

Semiconductor Diodes

Junction Phenomena; Diode Characteristics

Objective

The PN junction semiconductor diode is a real device whose steady-state terminal volt-ampere approximates that for the idealized diode definition. Although they have significantly different terminal characteristics it is nevertheless common practice for both devices to be represented by the same icon. Hence some care is needed to distinguish which diode the icon implies; however generally this is unambiguously clear from context. The diode is an important circuit element in its own right, as well as being important in understanding the operation of other devices, in particular but not limited to the bipolar junction transistor.

As is true generally for semiconductor devices the diode terminal volt-ampere relation is nonlinear. Approximating the diode by a simpler PWL model is very useful for reasons noted elsewhere. In this first of several introductory notes on diodes a qualitative description of junction phenomena leading to the basic diode terminal relation is presented. This is intended to provide a conceptual background against which to better appreciate junction characteristics. Then the characteristic of a representative diode (1N4004) is examined (via a computer analysis using a behavioral diode model) to illustrate theoretical expectations. Some often useful generalizations are made about diode characteristics. Then the idealized diode is used to describe modeling of real device characteristics. Finally a useful computer model approximating an idealized diode is described, primarily for a limited pedagogical use.

Addendum:

a) For the present we assume that the diode internal physical processes are fast enough to be able to adapt essentially instantaneously to changes in diode terminal voltage. The practical meaning of 'instantaneous' is that physical changes in the diode state are more or less complete before terminal voltages and currents change significantly. However the internal response time although it usually is very short still is finite, and in a number of increasingly common applications a particular diode may not respond quite fast enough for the reaction to be 'instantaneous'. This aspect of the dynamic behavior of the diode we consider separately later.

b) Computed terminal characteristics are used to illustrate the discussion of device characteristics. There are some aspects of this use that should be appreciated. The computations are derived from mathematical models, not real devices. However model parameters have been selected to closely characterize measured terminal properties, and the models are extremely valuable in predicting device behavior in a circuit. Still they do not in general reflect all the physical interactions in the operation of a real device, and the distinction between a model and the device modeled ought to be kept in mind always. Indeed in many cases the mathematical description of the model is not a description of the device physical processes themselves but rather is a 'behavioral' model simulating device terminal behavior mathematically without direct reference to physics. The point is simply that computer models are invaluable but not infallible.

Semiconductor diodes

Introduction: A detailed technical examination of the fabrication technology and the physical phenomena associated with a semiconductor junction is no simple undertaking, and it is not to be found here. On the other hand a broad qualitative appreciation of semiconductor junction physics can be of enormous help in the application of semiconductor devices in electronic circuits. Hence we start the examination of diodes with an abbreviated qualitative discussion of junction physics, providing a general understanding of the device terminal volt-ampere relation. Even this qualitative description is not a simple undertaking; it involves a preliminary discussion of a number of seemingly unrelated topics which ultimately are combined to describe a semiconductor junction. Be patient.

While there are several semiconductors of technological interest it is sufficient for our needs to consider only the dominant silicon semiconductors. Silicon exhibits a remarkable confluence of metallurgical,

chemical, electrical, and mechanical properties that give it commercial preeminence as a semiconductor material. Hence we approach a qualitative examination of the junction by first considering the element silicon itself. Silicon, the most plentiful element on Earth, is the principal constituent of sand (ordinary glass is primarily SiO_2), and is particularly interesting technologically because of what can be done to modify its properties. We consider a dominant method of making such changes, impurity doping of silicon. The purpose of this doping is modify the conduction properties of the intrinsic silicon, and this leads to a discussion of how charge carriers move about in a semiconductor. Finally all this is put together in a discussion of the behavior of a silicon junction.

Silicon Crystal: The attractive Coulomb force between electric charges of opposite sign is very strong, and because of this it is difficult to maintain a separation of opposite electrical charges. To do so requires a counter-force to balance the strong Coulomb attraction. The particular method of providing the counter force and maintaining a charge separation in a semiconductor junction involves the use of impurity-doped semiconductors. Although several semiconductor materials have commercial importance it is silicon which offers a particularly harmonious combination of chemical, metallurgical, electrical, and mechanical properties. Silicon, appropriately encouraged, forms in a crystalline (i.e., geometrically regular) atomic structure for which the interaction of atomic forces tend to firmly position the atoms symmetrically with respect to one another. (Strictly speaking it is not correct, although commonly done, to refer to silicon crystal 'atoms'; the interaction between an isolated positively charged atomic nucleus and its surrounding electrons, while derived from atomic properties, is changed significantly in the crystal because of the influence of neighboring atoms.)

