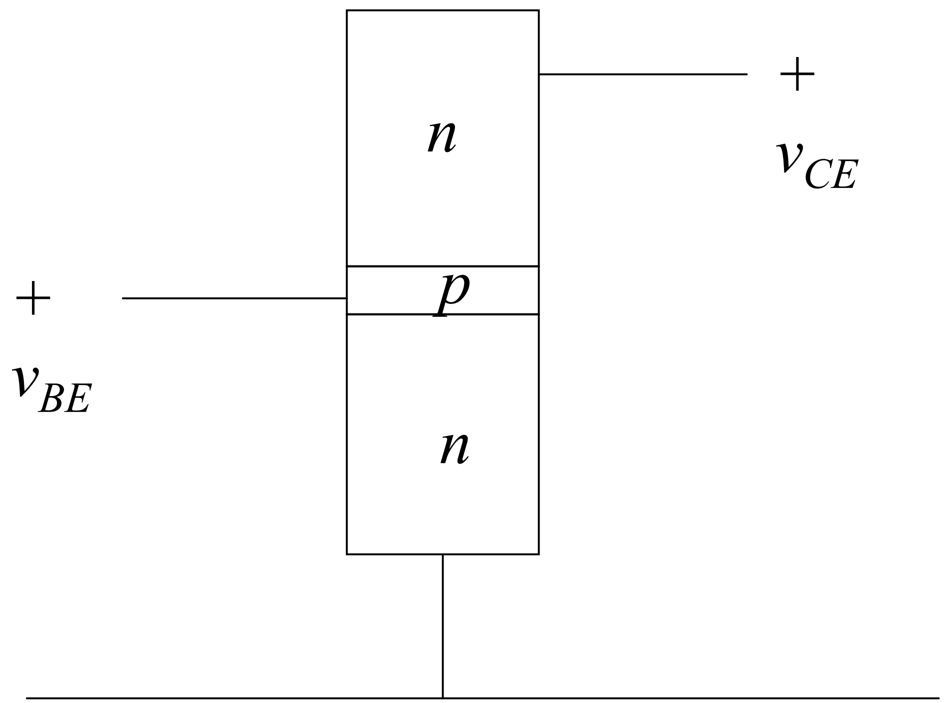
Lecture 3: Transistors

- Now that we know about diodes, let's put two of them together, as follows:



- At first glance, this looks like an insulator
  - but the actual behavior is far more interesting, if we apply external voltages properly

- Let's apply the following voltages:



- First, assume v_{CE} is 0. Then, if v_{BE} is bigger than the diode drop, a current flows through the forward-biased diode from base to emitter
  - call this current the "base current", i_B

- Now let's start cranking up v_{CE}
  - this attracts more electrons from the base to the collector
  - "collector current" i_c increases
  - this is called the "saturation region" of the transistor
  - at relatively small v_{CE}, most the electrons coming in to the base get scooped up by the collector
  - when this happens we enter the "active region" of the transistor
- Let's follow an electron up from the emitter in the active region
  - first it enters the base, where it has two choices:

  1. drop into a hole in the *p* doped base. This is called "recombining", and electrons that do this will end up contributing to i_B
  2. drift across the junction to the collector. Electrons that do this will end up contributing to i_C

- But recall that we made the base thin, and lightly doped (not many holes available)
  - the chances of recombining are not good!
- If the recombination probability is 1-α, the ratio of base and collector currents will be:

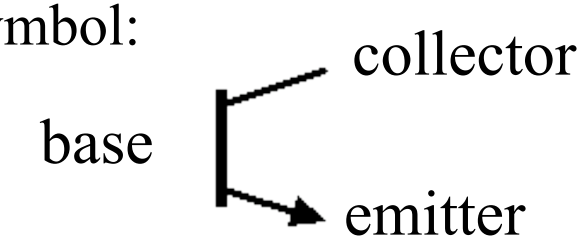
$$\frac{i_C}{i_B} = \frac{\alpha}{1-\alpha} \equiv \beta$$

$$i_C = \beta i_B$$

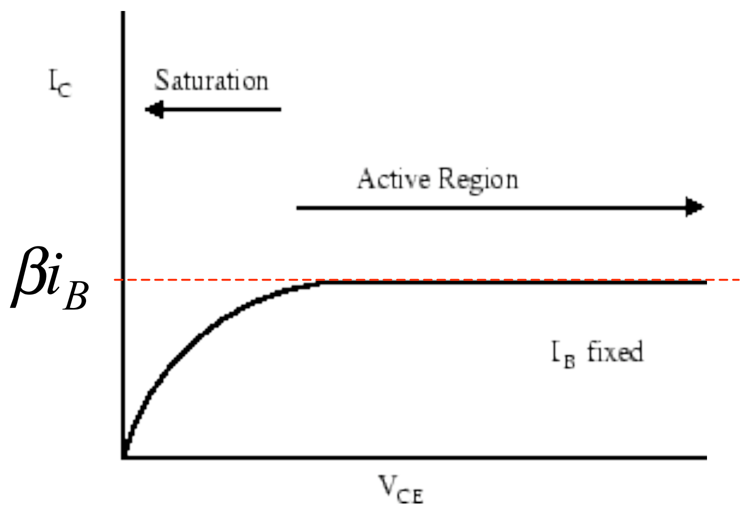
- Typically β is ~100
- Note what this means:
  - the transistor can control the (large) collector current by adjusting the (small) base current
  - the inverse is *not* true: the base current can't be changed much by adjusting v_{CE}

