Diodes

A diode is an electrical device allowing current to move through it in one direction with far greater ease than in the other. The most common type of diode in modern circuit design is the semiconductor diode, although other diode technologies exist. Semiconductor diodes are symbolized in schematic diagrams as such:

```
Semiconductor diode
<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>
```

To depict the permitted direction of electron flow:

```
Semiconductor diode
<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>
```

When placed in a simple battery-lamp circuit, the diode will either allow or prevent current through the lamp, depending on the polarity of the applied voltage:

```
Diode operation
```

When the polarity of the battery is such that electrons are allowed to flow through the diode, the diode is said to be forward-biased. Conversely, when the battery is "backward" and the diode blocks current, the diode is said to be reverse-biased. A diode may be thought of as a kind of switch: "closed" when forward-biased and "open" when reverse-biased.

Oddly enough, the direction of the diode symbol's "arrowhead" points against the direction of electron flow. This is because the diode symbol was invented by engineers, who predominantly use conventional flow notation in their schematics, showing current as a flow of charge from the positive (+) side of the voltage source to the negative (-). This convention holds true for all semiconductor symbols possessing "arrowheads:" the arrow points in the permitted direction of conventional flow, and against the permitted direction of electron flow.

Diode behavior is analogous to the behavior of a hydraulic device called a check valve. A check valve allows fluid flow through it in one direction only. Check valves are essentially pressure-operated devices: they open and allow flow if the pressure across them is of the correct "polarity" to open the gate (in the analogy shown, greater fluid pressure on the right than on the left). If the
pressure is of the opposite "polarity," the pressure difference across the check valve will close and hold the gate so that no flow occurs.

Like check valves, diodes are essentially "pressure-" operated (voltage-operated) devices. The essential difference between forward-bias and reverse-bias is the polarity of the voltage dropped across the diode. Let's take a closer look at the simple battery-diode-lamp circuit shown earlier, this time investigating voltage drops across the various components:
When the diode is forward-biased and conducting current, there is a small voltage dropped across it, leaving most of the battery voltage dropped across the lamp. When the battery's polarity is reversed and the diode becomes reverse-biased, it drops *all* of the battery's voltage and leaves none for the lamp. If we consider the diode to be a sort of self-actuating switch (closed in the forward-bias mode and open in the reverse-bias mode), this behavior makes sense. The most substantial difference here is that the diode drops a lot more voltage when conducting than the average mechanical switch (0.7 volts versus tens of millivolts).

This forward-bias voltage drop exhibited by the diode is due to the action of the depletion region formed by the P-N junction under the influence of an applied voltage. When there is no voltage applied across a semiconductor diode, a thin depletion region exists around the region of the P-N junction, preventing current through it. The depletion region is for the most part devoid of available charge carriers and so acts as an insulator:
If a reverse-biasing voltage is applied across the P-N junction, this depletion region expands, further resisting any current through it:

Conversely, if a forward-biasing voltage is applied across the P-N junction, the depletion region will collapse and become thinner, so that the diode becomes less resistive to current through it. In order for a sustained current to go through the diode, though, the depletion region must be fully collapsed by the applied voltage. This takes a certain minimum voltage to accomplish, called the **forward voltage**.
For silicon diodes, the typical forward voltage is 0.7 volts, nominal. For germanium diodes, the forward voltage is only 0.3 volts. The chemical constituency of the P-N junction comprising the diode accounts for its nominal forward voltage figure, which is why silicon and germanium diodes have such different forward voltages. Forward voltage drop remains approximately equal for a wide range of diode currents, meaning that diode voltage drop not like that of a resistor or even a normal (closed) switch. For most purposes of circuit analysis, it may be assumed that the voltage drop across a conducting diode remains constant at the nominal figure and is not related to the amount of current going through it.