The character of bonding between silicon crystal atoms is important. A silicon atom is a member of Group IV of the periodic table, meaning that there are four 'outermost' electrons around the atomic nucleus. The special importance of these electrons is basically that by virtue of their relatively remote distance from the nucleus they are the electrons which are least tightly bound to the positively charged nucleus by the atomic Coulomb forces, and so most easily influenced by externally applied forces. It happens that an arrangement of eight outermost electrons is a particularly stable atomic arrangement, almost indifferent to external forces. Examples of such a highly stable atom with eight outermost electrons are provided by the inert noble gases, e.g., neon. When a silicon crystal is formed the atomic geometric positioning is such that each silicon atom is proximate to four nearest-neighbor atoms. Each atom then shares one of its outermost (valence) electrons with each of the four neighbors. The net effect of this sharing is that each atom enjoys an average electron configuration similar to although not quite the inert eight-electron configuration.

Because the shared valence electrons are so tightly bound to the crystal nuclei it turns out that a perfect silicon crystal actually would be a non-conductor. That is the valence electrons are so tightly bound in such a case that none can be moved by an external (non-destructive) electric force to produce a current. In practice however crystals are never perfect, and there always are a few 'movable' electrons; thermal excitation causes the silicon atoms to 'jiggle' a bit in place so as to not quite form the perfect geometry, slightly upsetting the balanced sharing of the valence electrons. Some, it is an exceedingly small fraction, of the valence electrons have enough thermal energy added to be able to break away from a nucleus in response to an external electric field. (This is a statistical process, i.e., it is not the same electrons that are thermally excited at any one time. Atoms that have 'lost' an electron continually recapture excited electrons, while other atoms continually release electrons.) The thermally excited valence electrons, less tightly bound to atoms, can be coerced by an external electric field. Thus silicon in practice has a small conductivity, very much smaller than, say, copper but still much greater than glass (hence 'semi-conductor'). This modest conductivity is very temperature-sensitive, increasing as rising temperature increases the thermal energy available to excite valence electrons.

The same thermal excitation process, which produces the mobile excited conduction electrons, indirectly provides means for other electrons to move about. Consider a valence electron in a crystal atom which, for the moment, is not thermally excited. For that electron the overall resultant force of all the atomic forces acting on that electron must be essentially zero; the electron is after all not mobile. If that electron then is thermally excited, and departs the atom to wander about the crystal, the net force from other crystal atoms where it had been located acts so as to encourage an electron to replace the departed electron, i.e., rebalance the net force. A comparatively weak external electric field, when supplemented by this strong crystalline

background force unbalance, can cause a valence electron to shift from one atom into a vacated position of an electron in a neighboring atom. This electron transport while quite small nevertheless constitutes an electric current. Note that the electrons involved in this sort of transport are not the conduction electrons that have been thermally excited away from the influence of the parent atoms. It is the otherwise firmly bound valence electrons that are moved from one tightly bound atomic position in the crystal to another by slight upsets of the strong but nearly balanced internal forces.

It turns out to be easier mathematically, although conceptually more subtle, not to track the valence electron movement directly but rather to track the movement of the 'hole' into which a valence electron is attracted. (Consider, as a crude analogy, an air bubble in a glass of water. The description of the movement of all the individual water molecules is rather complicated. On the other hand the collective result of their movement is described rather more easily by describing the movement of just the air bubble.)

In many respects 'hole' movement can be described conveniently as the movement of a fictional positive charge; if an external field moves a valence electron in one direction it moves the 'hole' in the other. (The mathematics involved are not quite as simple as this description may suggest; it is a statistical process, and it is simply a mistake to think there exist real positive charges moving about in the semiconductor.) The holes are used to describe the net transport of a certain class of valence electrons in an environment that includes immobile positively charged nuclei.

A useful ancillary effect of the hole concept is that it inherently separates the electrons involved in charge transport into two conceptually distinguishable groups. The 'conduction' electrons are those electrons thermally excited away from the atoms, and the 'holes' are associated with valence electrons still tightly bound to the atoms that have some mobility because of assistance from internal crystal forces. Both currents associated with these mechanisms in an intrinsic semiconductor are small. Conduction current is small because although conduction electrons move relatively fairly freely in response to an external force there are few conduction electrons available to move. On the other hand while there are considerably more non-conduction valence electrons they do not respond readily to external forces because there are few 'vacated' positions for them to move into. (The 'hole' point of view describes this analogous to conduction electron current; holes move relatively freely but there are few holes present to move.)

Impurity Semiconductors

An intrinsic (pure) semiconductor is simply that; it is neither a good conductor nor a good insulator. However a semiconductor made impure in an appropriate way is of considerable technological importance. What is involved is not a casual contamination of a silicon crystal but rather the use of selected chemical impurities introduced so as to minimally disrupt the basic crystal structure in specific ways. On using appropriate metallurgical methods a small percentage of silicon atoms (1 atom in a 100 is a large proportion) distributed uniformly over the crystal are replaced by atoms from Group V of the periodic table in one important case, or from Group III in another important case.

