The ‘idealized’ diode, which is introduced here, is an abstract approximation to a property of a ‘real’ diode, in the same sense that an idealized resistor is an abstract approximation to a property of a ‘real’ resistor. The idealized lumped-circuit diode is defined as shown to the right. Note particularly the polarity assumptions for the voltage and the current in terms of which the diode volt-ampere property is defined. The diode is in a limited sense analogous to a perfect mechanical knife switch, i.e., current flows without dissipating energy when the switch is closed (the ‘forward bias’ or ON state of the diode) and no current flows when the switch is open (the reverse bias or OFF state of the diode). A mechanical switch requires some sort of additional mechanism actually to open or close the switch; the idealized diode however switches state automatically on the basis of the voltage or current level. (To avoid an occasional conceptual difficulty note carefully that the diode characteristic is not discontinuous at the origin; it is the slope of the diode characteristic that is discontinuous. If it were otherwise there would be difficulties with Conservation of Energy (KVL) and Conservation of Charge (KCL) principles.

The analysis of circuits including idealized diode elements is, with one qualification, no different than the analysis of linear circuits in general. All the means of extracting information about circuit performance remain applicable. It remains generally true that a solution exists, that the solution is unique, and that to be valid it is necessary and sufficient that a proposed solution satisfy KVL, KCL, and the circuit element constitutive relations. However except in the simplest of circumstances it is generally not readily obvious whether a diode is forward or reversed-biased; this uncertainty often persists until an analysis is substantially completed. However the situation can be resolved in a straightforward manner. The first step simply is recognizing that a diode either is forward-biased or it is reversed-biased, one state or the other, and that there is just one solution to the analysis. While the procedure can be organized efficiently for the present simply assume (absent any insight make an arbitrary choice) one or the other state for the diode. Analyze the circuit based on the assumed diode state. Note that once a diode state is chosen the diode effectively becomes either an open circuit (reverse bias) or a short circuit (forward bias). There is no special problem in the analysis of the circuit. Some analysis procedures may be more convenient than others but all lead to the same unique solution.

Assuming no procedural errors are made (KVL, KCL, etc. are applied properly) the process of solution ordinarily inherently will provide for satisfaction of Kirchoff’s Laws and the element constitutive...
relations, and therefore verified that it is the unique solution. However a question as to the validity of
the solution arises in association with the assumption made about a diode state, i.e., is the diode
constitutive relation satisfied. It is necessary to check for consistency of the assumed state of each diode
(where there can be more than one) with the diode volt-ampere relation. Does the calculated voltage across a
diode assumed to be reverse-biased have the proper polarity? Does the current through a diode assumed
to be forward-biased have the proper polarity? If so the uniqueness of the solution assures that the
assumption made was the correct one. If not the assumption made simply is incorrect. But then, since a
diode has but two states, the correct assumption is obvious, and the analysis can be performed again
with assurance that a valid solution can be obtained.

**Idealized Diode Circuit Analysis Illustration**

Calculate V and I in circuit (a) drawn below, using the idealized diode model. The diode is either ON or
OFF. For a given assumption about the diode state, the diode is replaced either by an open-circuit or by
a short-circuit, and the circuit analyzed for the assumed conditions. Note that once the diode state is
assumed the circuit is effectively linear. Verify whether or not the assumed diode state is consistent
with the calculated currents and voltages. The voltage across an OFF diode must be a reverse-bias
voltage. The current through an ON diode must flow in the 'easy' current flow direction. If this is so the
conditions for a solution exist, and since the solution is unique it is the solution. On the other hand
inconsistency means the assumption made about the diode state is incorrect; since there are only two
possibilities for the diode state the correct assumption is clear.

Suppose the diode is reverse-biased (as in (b)); then I = 0. The voltage at (y) is 2V. The voltage at (x)
is 4\( \frac{2}{5} \) = 1.6V. Then \( V_0 = 0.4 \) V; this is inconsistent with the assumed reverse-biased diode state. The
diode, therefore, must be forward-biased. It is a straightforward calculation using the revised
assumption about the diode state, as shown in (c), to determine the diode current is 2/11 ma.

