

## CHAPTER 4

# Semiconductors (Diodes and Transistors)

### HERDING WILD ELECTRONS

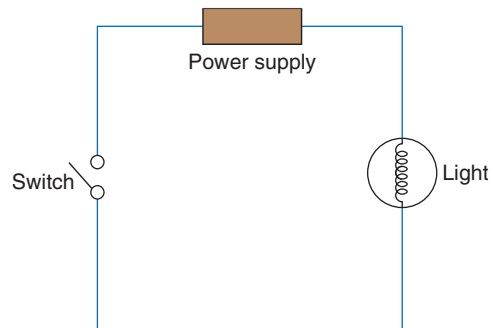
As we noted in *Chapter 2: Atoms, Molecules, and Crystals*, electricity may be considered to be vast herds of electrons migrating from one place to another, while electronics is the art and science of controlling these herds: starting them, stopping them, deciding where they can roam, and determining the activities they are going to perform while on their way. Ever since humans discovered electricity (as opposed to electricity—in the form of lightning—discovering us), taming the little rascal and bending it to our will has occupied a lot of thought and ingenuity.

The first, and certainly the simplest, form of control is the humble mechanical switch. Consider a circuit consisting of a switch, a power supply (say a battery), and a light bulb ([Figure 4.1](#)).

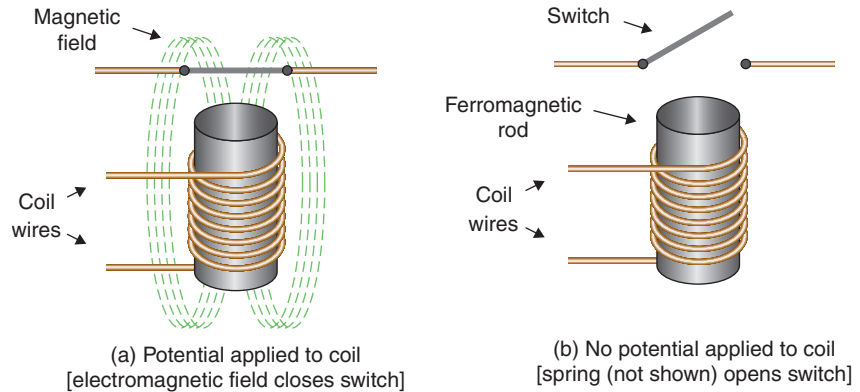
When the switch is CLOSED, the light is ON; when the switch is OPEN, the light is OFF. As we'll see in *Chapter 5: Primitive Logic Functions*, we can actually implement some interesting logical functions by connecting switches together in different ways. If mechanical switches were all we had to play with, however, the life of an electronics engineer would be somewhat boring. The folks in the dim-and-distant past agreed, so they decided that something with a little more “zing” was required ...

### THE ELECTROMECHANICAL RELAY

By the end of the 19th century, when Queen Victoria still held sway over all she surveyed, the most sophisticated form of control for electrical systems was the electromechanical relay ([Figure 4.2](#)). This device consisted of a rod of iron (or some other ferromagnetic



**FIGURE 4.1**  
The simplest control device is a switch.



**FIGURE 4.2**  
The electromechanical relay.

material) wrapped in a coil of wire (the wire would be coated by a layer of insulating material to prevent the coil windings forming electrical connections with each other or with the rod). Applying an electrical potential across the ends of the coil caused the iron rod to act like a magnet. The magnetic field could be used to attract another piece of iron acting as a switch [Figure 4.2(a)]. Removing the potential from the coil caused the iron bar to lose its magnetism, and a small spring would return the switch to its inactive state [Figure 4.2(b)].

The relay is a digital component, because it is either ON or OFF. Although simple in concept, these devices were to prove enormously important with regard to early control systems. This is because the outputs from one or more relays can be used to control other relays, and the outputs from these relays can be used to control yet more relays, and so on and so forth. In fact, by connecting relays together in different ways it's possible to create all sorts of things.

Perhaps the most ambitious use of relays was to build gigantic electromechanical computers, such as the Harvard Mark 1. Constructed between 1939 and 1944, the Harvard Mark 1 was 50 feet long, 8 feet tall, and contained over 750,000 individual components.

The problem with relays (especially the types that were around in the early days) is that they can only switch a limited number of times a second. This severely limits the performance of a relay-based computer. For example, the *Harvard Mark 1* took approximately six seconds to multiply two numbers together, so engineers started to look around for something that could switch faster ...

## THE FIRST VACUUM TUBES

In 1879, the legendary American inventor Thomas Alva Edison (1847–1931) publicly demonstrated his incandescent electric light bulb for the first time.<sup>1</sup> This is the way it worked. A filament was mounted inside a glass bulb. Then all the air was sucked out, leaving a vacuum. When electricity was passed through the filament, it began to glow brightly (the vacuum stopped it from bursting into flames).