Consider first the use of a Group V element; phosphorous (P) provides an example. As a Group V element phosphorous has five valence electrons. If a phosphorous atom replaces a silicon atom at an occasional crystal lattice point four of the P valence electrons fit conveniently in the crystal environment more or less as would the four valence electrons of silicon. It is the 'fifth' valence electron that is the significant one. This electron remains bound to the phosphorous nucleus as for the P atom, but the binding force in the crystal is weakened by the crystal lattice forces compared to that for the isolated atom. The partial shielding effect of the crystal is sufficient so that at room temperature the 'extra' valence electron is thermally excited in the same manner as a conduction electron in the intrinsic crystal. Essentially all the impurity atoms contribute to the conduction electrons in this fashion. What makes this particularly important is that even a very small percentage doping (e.g., 1 part in a 1000) contributes far more conduction electrons than are present in the intrinsic crystal. Hence the number of conduction electrons available for conduction can be controlled readily by doping. Because the conductivity of the crystal is proportional to the number of charge carriers available the crystal conductivity is controlled essentially entirely by the impurity atoms.

Elements of Group V are called 'donor' atoms; they 'donate' a conduction electron. (Incidentally keep in mind that the crystal as a whole remains electrically neutral, i.e., there is a positively charged donor atom proton for each donor electron.) A crystal modified with donor atoms is called an n-type crystal ('n' for negative, referring to the conduction-electron enhancement). Doping a crystal n-type concomitantly reduces the hole concentration because there are many more conduction electrons available to be 'caught' by the force imbalance associated a hole. The hole concentration turns out to be reduced in the same proportion that the electron concentration is increased, so emphasizing further the influence of the conduction electrons.

Doping, i.e., introducing impurities, also is done using elements from Group III of the periodic table, for example arsenic (As). Group III atoms are called 'acceptor' impurities for reasons to appear. Acceptor impurities have three valence electrons, and as with the donor atoms an occasional acceptor atom substitutes nicely into the crystal for a silicon atom, contributing replacements for three of the four silicon valence electrons. The absence of a fourth valence electron means that the crystalline force imbalance establishes a 'hole' where the fourth electron could be accepted (ergo 'acceptor'). As before even a modest doping density establishes far more holes than occur in the intrinsic crystal, increasing the impure crystal conductivity by increasing the number of holes. And similar to the n-type case p-type doping ('p' for the positively charged fictional holes) reduces the conduction electron concentration by the same factor the p concentration is increased.

It is important enough to emphasize again that impurity doping is not an arbitrary introduction of impurities into a silicon crystal. The type of impurity is chosen with the specific intent of emphasizing one or the other of the intrinsic charge carrier types.

There is an implicit aspect of impurity doping that also needs to be emphasized. Consider, for example, an n-type crystal. Most of the atoms in the crystal are remote from an occasional donor impurity atom location. After a donor atom is ionized, i.e., the 'fifth' electron has been thermally excited and departed the vicinity the initially electrically neutral donor atom presents a local net positive charge. This is an immobile charge, i.e., the donor atom is locked in place in the crystal by the atomic forces of the other atoms. Except close to the donor atom the attraction of that positive charge is small; elsewhere in the crystal the atomic forces largely cancel one another. As will be described later it is the lattice forces locking an ionized (charged) impurity atom in place that provides a counter force allowing the charge separation essential for junction operation.

Similarly for an acceptor impurity. An acceptor impurity when ionized has locked onto an electron to balance the forces of the surrounding crystal, and so provides an immobile local negative charge.

Transport Processes A doped crystal provides an environment with an enhanced charge concentration of one or the other type of carrier. But to discuss the formation of a semiconductor junction we need first to consider the mechanisms by which carriers, electrons (conduction band electrons) or holes (valence band electrons), can be caused to move and so form a current.

