Transistors

History of Transistors

Introduced in 1948.

John Pierce credited for naming it: transfer resistance.

Large scale commercial use not until much later.

Early were germanium type. Took off when silicon transistors were developed.

First advantages were relatively low power consumption at low voltage levels which made large scale production of portable entertainment devices feasible.

Growth of batter industry paralleled transistors.

Introduction to Transistors?

"fastest acting fuse known to mankind"

Bipolar transistors are NPN more common or the rarer PNP

FET transistors have high input inpedance.

Include MOSFETS and JFETs.

Modern FETs are very rugged.

How do Transistors Work?

Certain materials can perform as solid state devices.

Impurities in silicon crystals add free electrons/conductivity.

Gallium and arsenic are impurities. Arsenic adds electrons, gallium adds holes/deficiencies.

N-type material conducts via free electrons, P-type material conducts via electron deficiency.

This material is called a semi-conductor.

They take a bar of p-type material and they dope the center to make it into n-type.

How do Holes and Electrons Conduct in Transistors?

Connect P-type and N-type materials, apply voltage, then current will flow.

Electrons attracted across the junction of the p and n material.

Current flow by electrons going one way and holes going the other way. If battery polarity reversed then current flow would cease.

Electron Flow in a p-n Junction of a Diode

Diagram 1

A junction of p and n types constitutes a rectifier diode.

With zero spacing between p and n junctions we have a relatively high value **capacitor**. **Varactor** is a diode used as a variable capacitor. Recall that capacitance proportional to **area of plates** and inverse to **distance between plates** and proportional to a **dielectric constant**. This construction places upper frequency limit at which the device will operate. This was a severe early limitation on transistors at radio frequencies.

Modern times: some bipolar transistors have F_t's (total frequency i.e., maximum excursion, highest operating frequency) beyond 1 Ghz. Capacitance at junction taken advantage of in form of varactor diodes.

A transistor is merely a sandwich of these devices (see **diagram 2**). The inter-terminal capacitance doesn't mean that transistors can't pass DC. Probably better to think about the capacitance being in parallel with the junction.

Transistor Amplification

Because collector current where voltage is relatively high is pretty much the same as the emitter current and also controlled by the emitter current (where the voltage is usually much lower) it can be shown by Ohm's law $P = I^2 * R$ that amplification occurs. A small amount of current at the base controls a large amount of current between the emitter and collector.

Current always flows into the arrow in a transistor. The arrow points in the direction of hole flow. The arrow being the symbol within the schematic symbol for a transistor; there is no actual arrow within the transistor.

Signals don't always come in through the base.

The NPN Transistor

NPN = “not pointing in”, **PNP** = “pointing in (perpetually)” (diagram 3).

Why use NPNs over PNPs? NPNs are quieter (at least that is historically true even if not still true); they work like vacuum tubes (always a positive at the collecting point of the tube (the **plate**)).

PNPs have more negative voltage at the collector than the emitter, while NPNs are the opposite, but don't commit this to memory because in circuits you will see many different things.

Differences in manufacturing and of much importance in the biasing.

The work horse is the NPN.

A silicon NPN transistor needs to be forward biased by about 0.65 volts for it to turn on.

FETs as Transistors

A J-FET and a dual gate MOSFET. Diagram 4.

Metal Oxide Substrate.

FET characterized by its extremely high input impedance.

Bipolar transistor has moderate input impedance depending on configuration while some FETs can/do have input impedances measured in megohms.

Advantages of high input impedance? A great input stage for a meter or oscilloscope.

Video is 2V P-P. If load it down, would get a lot lower voltage. High impedances present a lot smaller load so they don't affect the circuit nearly as much.

Bipolar transistors are essentially “current” amplifiers while FETs could be considered voltage amplifiers. Some controversy, but most people view them this way.

Testing Transistors (a side discussion)

(Transistors short out – the depletion region shorts out).

Checking transistors is like checking two diodes.