A few years later in 1883, one of Edison's assistants discovered that he could detect electrons flowing through the vacuum from the lighted filament to a metal plate mounted inside the bulb. Unfortunately, Edison didn't develop this so-called *Edison Effect* any further. In fact, it wasn't until 1904 that the English physicist Sir John Ambrose Fleming (1849–1945) used this phenomenon to create the first vacuum tube.<sup>2</sup> This device, known as a *diode*, had two terminals and conducted electricity in only one direction (a feat that isn't as easy to achieve as you might think).

In 1906, the American inventor Lee de Forest (1873–1961) introduced a third electrode into his version of a vacuum tube. The resulting *triode* could be used as both an amplifier and a switch. De Forest's triodes revolutionized the broadcasting industry (he presented the first live opera broadcast and the first news report on radio). Furthermore, their ability to act as switches was to have a tremendous impact on digital computing.

One of the most famous early electronic digital computers is the *Electronic Numerical Integrator and Calculator* (ENIAC), which was constructed at the University of Pennsylvania between 1943 and 1946. Occupying 1000 square feet, weighing in at 30 tons, and employing 18,000 vacuum tubes, ENIAC was a monster ... but it was a monster that could perform fourteen multiplications or 5000 additions a second, which was way faster than the relay-based Harvard Mark 1.

However, in addition to requiring enough power to light a small town, ENIAC's vacuum tubes were horrendously unreliable, so researchers started looking for a smaller, faster, and more dependable alternative that didn't demand as much power ...

---

<sup>1</sup>Contrary to popular believe, this wasn't the world's first incandescent bulb. In 1878, the English physicist and electrician, Sir Joseph Wilson Swan (1828–1914) successfully demonstrated a true incandescent bulb, a year earlier than Edison.

<sup>2</sup>*Vacuum tubes* are known as *valves* in England. This is based on the fact that they can be used to control the flow of electricity, similar in concept to the way in which their mechanical namesakes are used to control the flow of fluids.

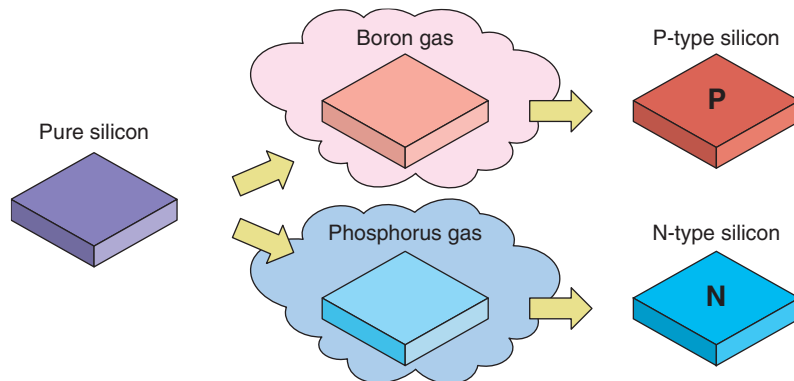
## SEMICONDUCTORS

Most materials are conductors, insulators, or something in-between, but a special class of materials known as *semiconductors* can be persuaded to exhibit both conducting and insulating properties.

The first semiconducting material to undergo serious evaluation was the element germanium (chemical symbol: Ge). However, for a variety of reasons, silicon (chemical symbol: Si) replaced germanium as the semiconductor of choice. As silicon is the main constituent of sand and one of the most common elements on earth (silicon accounts for approximately 28% of the earth's crust), we aren't in any danger of running out of it in the foreseeable future.

Pure crystalline silicon acts as an insulator; however, scientists at Bell Laboratories in the United States found that by inserting certain impurities into the crystal lattice, they could make silicon act as a conductor. The process of inserting the impurities is known as *doping*, and the most commonly used *dopants* are boron atoms (chemical symbol: B) with three electrons in their outermost electron shells, and phosphorus atoms (chemical symbol: P) with five.

If a piece of pure silicon is surrounded by a gas containing boron or phosphorus and heated in a high-temperature oven, the boron or phosphorus atoms will permeate the crystal lattice and displace some silicon atoms without disturbing other atoms in the vicinity. This process is known as *diffusion*. Boron-doped silicon is called *P-type silicon* and phosphorus-doped silicon is called *N-type silicon* (Figure 4.3).<sup>3</sup>



**FIGURE 4.3**  
Creating P-type and  
N-type silicon.

<sup>3</sup>If you ever happen to run across a “full-up” illustration of an integrated circuit, as shown in *Chapter 14: Integrated Circuits (ICs)*, you may see annotations like  $n$ ,  $n+$ ,  $n++$ ,  $p$ ,  $p+$ , and  $p++$ . In this case, the  $n$  and  $p$  stand for standard N-Type and P-type material (which we might compare to an average guy); the  $n+$  and  $p+$  indicate a heavier level of doping (say a bodybuilder or the author flexing his rippling muscles on the beach); and the  $n++$  and  $p++$  indicate a really high level of doping (like a weightlifter on steroids).