In actuality, things are more complex than this. There is an equation describing the exact current through a diode, given the voltage dropped across the junction, the temperature of the junction, and several physical constants. It is commonly known as the *diode equation:*
The equation \( kT/q \) describes the voltage produced within the P-N junction due to the action of temperature, and is called the thermal voltage, or \( V_t \) of the junction. At room temperature, this is about 26 millivolts. Knowing this, and assuming a "nonideality" coefficient of 1, we may simplify the diode equation and re-write it as such:

\[
I_D = I_S \left( e^{V_D/kT} - 1 \right)
\]

Where,

- \( I_D \) = Diode current in amps
- \( I_S \) = Saturation current in amps (typically 1 x 10^{-12} amps)
- \( e \) = Euler’s constant (~ 2.718281828)
- \( q \) = charge of electron (1.6 x 10^{-19} coulombs)
- \( V_D \) = Voltage applied across diode in volts
- \( N \) = "Nonideality" or "emission" coefficient (typically between 1 and 2)
- \( k \) = Boltzmann’s constant (1.38 x 10^{-23})
- \( T \) = Junction temperature in degrees Kelvin

The equation \( kT/q \) describes the voltage produced within the P-N junction due to the action of temperature, and is called the thermal voltage, or \( V_t \) of the junction. At room temperature, this is about 26 millivolts. Knowing this, and assuming a "nonideality" coefficient of 1, we may simplify the diode equation and re-write it as such:

\[
I_D = I_S \left( e^{V_D/0.026} - 1 \right)
\]

Where,

- \( I_D \) = Diode current in amps
- \( I_S \) = Saturation current in amps (typically 1 x 10^{-12} amps)
- \( e \) = Euler’s constant (~ 2.718281828)
- \( V_D \) = Voltage applied across diode in volts

You need not be familiar with the "diode equation" in order to analyze simple diode circuits. Just understand that the voltage dropped across a current-conducting diode does change with the amount of current going through it, but that this change is fairly small over a wide range of currents. This is why many textbooks simply say the voltage drop across a conducting, semiconductor diode remains constant at 0.7 volts for silicon and 0.3 volts for germanium. However, some circuits intentionally make use of the P-N junction's inherent exponential current/voltage relationship and thus can only be understood in the context of this equation. Also, since temperature is a factor in the diode equation, a forward-biased P-N junction may also be used as a temperature-sensing device, and thus can only be understood if one has a conceptual grasp on this mathematical relationship.
A reverse-biased diode prevents current from going through it, due to the expanded depletion region. In actuality, a very small amount of current can and does go through a reverse-biased diode, called the *leakage current*, but it can be ignored for most purposes. The ability of a diode to withstand reverse-bias voltages is limited, like it is for any insulating substance or device. If the applied reverse-bias voltage becomes too great, the diode will experience a condition known as *breakdown*, which is usually destructive. A diode's maximum reverse-bias voltage rating is known as the *Peak Inverse Voltage*, or *PIV*, and may be obtained from the manufacturer. Like forward voltage, the PIV rating of a diode varies with temperature, except that PIV *increases* with increased temperature and *decreases* as the diode becomes cooler -- exactly opposite that of forward voltage.

![Diagram showing forward and reverse bias of a diode with breakdown point at 0.7V](image)

Typically, the PIV rating of a generic "rectifier" diode is at least 50 volts at room temperature. Diodes with PIV ratings in the many thousands of volts are available for modest prices.

- **REVIEW**:  
  - A *diode* is an electrical component acting as a one-way valve for current.  
  - When voltage is applied across a diode in such a way that the diode allows current, the diode is said to be *forward-biased*.  
  - When voltage is applied across a diode in such a way that the diode prohibits current, the diode is said to be *reverse-biased*.  
  - The voltage dropped across a conducting, forward-biased diode is called the *forward voltage*. Forward voltage for a diode varies only slightly for changes in forward current and temperature, and is fixed principally by the chemical composition of the P-N junction.  
  - Silicon diodes have a forward voltage of approximately 0.7 volts.  
  - Germanium diodes have a forward voltage of approximately 0.3 volts.
• The maximum reverse-bias voltage that a diode can withstand without "breaking down" is called the *Peak Inverse Voltage*, or *PIV* rating.