Probably the most familiar electrical transport process is 'drift' motion, movement in response to an applied electric force. However a crystal is an arrangement of atoms, not a vacuum, and as a thermally excited conduction electron (for example) moves into close proximity to an atom it experiences the extremely strong local electric forces of that atomic core. As a result the electron is deflected and moves away in some other direction (the interaction is called a 'collision' although there is no direct contact in the macroscopic sense). This randomizing deflection process occurs repeatedly, and absent an external force the electron although it travels extensively makes no net displacement on average over a period of time. If however an external force is applied the motion becomes partially organized. Small as the external force may be relative to the internal atomic forces, it is a uniform force during each otherwise random collision, and so biases each otherwise random displacement in the same way on every leg. Over a period of time, there is a net displacement parallel to the external force. (Imagine taking a snapshot of an electron displacement periodically. Most of the motion occurring between snapshots would have a zero net displacement, but the regularity of the external force would regularly add a small net displacement in the same direction from snapshot to snapshot.) The velocity of this net drift (average displacement between

snapshots÷snapshot time interval) is proportional to the external force. This is not a contradiction of Newton's Law; rather it is a consequence of Newton's Law averaged over a period of time to eliminate the effect of the collision forces. The resulting charge drift transport is what is described, for example, by Ohm's Law.

Whereas drift motion is a consequence of the application of an external force there is another important transport mechanism that results in charged carrier transport without an external applied force; this is diffusion along a concentration gradient. This is the mechanism by which, for example, a small amount of dye dropped in one corner of a bowl of water spreads out to fill the entire bowl uniformly. Consider, to be specific, an n-type semiconductor in a steady-state condition at a uniform temperature for which (somehow) the doping level varies with position. Consider further a small volume element at a position x (one-dimensional gradient for simplicity). The conduction electrons have a random thermal motion, and at any time half the electrons in that volume element move in the $+x$ direction, the other half move in the $-x$ direction. The same statement applies to an adjacent volume element. But if that second volume element has, say, a lower electron concentration because of the concentration gradient then there are numerically more electrons moving from the first volume element into the second than move in the opposite direction. Hence there is a net transport of electrons from the higher concentration region towards the lower concentration region. The diffusive transport always is in the direction of the gradient, tending to change the concentration so as to reduce the gradient. No electrical force is involved; the same transport mechanism occurs with uncharged particles. For example in one commercial process n-type doping of integrated circuits is done by passing a phosphorous-rich gas at high temperature over a silicon crystal; phosphorous atoms diffuse over time from a high concentration at the crystal surface into the interior.

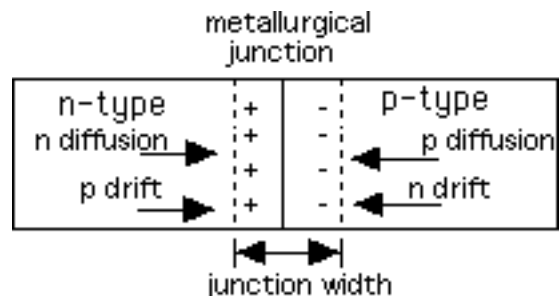
Semiconductor Junction

A semiconductor junction is commonly formed by doping a semiconductor so that the impurity concentration varies from n-type to p-type. The electrical junction forms about the 'metallurgical' junction where the transition from one doping type to the other occurs. It is important to note that simply butting an n-type semiconductor against a p-type semiconductor does not form a junction. The irregular surface atomic forces at such a physical discontinuity at the interface simply prevent junction formation. It is essential that the crystalline background forces are uniform across the junction, i.e., the basic internal crystal regularity is maintained across the junction.

Externally supplied voltages applied to a junction are applied to act on the electron and hole carriers. Compared to the internal atomic forces, only very small forces can be applied externally. But by using the crystal lattice to balance atomic forces so that they roughly cancel one another the small externally applied force becomes the significant one. Modest doping in an otherwise uniform crystal enables this balance to be maintained across the junction.

Consider for simplicity a 'step' junction, for which there is a relatively abrupt change from n-type to p-type (see figure). Doping a semiconductor n-type increases the conduction electron density above the intrinsic density by some large factor, concurrently reducing the hole density by the same factor. Similarly for hole density for p-type doping. We focus attention here on the electrons; the discussion for holes is formally the same.

At the metallurgical junction there is a conduction electron concentration gradient from the n-type side to the p-type side. Hence there is a diffusion of electrons from the n into the p material, and a consequent separation of charge. The conduction electrons are provided essentially entirely from donor atoms doping the n material. Note that these donor atoms are electrically neutral when introduced. Hence when the electron is removed from its donor it leaves behind the uncompensated positive charge of the nucleus. A key point to keep in mind is that the ionized donor atom is not mobile; it is locked into the crystal structure by very strong atomic forces and thus provides an immobile positive charge.



There is a corresponding diffusion of holes from the p side into the n region; the source of the holes is the acceptor impurities doping the p material. Here also note that acceptor impurities are electrically neutral when introduced. When an acceptor releases a hole, meaning the localized forces of the crystal have captured an electron, the initially neutral acceptor becomes negatively charged. This is an immobile negative charge, i.e., the acceptor is bound in place by the crystal forces.