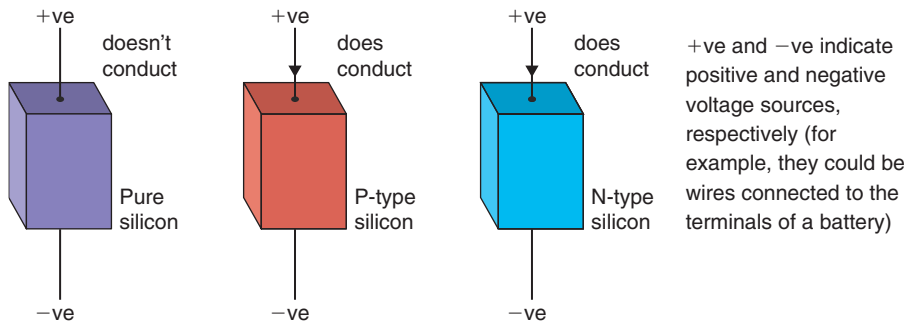
Due to the fact that boron atoms have only three electrons in their outermost electron shells, they can only make bonds with three of the silicon atoms surrounding them. This leaves the fourth silicon atom un-sated and eager to fill its outermost electron shell. Thus, the site (location) occupied by a boron atom in the silicon crystal will accept a free electron with relative ease and is therefore known as an *acceptor* (it's also called a *hole*). So, why do we call this "P-Type silicon?" Well (unofficially), due to the fact that each site of a boron atom is happy to accept an electron, we can visualize this site as being "sort-of" positive. However, the more "official" reason is that we can regard holes as being positive charge carriers (sort of the opposite of electrons).

By comparison, as phosphorous atoms have five electrons in their outermost electron shells, the site of a phosphorous atom in the silicon crystal will donate an electron with relative ease and is therefore known as a *donor*. Due to the fact that it is happy to donate an electron, we can visualize this site as being "sort-of" negative. But the real reason we call this "N-Type silicon" is that the conducting electrons are, of course, negative charge carriers.

## SEMICONDUCTOR DIODES

As was noted above, pure crystalline silicon acts as an insulator. By comparison, both P-type and N-type silicon are reasonably good conductors (Figure 4.4).

The point when things start to become really interesting, however, is when a piece of silicon is doped such that part is P-type and part is N-type. In order to wrap our brains around this, consider what would happen if we were to take a wooden toothpick, dip half of it in a pool of melted wax, remove it from the wax and let it harden, briefly immerse the whole thing in a cup of colored dye, take it out of the dye, dry it off, and scrape away the wax. This leaves one-half of our toothpick in its original state while the other half has been colored.

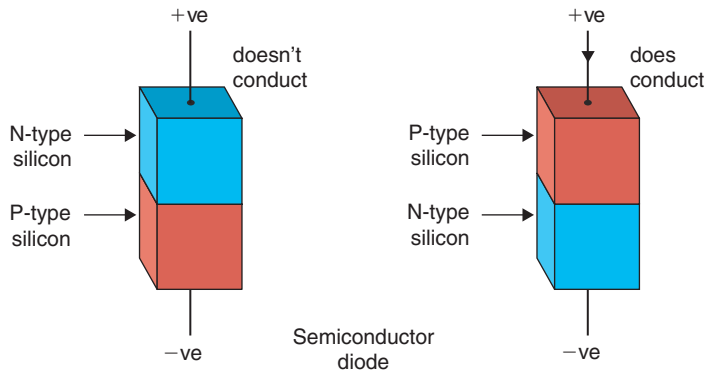


**FIGURE 4.4**

Pure silicon compared to P-type and N-type silicon.

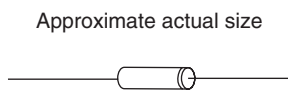
Next, we flip the toothpick upside down, dip the half of the toothpick we just tinted in melted wax, remove it from the wax and let it harden, briefly immerse the whole thing in a cup containing a different colored dye, take it out of the dye, dry it off, and scrape away the wax. Now, one-half of our toothpick will be one color, while the other half is a different color.

Although the actual process is much more complex, we can do something similar with a piece of silicon to make part P-type and part N-type as illustrated in [Figure 4.5](#). We'll return to consider this process in more detail in *Chapter 14: Integrated Circuits (ICs)*. The result is known as a *p-n junction*.