Transistor uses

- The type of transistor just described is a *npn bipolar junction* transistor
  - Schematic symbol:



  - One can also make *pnp* bipolar junction transistors
- Summary of i_c as a function of v_{CE} for a given i_B:



- If v_{CE} gets too big, the transistor breaks down
  - i_C becomes large
  - transistor might fry

- Transistor properties are useful for two reasons:

1. Can control large-power circuit with small-power input
2. Can isolate different regions of complex circuits
   - i.e., divide circuit into “input” section connected to base, and “output” section connected to collector
     - both input and outputs are connected to the emitter
   - This greatly simplifies the design of such circuits

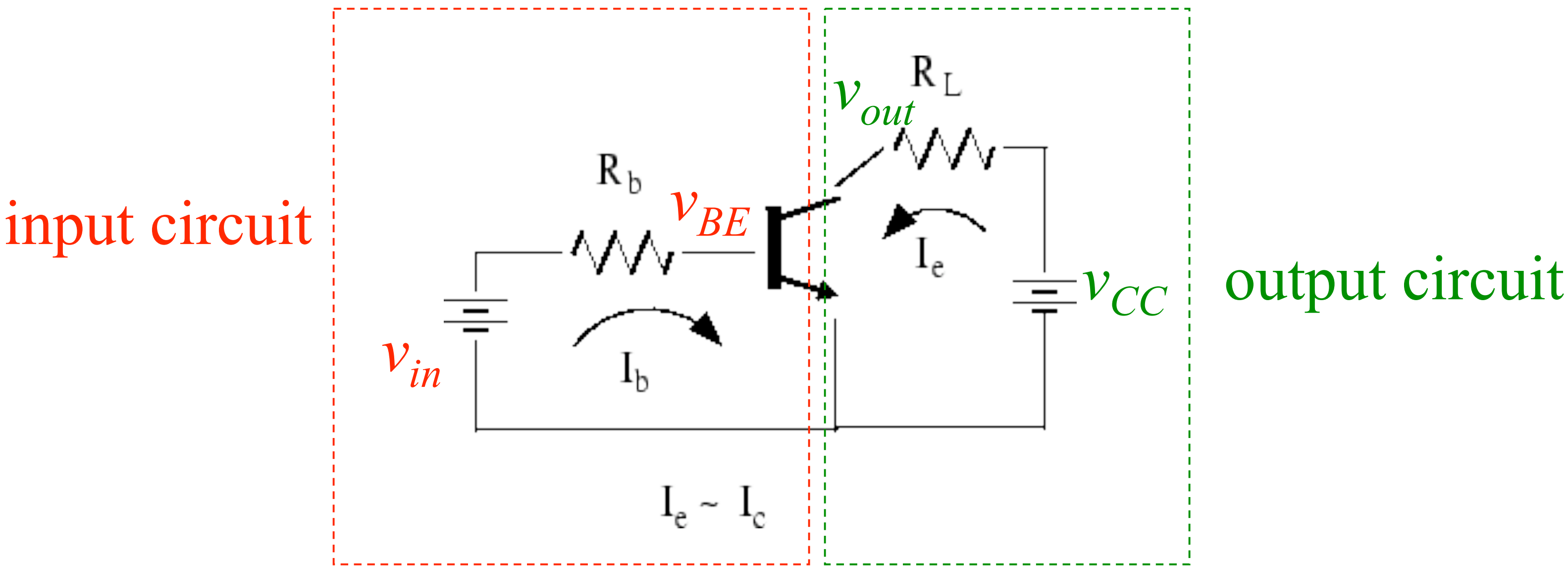
Transistor rules

- In order to take advantage of the nice behavior we want in the transistor, we must keep in mind the following rules:

1. V_C must be greater than V_E
2. Base-collector and base-emitter act like diodes
    - Base-emitter is forward-biased, base-collector is reverse-biased
3. $i_C \approx \beta i_B$
4. There are maximum values of i_B, i_C, and V_{CE} that can't be exceeded without destroying the transistor

Common emitter circuit

- One useful transistor circuit is the following:



- The input circuit can control what happens in the output circuit, but *not* vice-versa
- v_{CC} is a constant bias voltage
- want to see how v_{out} varies with v_{in}