The net result of these diffusions is that positive and negative electric charges have been separated, and this separation is maintained because strong crystalline atomic forces offset the Coulomb force attracting oppositely signed charges together. The distance of separation is not large and so the electric field associated with even a small amount of charge is quite strong.

Because of the separated charges there is an electric force present, the immobile ionized donor positive charge on the n-side and the electrons which have diffused into the p region and attached to acceptor impurities. This force is directed so as to cause electrons to drift from the p region back to the n region. The diffusion and drift processes are self-balancing; eventually (essentially immediately for ordinary purposes) an equilibrium is established in which electron diffusion into the p region is balanced by electron drift back to the n region, i.e., no net transport. This is a dynamic equilibrium; there is still electron transport (generally very large current densities in fact) but as much current flows in the one direction as in the other so that there is zero net charge transported.

A dual description applies to the hole diffusion and drift currents.

With a broad overall description of the junction formation in place we can refine some of the details of the junction phenomena. Consider electron flow to be specific. The concentration gradient causes a diffusion of electrons from the n region towards the p region. Because of the net positive charge of the ionized donor atoms on the n side conduction electrons diffusing towards the p side encounter a very strong electric field which opposes movement towards the p region. Few of the diffusing electrons have sufficient thermal energy to penetrate far against this field, and so after a short journey most are drawn back to the n region. Those few of the electrons that do have sufficient energy to cross the junction constitute the net diffusive flow into the p region. As long as a net flow of electrons into the p region persists charge separation increases (between fixed ionized donor atoms and diffusing conduction electrons) and the electric field consequent on the charge separation grows stronger.

Because of the electric field there is a drift movement of electrons from the p region back to the n region. Electrons in the p region, there are very few of these of course, whose random thermal motion carries them into the junction find a strong electric field directed so as to sweep them across to the n region. Eventually the electric field strength increases to a point where the net diffusion of electrons from the n into the p region (many electrons start, but few complete the journey) is matched by the drift of electrons back from the p to the n region (there are few electrons, but virtually all complete the journey). At this point there is no increase in charge separation and a dynamic balance is established.

In addition to the electron flow there is of course the corresponding balanced hole flow. Except that the holes (describing valence band electrons) act on average as if they were positive particles the process is formally the same as for electrons. There is a diffusion of holes from the p region into the n region, and this establishes an additional charge separation with a consequent hole drift back to the p region. The (negatively charged) ionized acceptors provide an immobile negative charge. (It is worth emphasizing yet again that there really is no positive-charge movement. The 'holes' simply provide a convenient mathematical means of describing the average transport of negatively charged valence electrons.)

In summary then the junction proper consists of adjacent very narrow regions of immobile ionized impurity charge to either side of the metallurgical junction; positive charges provided by ionized donor impurities and negative charges provided by ionized acceptor impurities. These charges are kept from recombining by the atomic forces which lock the crystal atoms, ionized or not, into fixed positions. Carriers, for example electrons, diffusing into the junction from the n region must overcome an electric force that opposes their flow; not all electrons entering the junction have enough thermal energy to pass through the junction and are forced back to the n region by the electric field. On the other hand electrons in the p region whose

thermal motion brings them into the junction are immediately swept across to the n region. A similar description applies for holes. Equilibrium occurs when the electric field becomes strong enough so that the net electron diffusion into the p region is balanced by the net electron drift from the n region, and similarly for the holes.

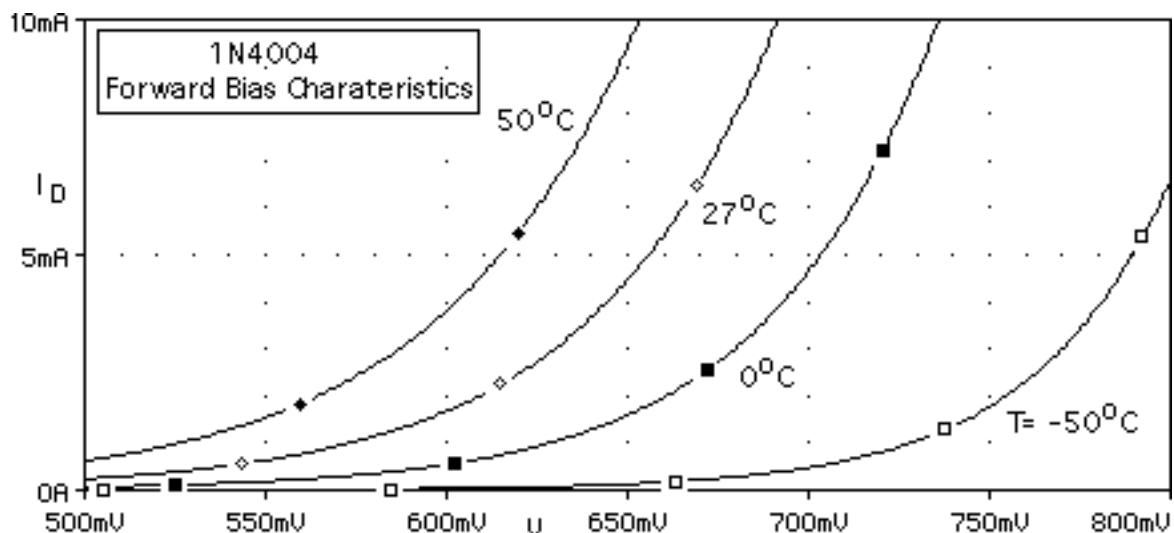
This delicate balance of very strong atomic forces is what is to be modified by applying an external voltage across the junction. (Actually not quite across the junction itself; some bulk material is required on either side of the extremely narrow junction simply to provide the physical strength needed to attach a connection). Because the very strong internal forces are balanced a small imbalance caused by a comparatively weak external force can have significant consequences.