**Idealized Diode Example: Two-diode circuit**

Things can get a bit involved with multiple diodes in a circuit. It is then
necessary, in a worst case, to examine \( 2^N \) possible combinations of diode
states. For example determine I in the circuit shown; assume idealized diodes.

Since there are two diodes there are four possible combinations of diode states.
We check them all for consistency. The four possibilities are illustrated
below; an OFF diode is effectively an open-circuit, and an ON diode is effectively a short-circuit. (A
small triangle is used as a reminder of diode orientation.)

Consider circuit (a), both diodes assumed OFF. But this condition can not occur because the circuit
currents and voltages are inconsistent with the assumed diode states. With the diodes OFF there is no
current flow, and so no voltage drop across the resistors. But then while the voltage across D1 is
consistent with reverse bias the voltage across D2 is not. Both diodes cannot be OFF concurrently.
For condition (b) I = [10 -(-15)] / (10k + 5k) = 5/3 ma. The direction of current flow is consistent with the assumed forward biased state of D2. The voltage (to ground) across D1 is 10 -(5/3)(10) = -20/3V and this is not consistent with the assumed reverse biased state of D1.

For condition (c) the direction of the current through D1 is inconsistent with its assumed forward biased state, and the voltage across D2 is inconsistent with its assumed reverse biased state.

The analysis of circuit (d) requires a bit of care because it is easy enough to jump to a hasty but false intuitive conclusion instead of making an explicit calculation using KVL and KCL. Thus I is 1 ma in this circuit. But the current in the 5 kΩ (and D2) is 3 ma, and flows consistent with forward bias. Hence the current in D1 would be 2 ma flowing in the correct direction for consistency with the assumed diode state. Circuit (d) is the diode state configuration consistent with KVL, KCL, and all the circuit element volt-ampere relations.

**Semiconductor Junction Diode**

A representative sketch of a PN junction diode characteristic is drawn to the right. The icon used to represent the diode is drawn in the upper left corner of the figure, together with the polarity markings used in describing the characteristics. The icon 'arrow' itself suggests an intrinsic polarity reflecting the inherent non-linearity of the diode characteristic. Although the icons are the same it is unlikely that there will be confusion about whether an idealized diode or semiconductor diode is referenced.

Note: The designations ‘P’ and ‘N’ have to do with the preparation of the semiconductor material. The ‘P’ material corresponds to the upper part of the icon, and the ‘N’ material corresponds to the base. The direction for ‘easy’ conduction is from P to N.

The diode characteristic has been separated (roughly) into three ranges of operation for purposes of description. The 'breakdown' range on the left side of the figure will be discussed separately later. For reasons considered later specially designed (and more costly) diodes are used for operation in this range. Ordinarily diodes operate only in the forward- and reverse-bias ranges. Forward bias is a range of 'easy' conduction, i.e., after a small threshold voltage level (≈ 0.6–0.7 volts for silicon devices) is reached a subsequent small voltage change produces an exponentially larger current change. The theoretical terminal relationship is
where $I_0$ is a function of material properties and temperature, $q_e$ is the electron charge, $k$ is Boltzmann's constant, and $T$ is the Kelvin temperature. $N$ is an empirical parameter included to account for variations of the manufacturing process; $N=1$ for diodes which are part of an integrated circuit, and $N=2$ for discrete diodes. The voltage $V$ and the current $I$ are positive for forward-bias (refer to the diode icon for polarity conventions used). For a representative signal diode the coefficient $I_0$ has a value of about $10^{-15}$ ampere. At room temperature (300K) the exponent is $(40/N)V$. When the diode current gets to be about a milliampere ($V \approx 0.7$ volts) the exponential term very much dominates and the '1' can be ignored. In forward bias a change in $V$ of roughly 50 millivolt produces an order of magnitude change in current. Note again that the diode icon 'points' in the direction of easy current flow.