**FIGURE 4.5**

Mixing P-type and N-type silicon.



(a) Diode component



(b) Symbol

**FIGURE 4.6**

Diode: Component and symbol.

As we see, the silicon with both P-type and N-type material conducts electricity in only one direction; in the other direction it behaves like an OPEN (OFF) switch.<sup>4</sup> These structures, known as *semiconductor diodes*,<sup>5</sup> come in many shapes and sizes; an example could be as shown in [Figure 4.6](#).

If the triangular body of the symbol is pointing in the classical direction of current flow (more positive to more negative), the diode will conduct. An individually packaged diode (which would be referred to as a *discrete component*) consists of a piece of silicon with connections to external leads, all encapsulated in a protective package (the silicon is typically smaller than a grain of sand). The package protects the silicon from moisture and other impurities and, when the diode is operating, helps to conduct heat away from the silicon.

<sup>4</sup>If you want to know more about how this works at the nitty-gritty level, including terms like *depletion zones*, then bounce over to *Appendix G: More on Semiconductors*.

<sup>5</sup>The *semiconductor* portion of *semiconductor diode* was initially used to distinguish these components from their vacuum tube-based cousins. As semiconductors took over, however, everyone started to refer to them simply as *diodes*.

Due to the fact that diodes (and transistors, as discussed below) are formed from solids—as opposed to vacuum tubes, which are largely formed from empty space—people started to refer to components formed from semiconductors as *solid-state electronics*.

## BIPOLAR JUNCTION TRANSISTORS (BJTs)

More complex components called *transistors* can be created by forming a sandwich out of three regions of doped silicon. The transistor and subsequently the integrated circuit must certainly qualify as two of the greatest inventions of the 20th century.

Unfortunately, serious research on semiconductors didn't really commence until World War II. At that time, it was recognized that devices formed from semiconductors had potential as amplifiers and switches, and could therefore be used to replace the prevailing technology of vacuum tubes, but that they would be much smaller, lighter, and would require less power. All these factors were of interest to the designers of electronic systems such as radar, which were to play a large role in the war.

Bell Laboratories in the United States began research into semiconductors in 1945, and physicists William Shockley (1910–1989), Walter Brattain (1902–1987), and John Bardeen (1908–1991) succeeded in creating the first point-contact germanium transistor on December 23, 1947. (They took a break for the Christmas holidays before publishing their achievement, which is why some reference books state that the first transistor was created in 1948.)

In 1950, Shockley invented a new device called a *Bipolar Junction Transistor* (BJT),<sup>6,7</sup> which was more reliable, easier and cheaper to build, and gave more consistent results than point-contact devices. BJTs are formed from three pieces of doped silicon, called the *collector*, *base*, and *emitter*. There are two basic types of bipolar transistors, called NPN and PNP,<sup>8</sup> where these names relate to the manner in which the silicon is doped (Figure 4.7).

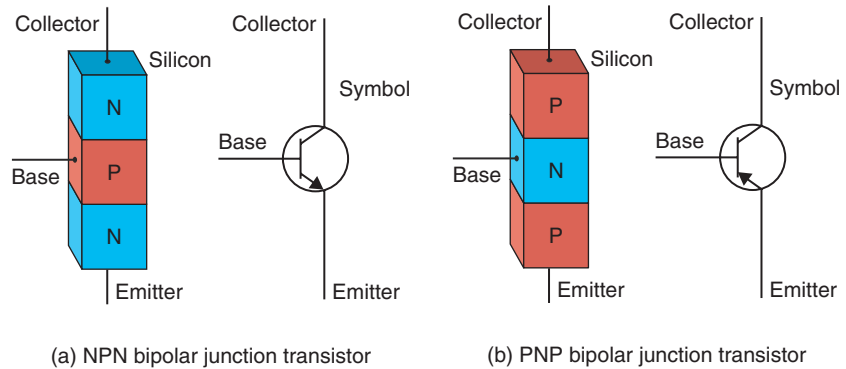
In the analog world, a transistor can be used as a voltage amplifier, a current amplifier, or a switch; in the digital world, a transistor is primarily considered to be a switch. The structure of a transistor between the collector and emitter

---

<sup>6</sup>In conversation, the term *BJT* is spelled out as “B-J-T.”

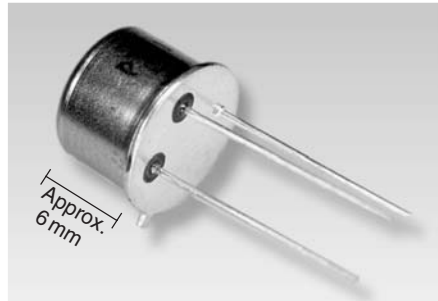
<sup>7</sup>Apropos of nothing at all, the first TV dinner was marketed by the C.A. Swanson company three years later.