Forward Bias

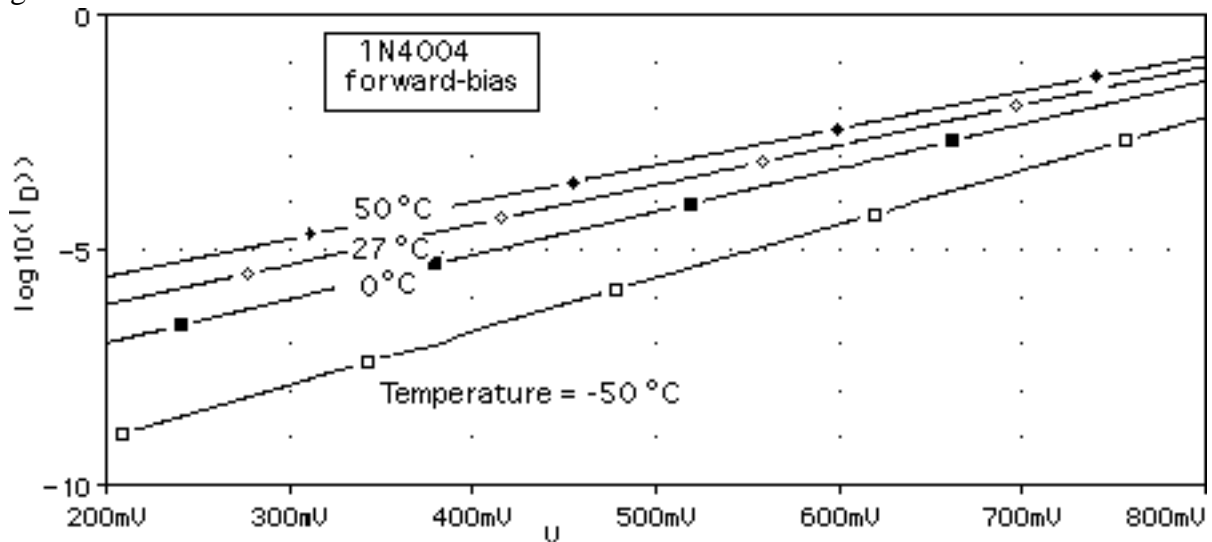
An external voltage can be applied across a junction with a polarity that either enhances the internal electric field of the junction, or reduces it. In either case the externally applied field will be very small compared to the built-in junction field. Consider the 'reduction' case first. For reasons to be seen this is described as 'forward biasing' the junction, and leads to a relatively large net junction current density.

Consider the electron transport first. If the junction is forward biased, i.e., the internal junction electric field is reduced, there will be essentially no effect on the drift transport across the junction of electrons initially in the p region which randomly wander into the junction and on to the n region. The electric field, even though slightly weakened, nevertheless remains strong and directed so as to sweep these electrons into the n region. On the other hand the diffusive electron flow from the n into the p region will be significantly affected. The reason is that the energy requirement to cross the junction is reduced (slightly) because the opposing electric force is reduced, and so more electrons have the initial thermal energy necessary to overcome the force and cross the junction. The change in the number of such electrons increases exponentially with an increase in forward-bias voltage. Thus a small voltage change causes a relatively large change in the electron flow into the p region. Recall that absent the forward bias the carrier flow is balanced. What forward bias creates is an imbalance in current flow. It is not so much a matter of adding to the current flow as it is enabling more carriers in an existing carrier flow to penetrate the force barrier imposed by the junction field.