On the other hand if the polarity of the voltage is reversed the exponent in the theoretical expression is negative, and the exponential rapidly becomes small compared to 1; the current flows in the reverse direction and the diode operates in reverse bias. The theoretical reverse bias current is very small. In fact it is so small that secondary phenomena unrelated to the junction produce much larger (compared to $10^{-15}$ ampere) currents, typically of the order of a microampere for a signal diode. Note that if contamination, e.g., skin oils, between the diode terminals provides a one megohm conduction path a current of 1 microampere flows for a reverse bias voltage of 1 volt.

It is not really the details of the non-linearity of the diode characteristic that we investigate here, although the exponential character of forward bias operation does have important applications. Rather it is the non-bilateral character of the characteristic that is emphasized, that is the recognition that the terminal volt-ampere relation is quite different depending on the polarity of the voltage applied across the diode. As noted before the diode behaves in some respects as does an electrical switch, open for one voltage polarity and closed for the other.

### Diode Comparison

Consider first a formal analysis of the circuit drawn to the left. The basic analysis procedure remains just as discussed earlier. If, for example, current variables are used apply KCL to define currents, then KVL to write loop voltage equations. Or, if voltage variables are used apply KVL to define voltages, then KCL to write node current equations. In either case use the circuit element volt-ampere relations to relate currents and voltages.

Suppose we use a mesh analysis. Define the loop current $I$ (KCL indicates all elements carry the same current). Write a loop equation (KVL) $E = IR + V$. Now apply the diode volt-ampere relation to relate $V$ to $I$, obtaining a single equation in one unknown. The equation is of course nonlinear (transcendental), but it can be solved for $I$ by an iterative trial-and-error numerical process.

$$E = IR + \left( \frac{1}{I} \right) \ln \left( \frac{1}{I_0} + 1 \right)$$

A graphical equivalent of the solution process could be used. Thus plot $V = E - IR$ on the same coordinate plane as the diode characteristic. The former is a line with intercepts as shown in the figure to the right. The circuit currents and voltages must be consistent both with the diode volt-ampere relations and with the 'load line'. The solution is the intersection of the two curves, i.e., $Q$. 

---

Diodes
Consider this illustrative calculation: given circuit parameters $E = 10\,\text{v}$, $I_0 = 10^{-15}\,\text{ampere}$ and $\eta = 40\,\text{v}^{-1}$ what value of $R$ will provide a current of about 2 mA? Using either method, even for this simple circuit, the analysis is not very convenient.

For many engineering purposes a good estimate (not the same as a guess) will provide useful information, particularly if it is available from a straightforward easy calculation. For this purpose the 'real' diode characteristics often are approximated by 'idealized' characteristics. This approximation is superimposed on the graphical analysis described above; see the figure to the right. Clearly the intersection of the load line with the idealized diode characteristic is not $Q$. But how significant is the difference? The load line is $E = IR + V$, and the degree of approximation in setting $V = 0$ for forward-bias depends on the magnitude of $E$. The actual value of $V$ will be of the order of 0.7 volts or less (remember exponential changes of current with voltage). If $E$ is, say 10 v, neglecting 0.7 volts is an error of the order of $0.7/10 = 7\%$.

For our purpose, i.e., obtaining good if not precise values for circuit voltages and currents sufficient to understand circuit operation and performance, the idealized diode model generally will suffice. Of course a more precise diode approximation might be used, for example a model with the characteristic illustrated by the dashed line in the figure. A diode model with such a characteristic is considered next.

**An Extended diode model**

Consider the composite model formed by the idealized diode in series with the voltage source as drawn to the right two states the necessary revised choice is clear. Assume (arbitrarily) that the diode is reverse-biased, i.e., it is effectively an open-circuit. Then clearly $I = 0$, and the voltage across the diode is $V - E$. If the diode is to be reverse-biased we require $V - E \leq 0$, or $V \leq E$. The solution will be valid provided the condition $V \leq E$ is met.

What happens when the condition $V \leq E$ is not met? The diode cannot then be reverse-biased, and so it must be forward-biased. In this case the diode is effectively a short-circuit and the current is whatever is required by the part of the circuit not shown. The terminal relationship for the series combination of idealized diode and voltage source then is $I = 0$ for $V \leq E$, and $V = E$ for $I \geq 0$. The combination of elements provides a somewhat closer approximation to a ‘real’ diode characteristic.