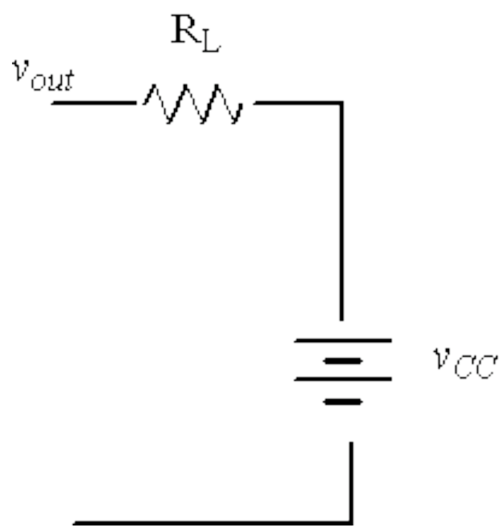
- First analyze the input circuit
- If v_{in} is less than the diode drop in the transistor (~0.7V):
  - $i_B = 0$
  - $v_{BE} = v_{in}$
- This is the *cutoff region* for the circuit
- For larger v_{in}, we have:

$$v_{BE} \approx \text{const} \approx 0.7\text{V}$$

$$v_{in} - i_B R_B - v_{BE} = 0$$

$$i_B = \frac{v_{in} - v_{BE}}{R_B}$$

- Now look at the output circuit
- First consider just the “load” part