The computed forward-bias volt-ampere characteristic of a 1N4004 diode (PSpice model used to compute each characteristic) is shown below for several temperatures; this is representative of small-geometry diode characteristics. Note that the forward-bias current does not reach, say, the milliamper level until a threshold ('knee') voltage of roughly 0.5 volt is crossed. It is useful to note, as is generally true for diodes of the 1N4004 type, that roughly 0.1-volt change in forward bias increases the current by about an order of magnitude. Note further that the forward-bias voltage needed for a given current decreases as the temperature increases, i.e., as the average thermal energy of the carriers increases. For silicon the bias voltage needed for a given current is reduced roughly 2 millivolt per °C temperature increase.



Because of the large current range for a forward-bias junction a logarithmic re-plot of the characteristics as shown below is informative. Note the close adherence to an exponential dependence of current on forward-bias voltage over the several orders of magnitude plotted. For diode operation in the milliampere current range around room temperature the bias will be roughly 0.6 to 0.7 volt. Note again that around room temperature a bias voltage increase of 50 to 100 millivolts or so increases the current by roughly an order of magnitude.



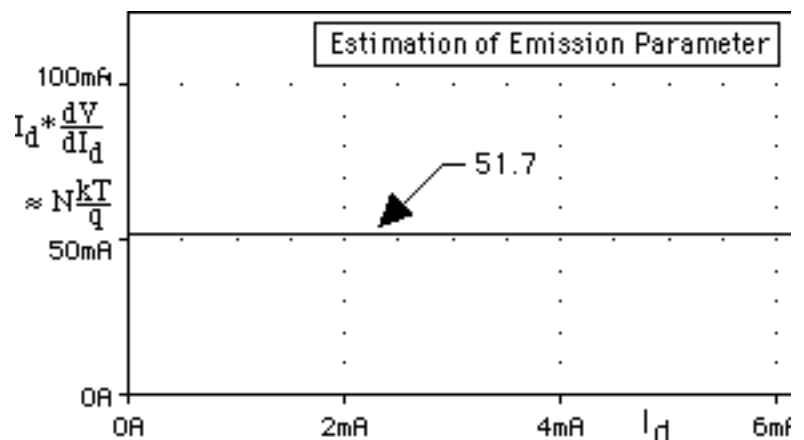
The equation on the right is a theoretical expression (not derived here) for the diode volt-ampere relation. The voltage and current polarity conventions are indicated on the conventional diode icon; the icon reflects the non-bilateral behavior of the junction, i.e., there is a direction (forward bias) of 'easy' current flow, and a direction (reverse bias) of 'hard' current flow. The icon is an arrow which 'points' in the direction of easy current flow; forward bias corresponds to $V > 0$. The 'emission' parameter N

$N = 2$ allows heuristically for the effect of different diode fabrication methods and construction. $N = 1$ for diodes imbedded in an integrated circuit, and $N = 2$ for discrete diodes.

q is the electron charge, k is Boltzman's constant, and T is the Kelvin temperature. I_s is a temperature-dependent device parameter reflecting various semiconductor material properties. At room temperature $qV/kT = 40\text{mv}$, so that a forward bias of about 0.2 volt or more makes the exponential term $\gg 1$. I_s is typically a very small current, nanoamperes or less for diodes not specifically designed to carry high currents. Hence forward-bias voltages of several tenths volt are necessary to bring the diode current above the microampere level; this is the semi-quantitative 'knee' voltage mentioned before. Values for N and I_s can be estimated for the 1N4004 from the forward-bias curves above.

$$I_d = I_s \left(e^{\frac{qV}{NkT}} - 1 \right)$$

At the risk of a compounded sophistry the PSpice 1N4004 diode model is used to estimate the emission parameter by plotting data as shown in the figure to the right. Actually the precise value of the emission parameter is 1.984 and the calculated value of the ordinate in the figure following is 51.6 at room temperature; this value is built into the model explicitly. On the other hand the model quite accurately reflects the actual device characteristics, and the computation serves to suggest what an experimental determination almost certainly would show, i.e.,