In the generally unlikely event that a still further improved diode model is desired consider the composite circuit to the left. Deriving the volt-ampere characteristic shown is left as an exercise.
More Diode Examples

There is a special objective to the following illustrations. The use of simplified device models is intended to provide approximate but nevertheless useful information without involvement of difficult calculations. Inexpensive computer circuit analysis programs executing on desktop machines are available for even involved nonlinear circuits. However such programs require specification of a circuit to analyze. If there is no circuit, i.e., the circuit is yet to be designed even a powerful analysis program does not help. Approximate calculations using simplified models can provide significant guidance in the initial design of a circuit, and that initial prototype then can be refined by computer computation. Keep this in mind in evaluating the following illustrations.

Single diode circuit

Determine V and I in the circuit shown; assume an idealized diode. However there is a special point to this calculation. The use of simplified device models is intended to provide approximate but useful information in a

The diode is either ON (forward-biased) or OFF (reverse-biased). Assume, subject to subsequent confirmation, that the diode is OFF. Then there is no current through the diode and so no current through the resistor. Hence the voltage drop across the diode (P->N) is calculated to be 3–(−7) = +10v. This is not a reverse-bias voltage, and so the assumption that the diode is open-circuit is contradicted, and the analysis is invalid.

But if the diode cannot be OFF then it must be ON. Verify this by reanalyzing the circuit. The ON diode is effectively a short-circuit, and so the voltage drop across the resistor is 3–(−7) = 10v, and so I is 0.5 ma. The current flow is properly directed through the forward-biased diode. Hence the assumption of an ON diode is consistent with KCL, KVL, and the volt-ampere relations for all the circuit elements, and the unique solution for the circuit voltages and currents has been obtained. Incidentally a PSpice analysis using a 1N4002 diode computes a current of 4.73 ma. If the composite model described above is used, with a 0.7 volt ‘knee’ voltage the estimated current is 4.7 ma. As will be emphasized in the discussion the point here is not a matter of accuracy but rather the usefulness of the information obtained with a simplified calculation.

Diode Clipping Illustration

A computer analysis of the diode circuit drawn to the right serves to further illustrate diode behavior. The (Thevenin equivalent) source presents a 10-volt peak triangular waveform across the series combination of a diode and DC offset voltage source; refer to the computer-generated plot below. In place of the idealized diode model the computer analysis uses the 1N4004 nonlinear PSpice diode model; this diode is a general-purpose low-power rectifier diode.

A formal analysis of this single idealized-diode circuit follows the earlier discussion; the difference here is that the source voltage is not fixed. It also is not difficult to recognize (for an idealized diode) rather directly that V(2) will be equal to the source voltage when that voltage is less than the offset voltage, since the diode will be OFF under these circumstances. On the other hand V(2) will be equal to the offset voltage otherwise.

The computational results of the computer analysis are plotted below. Note the ‘clipping’ of the input waveform at the different thresholds. In interpreting this plot recall that the idealized diode model has a
zero threshold (‘knee’) voltage while the 1N4004 threshold is 0.5 - 0.7 volts or so. Account for this in interpreting the plot.

In addition to the technical use of the diode to limit the output voltage level the use of the idealized diode model to estimate actual circuit performance should be noted.

Diode Voltage Regulator
Diodes occasionally are used as a convenient albeit crude kind of voltage reference in integrated circuits. Although there are other generally preferable means to serve the purpose than used in this illustration nevertheless it is instructive to examine the principles involved in this rather straightforward if crude circuit..

A Thevenin equivalent for a power supply consists of a voltage source in series with the ‘internal’ resistance of the supply. There is a voltage drop across this resistor that increases with increasing load current, so that the actual terminal supply voltage decreases. To limit this terminal voltage change in so far as the load is concerned a ‘regulator’ is added; in this illustration the regulator is the diode ‘tree’. The idea is to use the diode forward-bias property wherein large diode current changes involve exponentially smaller diode voltage changes.