**TESTING A DIODE:**
* An ohmmeter may be used to qualitatively check diode function. There should be low resistance measured one way and very high resistance measured the other way. When using an ohmmeter for this purpose, be sure you know which test lead is positive and which is negative! The actual polarity may not follow the colors of the leads as you might expect, depending on the particular design of meter.
* Some multimeters provide a "diode check" function that displays the actual forward voltage of the diode when it's conducting current. Such meters typically indicate a slightly lower forward voltage than what is "nominal" for a diode, due to the very small amount of current used during the check.

**Rectifier circuits**

Now we come to the most popular application of the diode: *rectification*. Simply defined, rectification is the conversion of alternating current (AC) to direct current (DC). This almost always involves the use of some device that only allows one-way flow of electrons. As we have seen, this is exactly what a semiconductor diode does. The simplest type of rectifier circuit is the *half-wave* rectifier, so called because it only allows one half of an AC waveform to pass through to the load:

*Half-wave rectifier circuit*

For most power applications, half-wave rectification is insufficient for the task. The harmonic content of the rectifier's output waveform is very large and consequently difficult to filter. Furthermore, AC power source only works to supply power to the load once every half-cycle, meaning that much of its capacity is unused. Half-wave rectification is, however, a very simple way to reduce power to a resistive load. Some two-position lamp dimmer switches apply full AC power to the lamp filament for "full" brightness and then half-wave rectify it for a lesser light output:
In the "Dim" switch position, the incandescent lamp receives approximately one-half the power it would normally receive operating on full-wave AC. Because the half-wave rectified power pulses far more rapidly than the filament has time to heat up and cool down, the lamp does not blink. Instead, its filament merely operates at a lesser temperature than normal, providing less light output. This principle of "pulsing" power rapidly to a slow-responding load device in order to control the electrical power sent to it is very common in the world of industrial electronics. Since the controlling device (the diode, in this case) is either fully conducting or fully nonconducting at any given time, it dissipates little heat energy while controlling load power, making this method of power control very energy-efficient. This circuit is perhaps the crudest possible method of pulsing power to a load, but it suffices as a proof-of-concept application. If we need to rectify AC power so as to obtain the full use of both half-cycles of the sine wave, a different rectifier circuit configuration must be used. Such a circuit is called a full-wave rectifier. One type of full-wave rectifier, called the center-tap design, uses a transformer with a center-tapped secondary winding and two diodes, like this:

This circuit's operation is easily understood one half-cycle at a time. Consider the first half-cycle, when the source voltage polarity is positive (+) on top and negative (-) on bottom. At this time, only the top diode is conducting; the bottom diode is blocking current, and the load "sees" the first half of the sine wave, positive on top and negative on bottom. Only the top half of the transformer's secondary winding carries current during this half-cycle:
During the next half-cycle, the AC polarity reverses. Now, the other diode and the other half of the transformer's secondary winding carry current while the portions of the circuit formerly carrying current during the last half-cycle sit idle. The load still "sees" half of a sine wave, of the same polarity as before: positive on top and negative on bottom:

One disadvantage of this full-wave rectifier design is the necessity of a transformer with a center-tapped secondary winding. If the circuit in question is one of high power, the size and expense of a suitable transformer is significant. Consequently, the center-tap rectifier design is seen only in low-power applications. Another, more popular full-wave rectifier design exists, and it is built around a four-diode bridge configuration. For obvious reasons, this design is called a full-wave bridge:
Current directions in the full-wave bridge rectifier circuit are as follows for each half-cycle of the AC waveform:
Remembering the proper layout of diodes in a full-wave bridge rectifier circuit can often be frustrating to the new student of electronics. I've found that an alternative representation of this circuit is easier both to remember and to comprehend. It's the exact same circuit, except all diodes are drawn in a horizontal attitude, all "pointing" the same direction:

*Full-wave bridge rectifier circuit (alternative layout)*

One advantage of remembering this layout for a bridge rectifier circuit is that it expands easily into a polyphase version:

*Three-phase, full-wave bridge rectifier circuit*

Each three-phase line connects between a pair of diodes: one to route power to the positive (+) side of the load, and the other to route power to the negative (-) side of the load. Polyphase systems with more than three phases are easily accommodated into a bridge rectifier scheme.