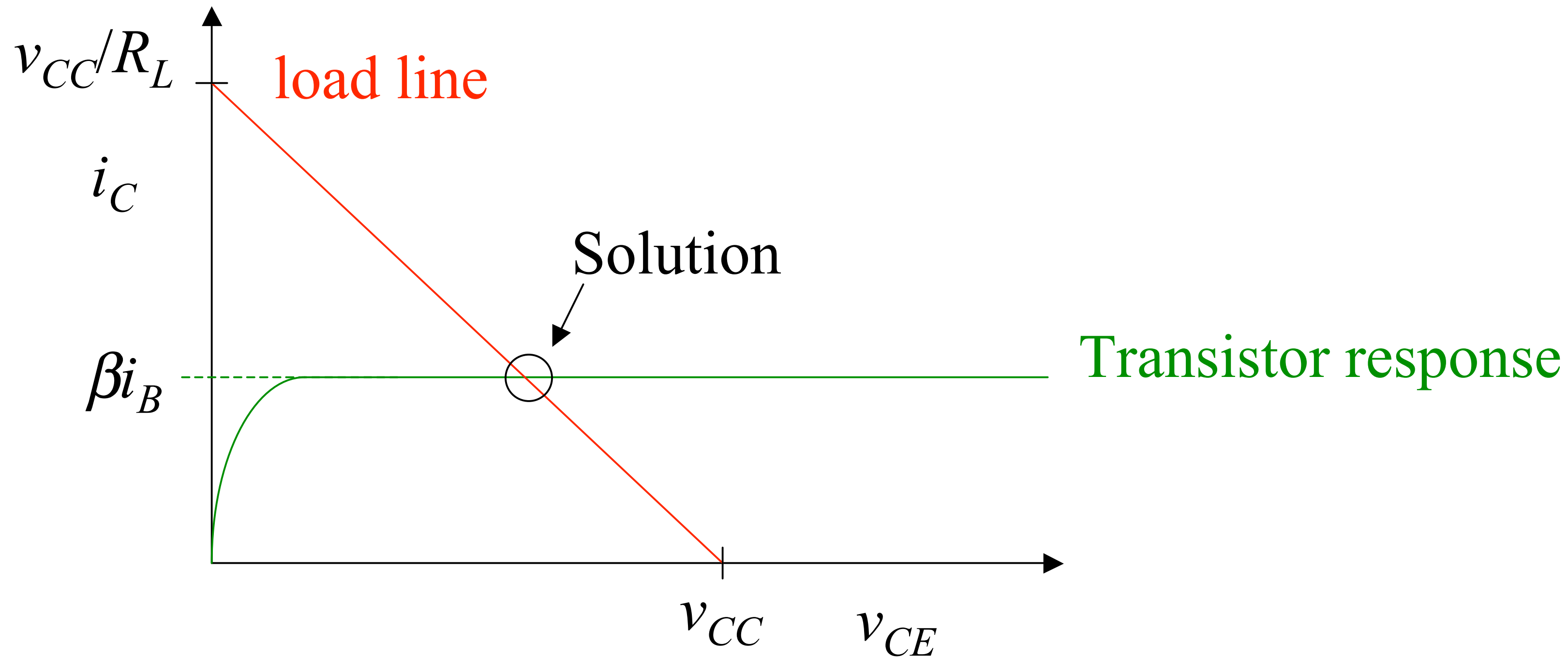


$$v_{CC} - i_C R_L - v_{out} = 0$$

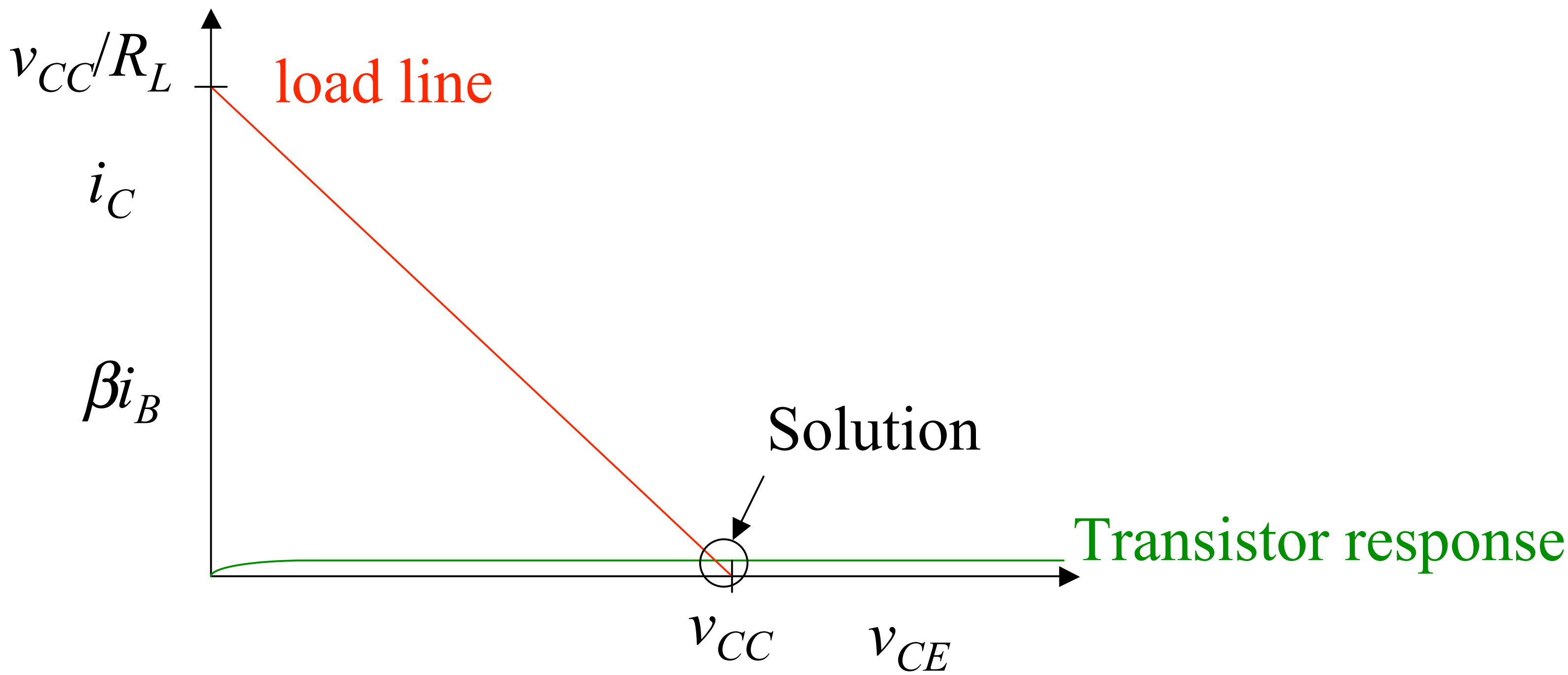
$$i_C = \frac{v_{CC} - v_{out}}{R_L}$$

- This linear dependence of i_C on v_{out} is called the *load line* for the circuit
- But we also know that due to the transistor, i_C depends on i_B, which in turn depends on v_{in}

- For a given base current, we can find i_C and v_{CE} by plotting both the load line and the transistor response curve on the same graph:

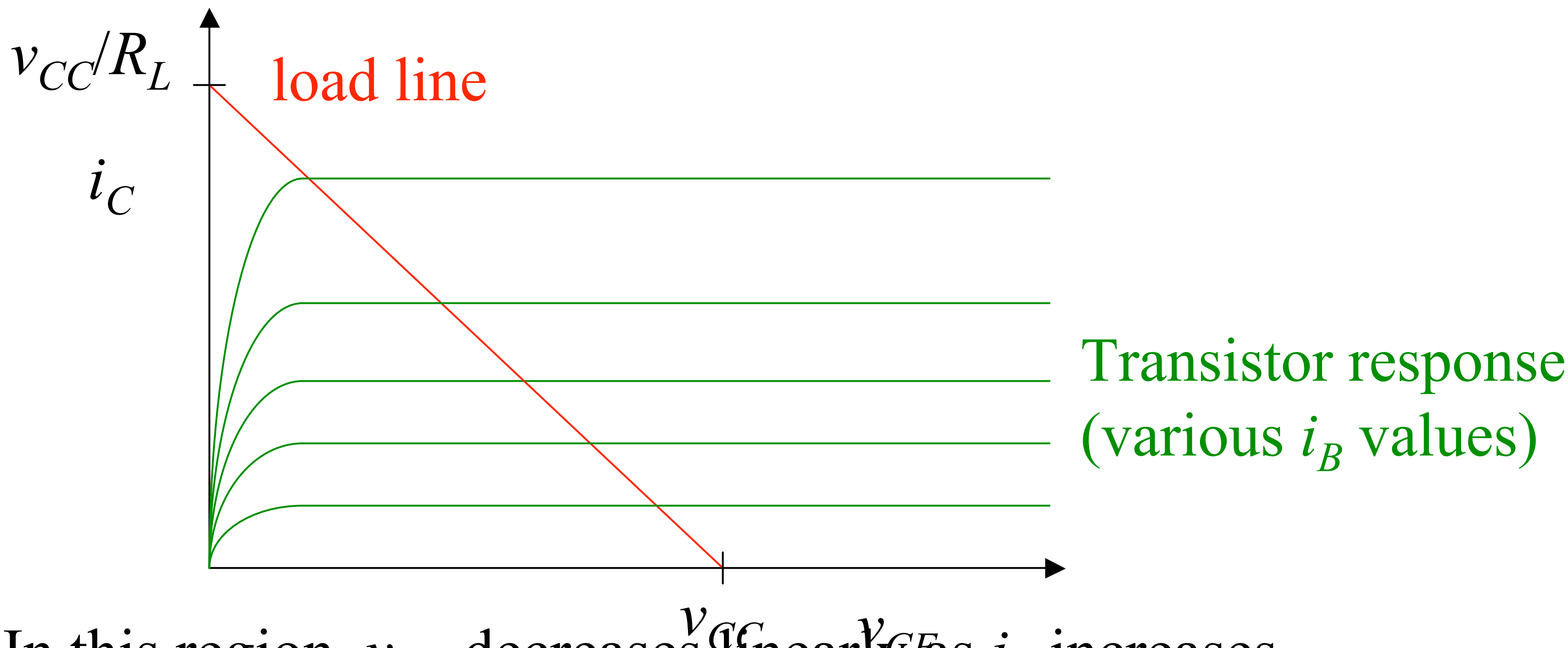


- If v_{in} is less than the diode drop, so there's no base current, the plot becomes:



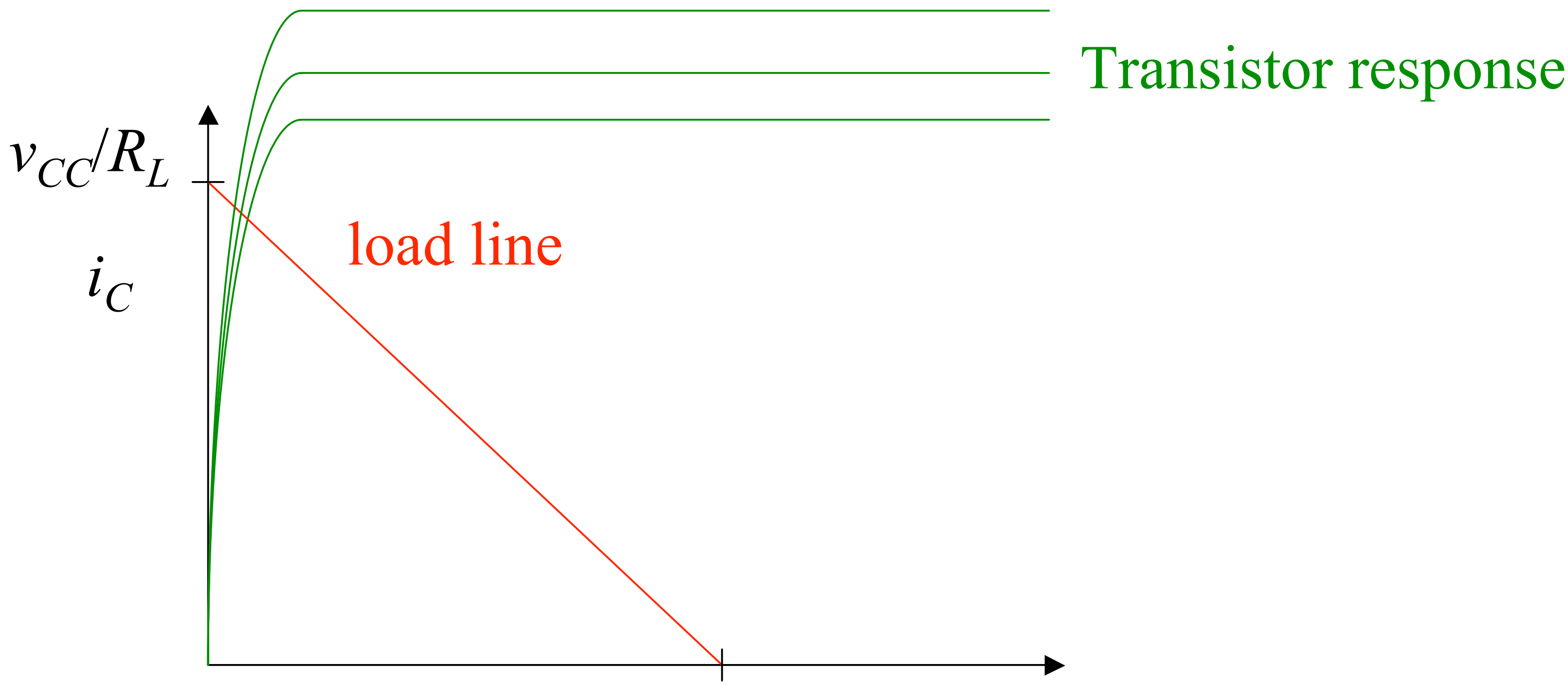
- The output voltage is near v_{CC}, independent of v_{in}
  – this is called the "cutoff region"