<sup>8</sup>In conversation, the terms *NPN* and *PNP* are spelled out as “N-P-N” and “P-N-P,” respectively.



**FIGURE 4.7**  
Bipolar junction  
transistors (BJTs).

**FIGURE 4.8**  
Individually packaged  
transistor. (Photo  
courtesy of Alan  
Winstanley)



terminals is similar to that of two diodes connected back-to-back. Two diodes connected in this way would typically not conduct; however, when signals are applied to the base terminal, the transistor can be turned ON or OFF. If the transistor is turned ON, it acts like a CLOSED switch and allows current to flow between the collector and the emitter; if the transistor is turned OFF, it acts like an OPEN switch and no current flows. We may think of the collector and emitter as *data* terminals, and the base as the *control* terminal.

As for a diode, an individually packaged transistor consists of the silicon, with connections to external leads, all encapsulated in a protective package (the silicon is typically smaller than a grain of sand). The package protects the silicon from moisture and other impurities and helps to conduct heat away from the silicon when the transistor is operating. Transistors may be packaged in plastic or in little metal cans about a quarter of an inch in diameter, with three leads sticking out of the bottom (Figure 4.8).

By the late 1950s, bipolar transistors were being manufactured out of silicon rather than germanium (although germanium had certain electrical advantages, silicon was cheaper and easier to work with). The original bipolar transistors were manufactured using the *mesa process*, in which a doped piece of silicon called the *mesa* (or base) was mounted on top of a larger piece of silicon forming the collector, while the emitter was created from a smaller piece of silicon embedded in the base.



In 1959, the Swiss physicist Jean Hoerni (1924–1997) invented the *planar process*, in which optical lithographic techniques were used to diffuse the base into the collector and then to diffuse the emitter into the base. One of Hoerni's colleagues, Robert Noyce (1927–1990), invented a technique for growing an insulating layer of silicon dioxide over the transistor, leaving small areas over the base and emitter exposed and diffusing thin layers of aluminum into these areas to create wires. The processes developed by Hoerni and Noyce led directly to modern integrated circuits. These techniques are discussed in more detail in *Chapter 14: Integrated Circuits (ICs)*.

## METAL-OXIDE SEMICONDUCTOR FIELD-EFFECT TRANSISTORS (MOSFETS)

In reality, the properties of semiconductors did not start to become well understood until the 1950s. Having said this, as far back as 1925, the Austro-Hungarian scientist Dr. Julius Edgar Lilienfeld (1881–1963) proposed the basic principles behind what we would now recognize as a semiconductor device called a *Metal-Epitaxial Semiconductor Field-Effect Transistor* (MESFET) being used as an amplifier.<sup>9</sup>

In 1926, Dr. Lilienfeld immigrated to America and applied for a patent for this device. On the off-chance you're interested; the title of this little scam (US Patent 1,745,175) was "*Method and apparatus for controlling electric currents.*" Two years later, in 1928 (US Patent 1,900,018), he described what we would now recognize as a depletion-mode MOSFET.<sup>10</sup>

The term MOSFET (or MOS-FET or MOS FET) stands for *Metal-Oxide Semiconductor Field-Effect Transistor* (we'll explain what this mouthful means in a moment).<sup>11,12</sup> In 1960, Dawon Kahng and Martin M. (John) Atalla at Bell Labs fabricated the first successful MOSFET. In 1962, Steven Hofstein and Fredric Heiman at the RCA research laboratory in Princeton, New Jersey, created an experimental integrated circuit comprising 16 MOSFETs.

---

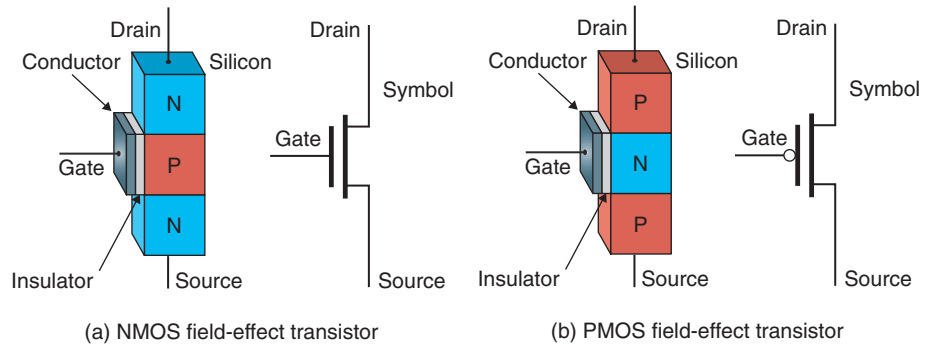
<sup>9</sup>As an aside, during the early 1900s, Lilienfeld did some work with Count Ferdinand von Zeppelin (1838–1917) on designing hydrogen-filled dirigibles.