Two p-n junctions in an NPN transistor. If took a diode (n-p junction), put a meter on , diode will either appear as dead short or an open. This is characteristic of a diode. Diode has one p-n junction. The p is at the anode (the triangle in the symbol), the n is at the cathode (the straight line in the symbol). N is at emitter and collector, P is at base. Can check transistor same as a diode by putting voltage across base and emitter and base and collector. Always check both ways when testing a transistor. Will have an open one way and some nonzero resistance in the other way when testing b-c and b-e. Also good to test from collector to emitter since may detect an internal short.

Must check the transistor outside of a circuit because the associated resistances will conduct current and get in the way of testing.

Use an analog meter not a digital meter for this testing..

Most cases transistors melt down and dead short. Some cases where manufacturing flaw causes other problems, so then check AC in AC out and if load resistance is dropping the full voltage.

This will be on a test.

On employment exam, should be able to tell if transistor conducting and if it is NPN or PNP using an ohmmeter.

How are Semiconductors Made?

Consists of more than 100 steps, during which hundreds of copies of an IC are formed on a single wafer (a substrate).

Generally, the process involves the creation of 8 to 20 patterned layers on and into the substrate, ultimately forming the complete IC. This layering process creates electrically active regions in and on the semiconductor wafer surface.

What are Small Signal Amplifiers?

An amplifier, with or without negative feedback, having the greatest fidelity in faithfully reproducing the input with the least distortion.

This means the circuitry doesn't modify the signal (the waveform shape).

It is however the least efficient, in as much as the power delivered to the load is only a small percentage of the DC power used in the amplification process.

This means that an amplifier wastes power.

Amplifiers might use feedback, have good fidelity, have lousy efficiency.

Feedback prevents oscillations, controls fidelity.

Feedback

Where part of the output signal fed back to the input but 180 degrees out of phase (partially cancels the input).

If it were in-phase feedback then we would have an oscillator, which we don't want.

Diagram 6.

Steve had a tube audio amp that was oscillating (parasitic oscillation at 44 Khz using up 4/5 of the power of the amp).

Fidelity

Output must be an exact replica of the input but only magnified/amplified.

Distortion is when output is not an exact replica of the input.

Efficiency

Theoretical limit to this amplifier's efficiency is 50% so for every watt of output we must use up at least 2 watts DC power input to the amplifier.

Depending on the application this may or may not be significant. If miniaturization is important and battery supply is the only power source it is important.

50% efficiency is at best.

So why use class A (a type of amplifier design)? At low levels of signal and amplification, the losses through inefficiency are not significant and are far outweighed by the goal of fidelity of linearity.

Must find a very linear part of the operating range of the transistor.

Class A amp consumes power all the time, but is the most linear.

Some Transistor Basics and/or Characteristics of 2N2222A

This transistor is an NPN.

ID: a silicon NPN general purpose type

PD: 0.5 watts, meaning it capable of dissipating 500 mW. (kind of like a resistor). This is an absolute limit.

V_{ce} – 40V – meaning don't use it above 20V (preferably less) DC supply. Voltage between the collector and the emitter.

I_c – 0.8A meaning the maximum collector current allowed is 800 mA.

Hfe – 100 @ Ic of 150 mA. This means it has an **amplification factor** of 100. Also known as **beta** (the greek letter beta like a B with a tail) Diagram 7. Also known as gain. Beta has no units.

F_t – 300 Mhz this means that by the time the frequency reaches 300 Mhz the amplification factor has dropped off to 1. The maximum usable frequency for the transistor.

Case – TO-18 type package. Cases identified in some strange methods. It's what the metal case looks like. Aluminum or plastic.

Rule 1: do not always believe all of the specs.

Most specs are the conservative limits but important items such as the gain, Hfe, varies widely.

Gain figures may range from 60 to 170.

Always use your transistor well under published specifications.

Leads: base, collector, emitter. Schematic symbol may or may not have the surrounding circle.

On an NPN transistor, if the base is open, then a small amount of current flowing into the base will lead to a much larger current flowing in the collector.

The ratio is the gain of this particular transistor. Hence a transistor is known as a current amplifier.

Rule 2: do not take the EBC connections or leads of transistors for granted.

Different manufacturers have different pin-outs even for the same type. Always double check.

Can count on, if there is a tab near a pin, it is the emitter. Diagram 8.

Lab