- Once v_{in} is large enough for the transistor to turn on, we enter the situation shown a couple of slides ago:



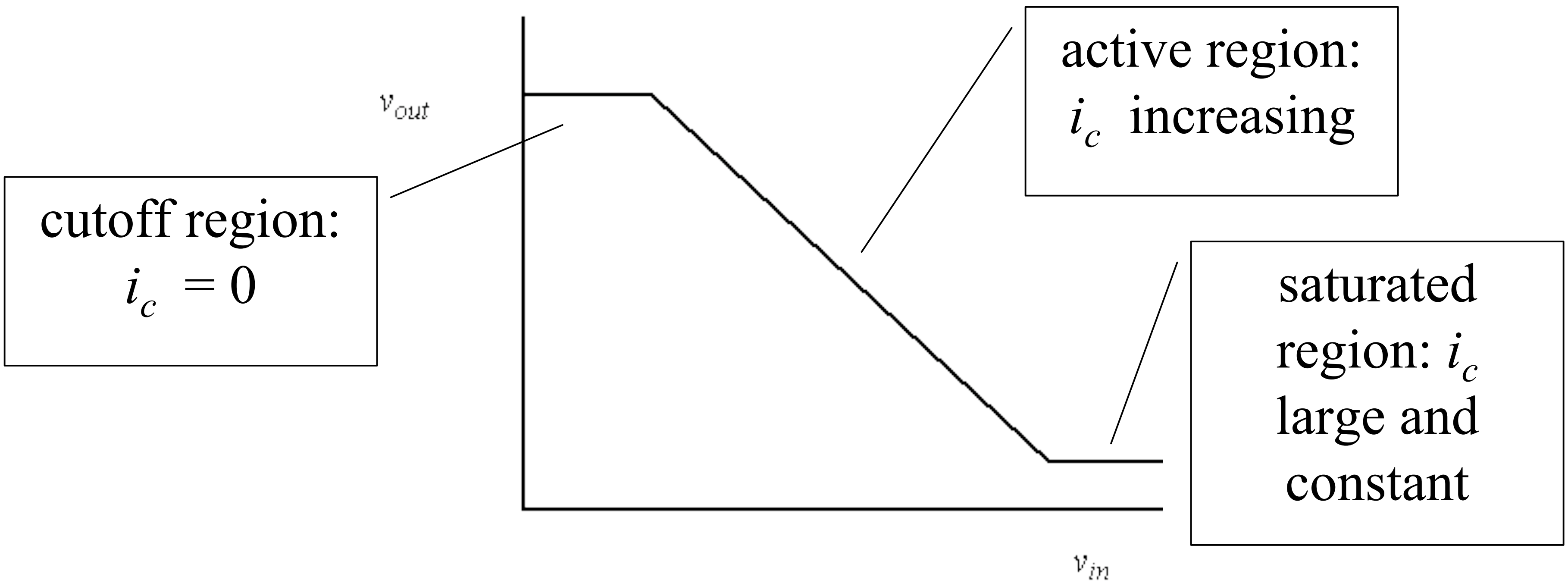
- In this region, v_{out} decreases linearly as i_B increases
  - and i_B increases linearly with v_{in}
  - this is the "active region"

- If we keep increasing v_{in}, we enter the following situation:



- Now the output voltage is small, and nearly independent of i_B
  - this is the “saturated region”

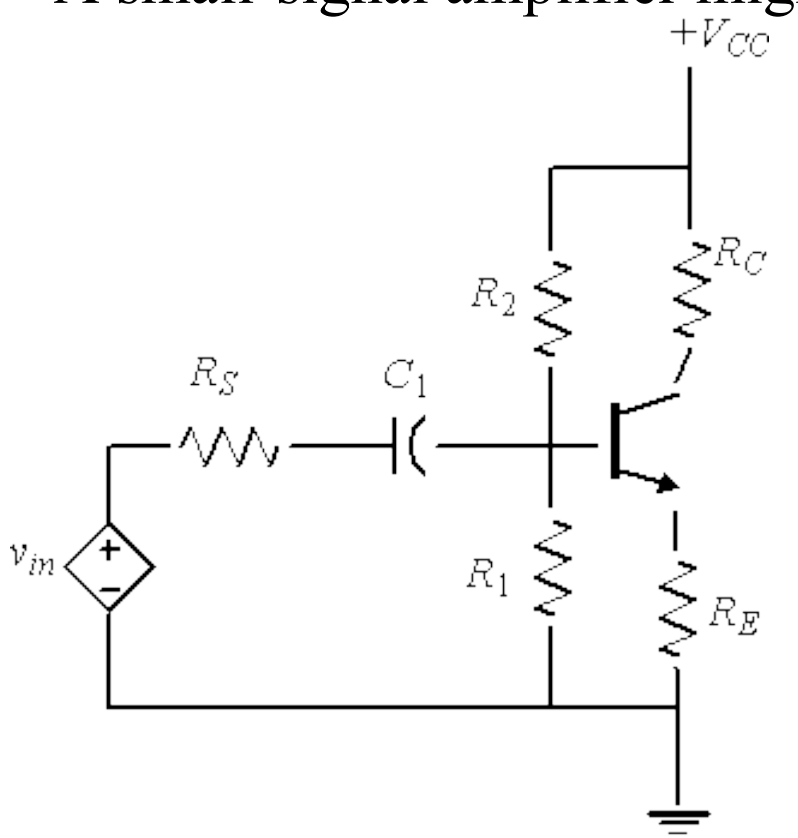
- Putting all of this together, we find that:



- As one goes from the cutoff region to the saturated region, the output circuit goes from OFF (no current) to ON (large current)
  - The transistor is acting like a switch!
- Transistor switches form the basis of digital electronics

Small signal amplification

- Amplifying signals is another very common use for a transistor
- "Small" means that the variations in the signal do not move the transistor outside of the active region
- A small-signal amplifier might look like:



- C_1 is a blocking capacitor
  - Keeps transistor in active region regardless of DC input voltage
  - Has very small impedance for the signal we want to amplify

Analysis of our circuit

- We'll set the circuit parameters as:
  - $R_S = 1\text{k}\Omega$, $R_1 = 5.6\text{k}\Omega$, $R_2 = 50\text{k}\Omega$, $R_C = 10\text{k}\Omega$, $R_E = 1\text{k}\Omega$
  - $V_{CC} = +10\text{V}$
  - Transistor $\beta = 100$
- First assume the signal generator is off, so all voltages are derived from V_{CC}
- V_{CC} is divided by R_1 and R_2 to give a voltage at the transistor's base of:

$$V_{BB} = V_{CC} \cdot \frac{R_1}{R_1 + R_2} = 1.0\text{V}$$

- This is greater than the 0.7V needed to start current flowing into the base

- To find the value of i_B, we divide the emitter voltage by the impedance given by R_1 and R_2 in parallel
  - good approximation since the internal impedance of the power supply is low, so both R_1 and R_2 can be considered as connected to ground

$$i_B = \frac{V_{BB} - 0.7}{R_1 \| R_2} = \frac{0.3\text{V}}{5\text{k}\Omega} = 60\mu\text{A}$$

  - This means that the collector current is:

$$i_C = \beta i_B = 6\text{mA}$$

- At this point we should verify that the transistor is in its active region
  - It is! See text for details…

- Looking at the output circuit, we have:

$$V_{out} = V_C = V_{CC} - I_C R_C$$

(true because the capacitors look like short circuits for the signals we care about)

- So the change in the output signal voltage is:

$$\Delta V_{out} = -R_C \Delta I_C$$

- The currents in the emitter and collector are nearly the same, so:

$$\Delta V_{out} = -R_C \Delta I_E$$

- The change in I_{E} is related to the change in V_E by:

$$\Delta I_E = \frac{\Delta V_E}{R_E}$$

- We also know that the emitter voltage is the base voltage – the diode drop, so:

$$V_E = V_{BB} - 0.7\text{V}$$
$$\Delta V_E = \Delta V_{BB} = \Delta v_{in}$$

- Which means that:

$$\Delta V_{out} = -R_C \frac{\Delta V_E}{R_E} = -\frac{R_C}{R_E}\Delta v_{in}$$

- For our example, this means that:

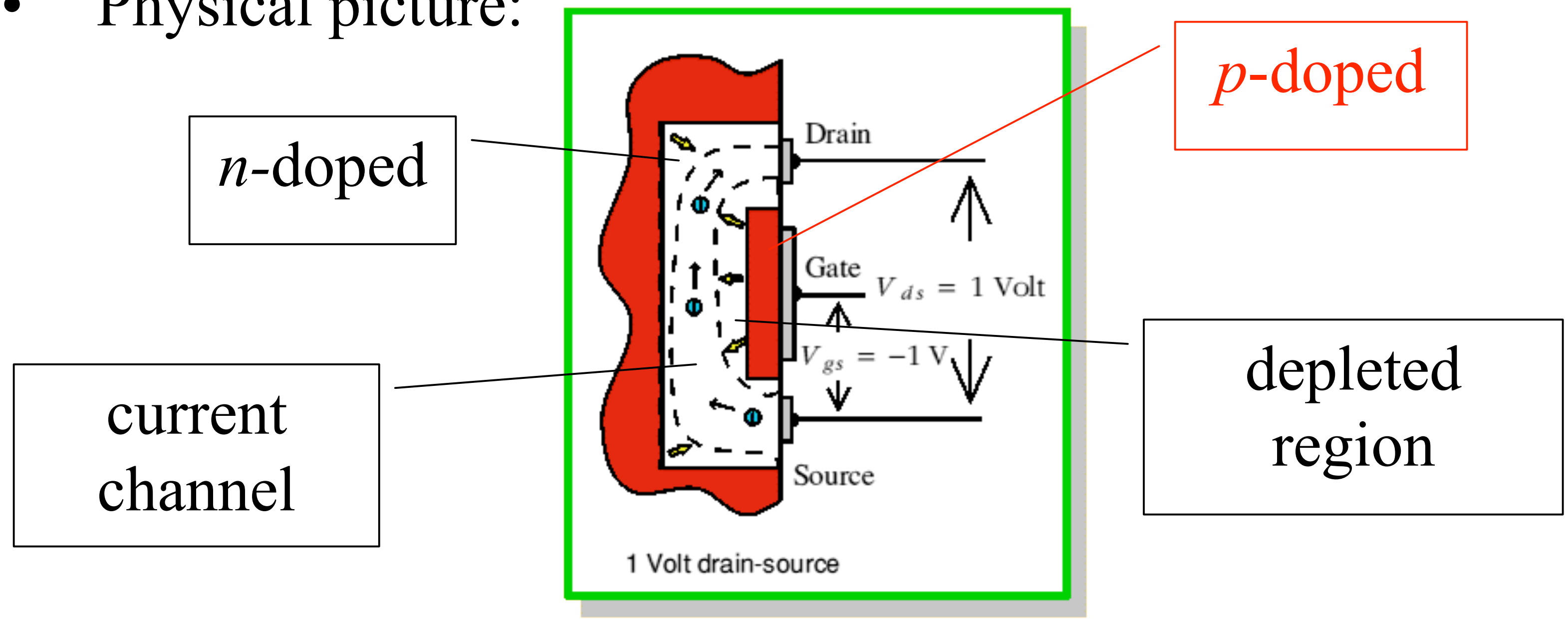
$$\Delta V_{out} = -R_C \frac{\Delta V_E}{R_E} = -10\Delta v_{in}$$