<sup>10</sup>Nothing is simple. The MOSFETs discussed in this chapter are *enhancement-type* devices, which are OFF unless a control signal is applied to the gate terminal to turn them ON. By comparison, *depletion-type* devices are ON unless a control signal is applied to turn them OFF. And then there are *Junction FETs* (JFETs) and MESFETs. See also *Appendix G: More on Semiconductors*, for more details on all of these little ragamuffins.

<sup>11</sup>In conversation, the term *MOSFET* is pronounced as a single word, where "MOS" rhymes with "boss" and "FET" rhymes with "bet."

<sup>12</sup>These may also be referred to as *Insulated Gate Field-Effect Transistors* (IGFETs).

**FIGURE 4.9**  
Metal-oxide  
semiconductor field-  
effect transistors  
(MOSFETs).



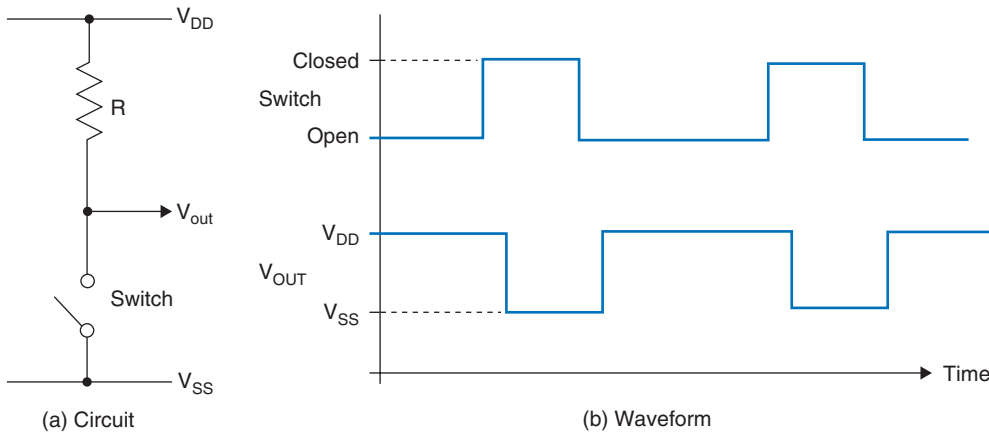
Although the original MOSFETs were somewhat slower than their bipolar transistors, they were cheaper, smaller, and used less power. Also of interest was the fact that modified metal-oxide semiconductor structures could be made to act as capacitors or resistors.

There are two basic types of MOSFETs, called *n-channel* and *p-channel*; once again these names relate to the way in which the silicon is doped (Figure 4.9).

In the case of these devices, the *drain* and *source* form the *data* terminals and the *gate* acts as the *control* terminal. Unlike bipolar devices, the control terminal is connected to a conducting plate, which is insulated from the silicon by a layer of nonconducting oxide. In the original devices the conducting plate was metal—hence, the term *metal-oxide*—but this is now something of a misnomer because modern versions tend to use a layer of *polycrystalline silicon* (*polysilicon*). When a signal is applied to the gate terminal, the plate, insulated by the oxide, creates an electromagnetic field, which turns the transistor ON or OFF—hence, the term *field-effect*.

Now, this is the bit that always confuses the unwary, because the term *channel* refers to the piece of silicon under the gate terminal; that is, the piece linking the drain and source regions. But the channel in the n-channel device is formed from P-type material, while the channel in the p-channel device is formed from N-type material.

At first glance, this would appear to be totally counterintuitive, but there is reason behind the madness. Let's consider the n-channel device. In order to turn this ON, a positive voltage is applied to the gate. This positive voltage attracts any negative electrons in the P-type material and causes them to accumulate beneath the oxide layer where they form a negative channel—hence, the term *n-channel*. In fact, saying “n-channel” and “p-channel” is a bit of a mouthful,



**FIGURE 4.10**  
Resistor-switch circuit.

so instead we typically just refer to these as *NMOS* and *PMOS transistors*, respectively.<sup>13</sup>

This book concentrates on MOSFETs, because their symbols, construction, and operation are relatively easy to understand as compared to their BJT cousins.