The circuit is designed to draw supply current greater than the maximum required by the load (represented here for convenience as a current source).. This source current divides, part drawn off by the load and the remainder shunted through the diodes. As the load current demand changes the current division ratio changes. However provided the minimum current through the diodes (which occurs at maximum load current) is sufficient for diode operation above the threshold (‘knee’) the voltage across the load will not vary greatly. (And of course the maximum diode current, which occurs for minimum load current, should not exceed the diode ratings.)

As an illustration the load voltage for the circuit is computed as the load current varies from 0 to 20 ma; VS = 10v, RS = 200Ω, RB = 150Ω. Note that, absent the regulator, the terminal voltage will have dropped by 7 volts from its no-load value for the maximum allowed 20 ma. Nonlinear 1N4004 diode
models are used for the computer analysis. Clearly an idealized diode approximation is not useful here. However we can use the extended model described before with a knee voltage of about 0.7 volts. In this basis estimate the regulated voltage as about 4 x 0.7 = 2.8 volts. Computed data is plotted below.

Note the load voltage varies only about 0.5 volt or so over a full 20ma change in load current; compare this to the 7 volts or so change absent the diode regulation. As a parenthetical observation the source current is compared to the load current in the lower plot. The source current is estimated as 
\[(10 - 4x0.7)/(200 + 150) = 20.6 \text{ ma}, \] and stays sensibly constant as the load current increases. The diode current, on the other hand, decreases to provided the increasing load current.

\[ \text{Square Root Transfer Function Synthesis} \]

An idealized diode operates in one or the other of two states; either effectively an open circuit (reverse bias) or effectively a short circuit (forward bias). A change of state of a diode corresponds to a change in the topology of the circuit in which the diode is embedded. On entering a reverse bias state the diode removes (electrically) a branch with which it is in series, or conversely electrically adds a series branch when it enters the forward bias state.

For example consider the illustrative circuit drawn below; assume \( E_1 < E_2 \). For \( V_{in} < E_1 \) neither diode is forward biased, and so the diode branches are electrically inactive. Hence \( V_{out} = V_{in} \). When \( E_1 \leq V_{in} < E_2 \) the diode \( D_1 \) becomes forward biased, activating the \( R_2 \) branch. The transfer relationship remains linear of course, but the slope is smaller. The new relationship continues until \( V_{out} \) reaches \( E_2 \), where \( D_2 \) becomes forward biased and effectively adds a shunt branch to the circuit.
Here is a more interesting application of this ‘piecewise linear’ circuit synthesis. The goal is a piecewise linear approximation to a circuit whose output voltage is the square root of the input voltage for \(0 \leq V_{in} < 25\) volts. The approximation is formed by segments connecting points where the input voltage is 0, 1, 4, 9, 16, and 25 volts (a more or less arbitrary choice here to simplify the endpoint coordinates).

The following idealized diode circuit synthesizes this transfer characteristic:

Note that for specific diode states the transfer response is linear, i.e., consists of a series of linear segments. As the input voltage increases a change in slope of a segment occurs only when a diode becomes forward biased, i.e. at output voltages of 1, 2, 3, 4, and 5 volts. Verify that the input voltage for each of these input voltages is the square of the voltage; the several segments then necessarily are determined by joining these coordinate points.

The idealized diode circuit provides a guide for realizing a circuit using real diodes. For example the circuit drawn below is derived from the idealized diode circuit. The resistors \(R\) and \(R/2\) are represented...
by the 6.6kΩ and 3.3kΩ resistors respectively; note the use of standard values rather than a strict proportion.

The idealized diodes are replaced by 1N4002 diodes, and in place of separate bias voltage sources a resistive voltage divider is used. The ‘experimental’, i.e. computed transfer characteristic matches the theoretical curve perfectly for $V_{in} = 0$. The resistor $R_{12}$ is adjusted ‘experimentally’ to provide a (close) match at $V_{in} = 25V$. The computed transfer characteristic is evaluated by plotting $V_{2out}$ vs $V_{in}$; this should be a line with unit slope. As a reference $V_{in}$ vs. $V_{in}$ is plotted also.
Half-Wave Rectifier
Assume, for the moment, that the capacitor in the diode circuit diagram to the right is absent. One cycle of a sinusoidal source voltage is sketched to the left of the circuit; the positive half-cycle is filled simply to distinguish it clearly from the negative half-cycle. The diode conducts only in forward-bias, i.e., only during the positive half cycle. Current flowing through the diode produces a voltage across the load resistance during this half-cycle. Since there is negligible conduction during the other half-cycle the load voltage during this period is zero. The waveform sketched on the right illustrates this rectification.