- The signal is amplified by a factor of 10!
  - the minus sign means the signal is also inverted

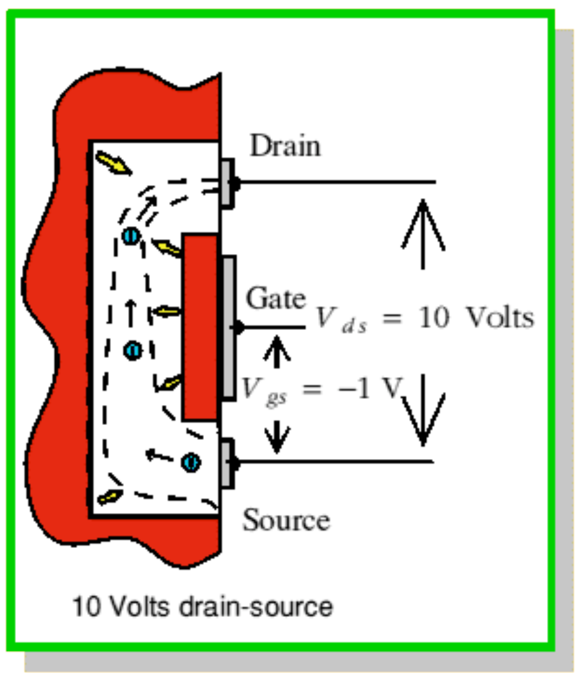
- Some notes:
  - The gain of the amplifier depends on the values of the resistors, *not* on the β of the transistor
  - That's a good design, since β can very substantially from transistor to transistor (even of the same model)
- What if we want a gain so large that the small-signal circuit can't be used?
  - just use the output as the input to another small-signal amp, and repeat as needed
  - only problem is that one will also be repeatedly amplifying noise on the signal
  - that's solved by the use of feedback (next week's lecture)

Field-effect transistors

- Another type of transistor is the field-effect transistor (FET)
  - Comes in two varieties
    1. junction FET (JFET)
    2. metal oxide semiconductor FET (MOSFET)
- They behave similarly, so we'll look at the JFET in detail
- Physical picture:



- It's basically a reverse-biased *pn* junction
  - no current through depleted region
  - means gate current i_G is zero – implies extremely large input impedance
- What happens as the gate voltage is increased?
  - The depleted region grows
  - The conduction channel gets smaller → resistance increases:



JFET operating regions

- To see how the JFET works, let's fix v_{GS} (it must be negative to reverse-bias the diode) and see what happens to i_D as v_{DS} increases
  - at first, i_D increases due to the increasing voltage
    - this is called the ohmic region, since the JFET behaves much like a resistor
  - but increasing v_{DS} also enlarges the depleted region, restricting current flow. Eventually current becomes constant as v_{DS} increases
    - this is the saturation region
  - if v_{DS} becomes very large, the transistor breaks down
- Note also that v_{GS} can be made more negative until the entire JFET is depleted – thus no current flows regardless of v_{DS}
  - this is the cutoff region of the transistor

JFET notes

- The “saturated” region of the JFET behaves similarly to the “active” region of the bipolar junction transistor
- FETs are useful because there is essentially no input current
  - Thus the output current can be controlled with nearly no input power
  - In this sense, FETs are more nearly ideal transistors than bipolar junctions are
- Integrated circuits (“chips”) are made by forming many FET’s on layers of silicon
- Main limitation of FETs is maximum current they can handle
  - For high-current applications the bipolar junction is a better choice