## THE TRANSISTOR AS A SWITCH

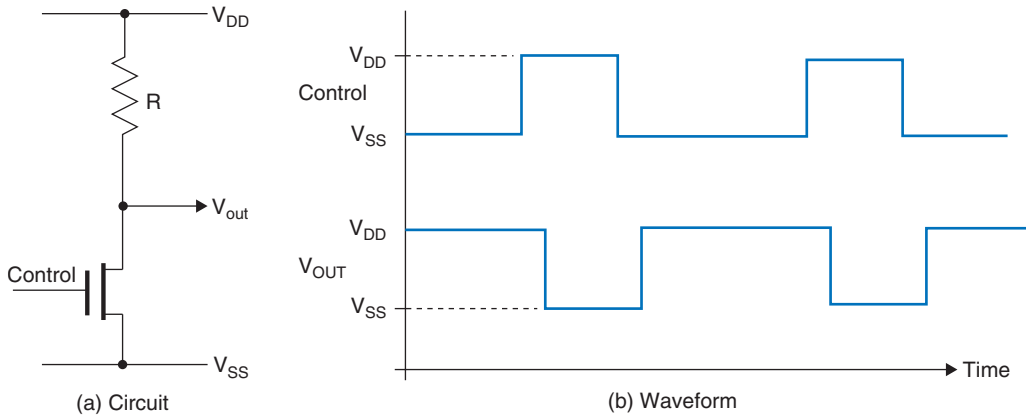
To illustrate the application of a transistor as a switch, first consider a simple circuit comprising a resistor and a real switch (Figure 4.10).

We'll consider the meaning behind the  $V_{DD}$  and  $V_{SS}$  power supply labels in a moment. For our purposes here, let's simply assume that  $V_{DD}$  is more positive than  $V_{SS}$ .

When the switch is OPEN (OFF),  $V_{OUT}$  is connected via the resistor to  $V_{DD}$ ; when the switch is CLOSED (ON),  $V_{OUT}$  is connected via the switch directly to  $V_{SS}$ . In this latter case,  $V_{OUT}$  takes the value  $V_{SS}$  because, like people, electricity takes the path of least resistance, and the resistance to  $V_{SS}$  through the closed switch is far less than the resistance to  $V_{DD}$  through the resistor. Observe that the waveforms in Figure 4.10 show a delay between the switch operating and  $V_{OUT}$  responding. Although this delay is extremely small, it is important to note that there will always be some elements of delay in any physical system.

Now consider the case where the switch is replaced with an NMOS transistor (Figure 4.11). Let's assume that there's a wire connected to the control input of

<sup>13</sup>In conversation, *NMOS* and *PMOS* are pronounced "N-MOS" and "P-MOS," respectively. That is, by spelling out the first letter followed by "MOS" to rhyme with "boss."



**FIGURE 4.11**  
Resistor-NMOS  
transistor circuit.

the transistor (this wire isn't shown here for simplicity), and that the other end of this wire can be switched back and forth between  $V_{DD}$  and  $V_{SS}$ .

When the control input to an NMOS transistor is connected to  $V_{SS}$ , the transistor is turned OFF and acts like an OPEN switch; when the control input is connected to  $V_{DD}$ , the transistor is turned ON and acts like a closed switch. Thus, the transistor functions in a similar manner to the mechanical switch. However, a mechanical switch controlled by hand can only be operated a few times a second, but a transistor's control input can be driven by other transistors, allowing it to be operated hundreds of millions of times a second.

Returning to the labels  $V_{SS}$  and  $V_{DD}$ , these are commonly used in circuits employing MOSFETs. At this point we have little interest in their actual values, and for the purpose of these examples, need only assume that the  $V_{DD}$  supply rail<sup>14</sup> is more positive than the  $V_{SS}$  rail. Why do we use these labels? Well, if you cast your mind back to *Chapter 3: Conductors, Insulators, and Other Stuff*, you will recall that current—in the form of electrons—flows from the more negative source to the more positive target. In the case of the circuit shown in [Figure 4.11](#), the current flows from the  $V_{SS}$  rail (the “source”) into the transistor's source terminal, through the transistor, and “drains away” out of its drain terminal into the  $V_{DD}$  rail.

## GALLIUM ARSENIDE SEMICONDUCTORS

Silicon is known as a *four-valence semiconductor* because it has four electrons available to make bonds in its outermost electron shell. Although silicon is the most

<sup>14</sup>I don't know where the term *rail* comes from in this context, but engineers say “supply rail” or “power supply rail” all the time.

commonly used semiconductor, there is another that requires some mention. The element gallium (chemical symbol: Ga) has three electrons available in its outermost shell, and the element arsenic (chemical symbol: As) has five. A crystalline structure of gallium arsenide (GaAs) is known as a III-V valence semiconductor<sup>15</sup> and can be doped with impurities in a similar manner to silicon.

In a number of respects, GaAs is preferable to silicon, not the least of which is that GaAs transistors can switch approximately eight times faster than their silicon equivalents. However, GaAs is hard to work with, which results in GaAs transistors being more expensive than their silicon cousins.