By converting the bipolar waveform to a unipolar one a DC waveform is ‘approached’, in the sense that unlike the sinusoidal waveform the average value of the voltage over a cycle is greater than zero. It is however not a constant DC waveform, which is what is needed to avoid interactions between supply voltage timing and circuitry being energized by the supply.

One approach towards reducing the variation involves separating the processes of providing and distributing the energy. Thus the source is not used to supply the load energy directly but rather to deliver the energy to an energy reservoir during the conducting half-cycle; the reservoir then is designed to release the stored energy more or less uniformly over the full cycle. This is the purpose of the capacitor that has been ignored until now. During the positive half-cycle the power source is used primarily to charge the capacitor. During the negative half-cycle, when the diode blocks current flow from the source, the capacitor discharges, acting as a temporary energy supply for the load resistor. The capacitor functions (in a loose manner of speaking) as a rechargeable battery, continually charged during one half-cycle and discharged during the other. That is the essential concept; the details are a bit more involved.

Suppose that the power supply has been operating long enough to reach a steady-state, i.e., initial turn-on transients have become negligible. The steady-state condition (as will be verified) for one full cycle of the source frequency is as shown in the figure below. For convenience consider the circuit behavior starting from the time the diode begins conduction at $t_1$. The AC source resistance is generally small (to avoid wasting power) so that the capacitor charges rapidly enough to track the source waveform. Once past the peak value the source voltage amplitude decreases sinusoidally. However the capacitor charge cannot flow back through the diode and so the capacitor voltage must decay exponentially through the load resistor. It is not difficult to show that the sinusoid decreases faster than the exponential, so that at some time $t_2$ the diode becomes reverse-biased! At that point the source no longer supplies energy to the load; the load current is provided entirely by the capacitor discharge. Eventually of course the
source begins the next positive half-cycle, and the diode is again forward-biased. This occurs one full cycle after $t_1$, and the waveform repeats. Refer to the text and notes for additional details.

**Full-Wave Bridge Rectifier**

In the half-wave rectifier circuit source power is provided from the source only during a small conduction angle near the peak of the positive half-cycle. Observe however that by reversing the diode orientation power would be supplied during the negative half-cycle rather than the positive half-cycle. This suggests powering the load using two half-wave rectifiers, one conducting during the positive half-cycle and the other during the negative half-cycle. It is necessary, of course, to assure that both rectifiers supply current that flows through the load in the same direction. A diode ‘bridge’ circuit to accomplish the desired end is drawn to the right.

As before temporarily ignore the capacitor. During the positive half-cycle, i.e., $V(1,3) \geq 0$, diodes $D_1$ and $D_2$ are forward-biased, and load current flows through $R_L$ from node 2 to node 0. The other two diodes are reverse-biased, and so not active in the circuit. Effectively the circuit during the positive half-cycle is as shown in the diagram below, left. Except for using two diodes rather than one this is the half-wave rectifier configuration.

For the negative half-cycle the circuit is effectively as shown below, right. Note that except for the

source the circuit configuration although geometrically flipped vertically is electrically the same as before. The special consequence of this topological reversal is to cause the rectified current to flow though the load resistor in the same direction for both half-cycles, so that the source is connected during both half-cycles to provide unipolar load power.

Circuit operation then is as described before, except that the capacitor is recharged on both half-cycles. Hence the exponential decay for a given filter capacitor is essentially halved, improving the peak-to-peak ripple for a given capacitor value. The reason for using two diodes for each half-wave circuit should be clear at this point; it is necessary to prevent short-circuiting. A secondary consequence of this is that the load voltage amplitude is lowered from the source voltage by two diode drops.