When polyphase AC is rectified, the phase-shifted pulses overlap each other to produce a DC output that is much "smoother" (has less AC content) than that produced by the rectification of single-phase AC. This is a decided advantage in high-power rectifier circuits, where the sheer physical size of filtering components would be prohibitive but low-noise DC power must be obtained. The following diagram shows the full-wave rectification of three-phase AC:
In any case of rectification -- single-phase or polyphase -- the amount of AC voltage mixed with the rectifier's DC output is called *ripple voltage*. In most cases, since "pure" DC is the desired goal, ripple voltage is undesirable. If the power levels are not too great, filtering networks may be employed to reduce the amount of ripple in the output voltage.

Sometimes, the method of rectification is referred to by counting the number of DC "pulses" output for every 360° of electrical "rotation." A single-phase, half-wave rectifier circuit, then, would be called a *1-pulse* rectifier, because it produces a single pulse during the time of one complete cycle (360°) of the AC waveform. A single-phase, full-wave rectifier (regardless of design, center-tap or bridge) would be called a *2-pulse* rectifier, because it outputs two pulses of DC during one AC cycle's worth of time. A three-phase full-wave rectifier would be called a *6-pulse* unit.

Modern electrical engineering convention further describes the function of a rectifier circuit by using a three-field notation of *phases*, *ways*, and number of *pulses*. A single-phase, half-wave rectifier circuit is given the somewhat cryptic designation of 1Ph1W1P (1 phase, 1 way, 1 pulse), meaning that the AC supply voltage is single-phase, that current on each phase of the AC supply lines moves in one direction (way) only, and that there is a single pulse of DC produced for every 360° of electrical rotation. A single-phase, full-wave, center-tap rectifier circuit would be designated as 1Ph1W2P in this notational system: 1 phase, 1 way or direction of current in each winding half, and 2 pulses or output voltage per cycle. A single-phase, full-wave, bridge rectifier would be designated as 1Ph2W2P: the same as for the center-tap design, except current can go *both* ways through the AC lines instead of just one way. The three-phase bridge rectifier circuit shown earlier would be called a 3Ph2W6P rectifier.

Is it possible to obtain more pulses than twice the number of phases in a rectifier circuit? The answer to this question is yes: especially in polyphase circuits. Through the creative use of transformers, sets of full-wave rectifiers may be paralleled in such a way that more than six pulses of DC are produced for three phases of AC. A 30° phase shift is introduced from primary to secondary of a three-phase transformer when the winding configurations are not of the same type. In other words, a transformer connected either Y-∆ or Δ-Y will exhibit this 30° phase shift, while a transformer connected Y-Y or Δ-Δ will not. This phenomenon may be exploited by having one transformer connected Y-Y feed a bridge rectifier, and have another transformer connected Y-∆ feed a second bridge rectifier, then parallel the DC outputs of both rectifiers. Since the ripple voltage waveforms of the two rectifiers' outputs are phase-shifted 30° from one
another, their superposition results in less ripple than either rectifier output considered separately: 12 pulses per 360° instead of just six:

3Ph2W12P rectifier circuit

- **REVIEW:**
  - *Rectification* is the conversion of alternating current (AC) to direct current (DC).
  - A half-wave rectifier is a circuit that allows only one half-cycle of the AC voltage waveform to be applied to the load, resulting in one non-alternating polarity across it. The resulting DC delivered to the load "pulsates" significantly.
  - A full-wave rectifier is a circuit that converts both half-cycles of the AC voltage waveform to an unbroken series of voltage pulses of the same polarity. The resulting DC delivered to the load doesn't "pulsate" as much.
  - Polyphase alternating current, when rectified, gives a much "smoother" DC waveform (less ripple voltage) than rectified single-phase AC.