## LIGHT-EMITTING DIODES (LEDS)

On February 9, 1907, one of Marconi's engineers, Mr. H.J. Round of New York, New York, had a letter published in *Electrical World* magazine as follows:

*To the editors of Electrical World:*

*Sirs: During an investigation of the unsymmetrical passage of current through a contact of carborundum and other substances a curious phenomenon was noted. On applying a potential of 10 volts between two points on a crystal of carborundum, the crystal gave out a yellowish light.*

Mr. Round went on to note that some crystals gave out green, orange, or blue light. This is quite possibly the first documented reference to the effect upon which special components called *light-emitting diodes* (LEDs) are based.<sup>16</sup>

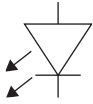
Sad to relate, no one seemed particularly interested in Mr. Round's discovery, and nothing really happened until 1922, when the same phenomenon was observed by O. V. Losov in Leningrad. Losov took out four patents between 1927 and 1942, but he was killed during World War II and the details of his work were never discovered.

In fact, it wasn't until 1951, following the discovery of the bipolar transistor, that researchers really started to investigate this effect in earnest. They found that by creating a semiconductor diode from a compound semiconductor formed from two or more elements—such as gallium arsenide (GaAs) as mentioned

---

<sup>15</sup>In conversation, the Roman Numerals "III-V" are pronounced "three-five."

<sup>16</sup>In conversation, the term LED may be spelled out as "L-E-D" (in which case you would say "an L-E-D") or pronounced as a single word to rhyme with "bed" (in which case you would say "a LED").



**FIGURE 4.12**  
Symbol for a LED.

in the previous topic—light is emitted from the p-n junction, that is, the junction between the P-type and N-type doped materials.

As for a standard diode, an LED conducts electricity in only one direction (and it emits light only when it's conducting). Thus, the symbol for an LED is similar to that for a normal diode, but with two arrows to indicate light being emitted (Figure 4.12).

An LED formed from pure gallium arsenide emits infrared light, which is useful for sensors, but which is invisible to the human eye. It was discovered that adding aluminum to the semiconductor to give aluminum gallium arsenide (AlGaAs) resulted in red light humans could see. Thus, after much experimentation and refinement, the first red LEDs started to hit the streets in the late 1960s.

LEDs are interesting for a number of reasons, not the least of which is that they are extremely reliable, they have a very long life (typically 100,000 hours as compared to 1000 hours for an incandescent light bulb), they generate very pure, saturated colors, and they are extremely energy efficient (LEDs use up to 90% less energy than an equivalent incandescent bulb).

Over time, more materials were discovered that could generate different colors. For example, gallium phosphide gives green light, and aluminum indium gallium phosphite can be used to generate yellow and orange light. For a long time, the only color missing was blue. This was important because blue light has the shortest wavelength of visible light, and engineers realized that if they could build a blue laser diode they could quadruple the amount of data that could be stored on, and read from, a CD-ROM or DVD.

However, although semiconductor companies spent hundreds of millions of dollars desperately trying to create a blue LED, the little rascal remained elusive for more than three decades. In fact, it wasn't until 1996 that the Japanese Electrical Engineer Shuji Nakamura demonstrated a blue LED based on gallium nitride. Quite apart from its data storage applications, this discovery also makes it possible to combine the output from a blue LED with its red and green cousins to generate white light. Many observers believe that this may ultimately relegate the incandescent light bulb to the museum shelf.

## ORGANIC LEDs (OLEDs)

The traditional LEDs discussed in the previous topic are based on semiconductors formed from metalloidal materials such as silicon. When current flows through the LED, positive and negative charges combine and light is emitted.

An *Organic Light-Emitting Diode* (OLED) performs the same trick, but it is based on thin layers of organic molecules (the term *organic* is used in this context because these molecules have a “backbone” formed from carbon atoms, and carbon is the key element for organic life as we know it).

When used to produce displays, OLED technology produces self-luminous displays that do not require backlighting. These properties result in compact displays that require very little power and are much thinner and brighter than their *Liquid Crystal Display* (LCD) counterparts.

## ACTIVE VERSUS PASSIVE AND ELECTRIC VERSUS ELECTRONIC

In this context, the term *active* is used to refer to a component that can use an electrical signal to control the current passing through it; for example, transistors are classed as *active devices*. By comparison, components that are incapable of controlling current by means of another electrical signal are referred to as *passive devices*. On this basis, resistors, capacitors, inductors, and even diodes—all of which simply respond (in a “passive” sort of way) to whatever electrical signals life throws at them—are therefore all classed as passive devices.

Some purists would say that in order for a circuit to be properly called “*electronic*,” it must contain one or more active devices. On this basis, a circuit comprising only resistors, capacitors, and inductors would be considered to be an “*electric circuit*” rather than an “*electronic circuit*.” (Personally, I don’t care what other people think and I would still call it an electronic circuit, but you can make your own decision on this point.)