It is interesting to note again that (for a small-ripple design) the source conducts only for a short fraction of each half-cycle. During the small conduction angle the capacitor is charged, and it is the discharge of the capacitor during the remainder of each half-cycle that provides the load current.

Circuits  Diodes  12  M H Miller
A bridge circuit characteristic of importance is that the load resistor and the source cannot have a common ground point; usually one end of the load is grounded as shown in the circuit diagram. Incidentally since a larger RC product promotes smaller ripple it follows that low-current (large load R) supplies have an advantage.

**Zener Diode Breakdown Characteristics**

Ordinarily the reverse-bias blocking action of a PN junction allows only a small 'leakage' current to flow. However if a sufficiently large reverse-bias is applied other junction phenomena develop which dominate the leakage current, allowing in effect much larger reverse-bias currents. This is the 'breakdown' part of the diode characteristic; 'breakdown' here refers to the overshadowing of the semiconductor junction behavior by other phenomena rather than to a destructive effect. While all diodes display this reverse-bias breakdown phenomenon Zener diodes are manufactured specifically for operation in the breakdown condition with guaranteed specifications. The breakdown parameters of these Zener (or voltage reference) diodes receive special processing attention during their manufacture.

Two distinct phenomena, acting individually or concurrently depending on diode details, are involved in the breakdown phenomena. One mechanism is associated with the acceleration of carriers across the very strong junction field. Kinetic energy gained by an accelerated carrier, if sufficiently great, can cause additional impurity atom ionization during a collision with the atom. Each additional carrier is then also accelerated and may cause additional ionization; the ionization grows exponentially. This is termed the ‘avalanche effect’, recalling the initiation of a massive snow slide by a small initial snowball. The second mechanism is a quantum mechanical effect more difficult to describe by a familiar analogy. Quantum mechanics predicts the possibility of a spontaneous crossing of a semiconductor junction by carriers subject to a strong electric field. This is called the ‘tunnel effect’; because the phenomena is not associated with ordinary mechanics it was suggested that some sort of metaphysical tunnel existed through which carriers somehow traveled out of ordinary sight.

While the breakdown characteristics for the two phenomena are not exactly the same they are close enough so that the distinction may be ignored in general for purposes of circuit design. Thus although the Zener effect originally referred to the quantum mechanical phenomena the label Zener diode is applied almost universally whatever the details of the breakdown mechanism.

An illustrative breakdown characteristic is drawn to the left; the scale is exaggerated for clarity. The nominal Zener reference voltage of the diode is the reverse-bias voltage at which a manufacturer-specified 'test' current $I_{ZT}$ flows, and typically represents a rated maximum diode current. In general the Zener voltage is a modest function of temperature; a representative temperature specification is 0.1 % change per °C change. The coefficient is negative for a diode with a reference voltage below about 5 volts, otherwise it is positive. (This is related to the dominance of one or the other of the two phenomena producing similar terminal breakdown characteristics.)

The inverse of the slope of the diode characteristic at the test point is called the 'dynamic resistance' of the diode, and is a parameter noted in the manufacturers' specifications. The slope of the characteristic does not vary greatly for currents in the range (roughly) between 0.1 $I_{ZT}$ and $I_{ZT}$, a usual range of operation of a Zener diode. (Note again that the scale in the figure is distorted for illustrative purposes.) The minimum usable current is conditioned by the necessity of operation above the knee, i.e., in the
breakdown region, and the general desirability of avoiding the rapid change of slope in the immediate vicinity of the knee.

Refer to the text and notes for details.

**Zener Test**

A test circuit as shown to the right is analyzed numerically using the PSpice computer analysis program. It should be clear that for the positive half-cycle of the sinusoidal input voltage the diode is reverse-biased and $V_o$ will 'stick' at the breakdown voltage ($\approx 5.1V$). For the negative half-cycle the diode is forward-biased and $V_o$ will be small.

The plot shows the output voltage 'clipped' by the forward-bias characteristic of the diode. Note that $V_o$ is slightly negative (diode threshold $\approx 0.7V$) for forward-bias operation. For reverse-bias operation ($V_o > 0$) the voltage again is clipped, this time at the Zener voltage.