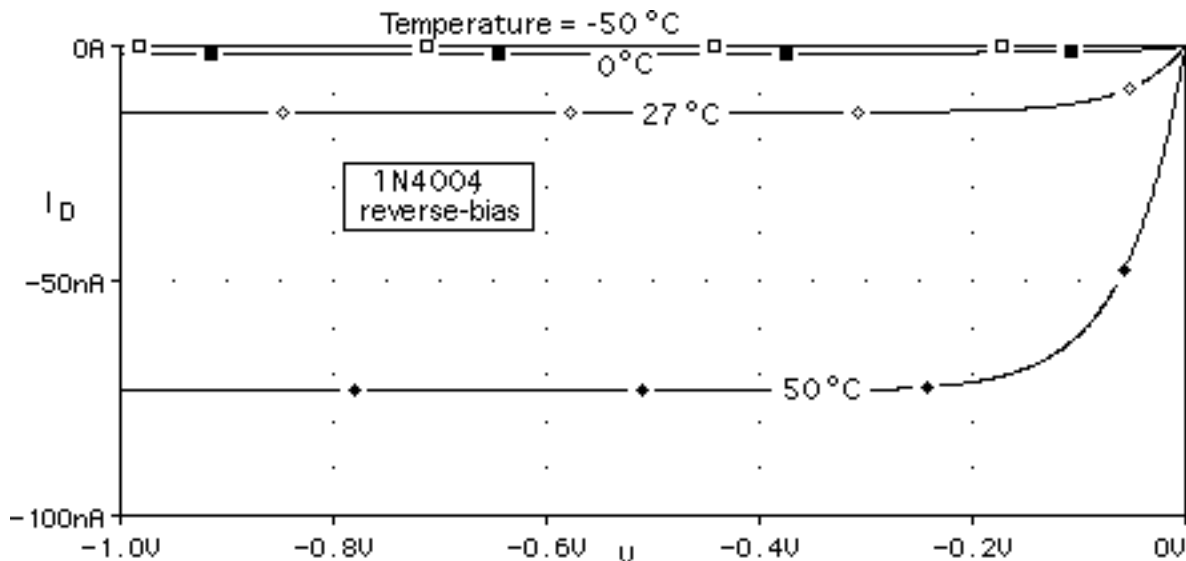


$$\frac{dI_d}{dV} = \frac{qV}{NkT} I_S e^{\frac{qV}{NkT}} \approx \frac{qV}{NkT} I_d$$

Reverse Bias

If the external voltage polarity is such as to cause the intrinsic junction electric field to be increased the diode is said to be 'reverse-biased'. In this case the energy requirement to cross the junction is increased, and the number of electrons which have sufficient energy to diffuse across decreases exponentially. However the thermal drift component from the p region still does not change appreciably; any electrons in the p region whose thermal meandering bring them into the junction still are swept across. But even a small reverse-bias voltage exponentially reduces the diffusive component of the current flow balance to a negligible level, leaving just the small, essentially constant drift component.

In the theoretical expression for the diode current this corresponds to the exponential becoming much less than 1 with a few tenths volt reverse bias. The following figure plots the reverse-bias current for several temperatures. Note the considerable difference between the reverse-bias current level and the forward-bias current range.



Incidentally what is measured experimentally as the reverse-bias current ordinarily is not actually I_S . I_S generally is so small that it is difficult to measure directly and, in addition, other conduction phenomena tend to mask the current. Suppose, for example, there is a resistance of 10 Megohm shunted across the 1N4004 diode; contamination of one sort or another, e.g., skin oils, could do this. A reverse-bias voltage of just 1 volt then leads to a 'reverse' current through this shunting resistance of 100 nA, an order of magnitude larger, for example, than I_S for the 1N4004. For forward bias of course this comparatively small current is not significant. This sort of shunting is evidenced by an apparent reverse-bias current that increases as the reverse-bias voltage magnitude increases. To determine I_S experimentally preferably measure the forward-bias current for $I \gg I_S$, and plot $\ln(I)$ vs. junction voltage V . This is a line which extrapolated to $V = 0$ provides $\ln(I_S)$.

Idealized Diode (computer model)

The idea of creating an idealized diode model to simplify 'hand' calculations in making design estimates was introduced as part of a discussion of PWL analysis. Idealized diodes are useful to simplify 'hand' calculations, enabling relatively quick and often useful approximate analytic calculations. Such calculations often provide a good indicator of the interaction of circuit components to produce a given circuit performance, as well as some idea of the relative importance of the component contributions. This kind of understanding is more difficult to come by using a numerical calculation. It also is often useful to suppress secondary aspects of actual circuit behavior to concentrate on the dominant aspects; this is also something an approximation can do. Use of simplified models supports the application of judgment to design situations.

The more accurate nonlinear diode models would be used for computer calculations; ordinarily there is no particular computational advantage to temporizing with a less accurate computer model. It may appear peculiar therefore to degrade a nonlinear diode model to approximate an idealized diode for a computer computation.. But this apparent contradiction should be evaluated in terms of objectives. At the very least for instructional purposes, it is often convenient to use a computer analysis to compare the results of an idealized diode PWL circuit analysis with those for an analysis using a nonlinear device model. For this and similar purposes a computer model for the idealized diode is useful. This requires rather more care to develop that might appear to be the case at first. The discontinuity in slope at the origin creates some computational difficulties. It turns out, almost perversely, that a continuously differentiable nonlinear model for the idealized diode is to be preferred. (This may be considered an indication of an important distinction between a hand calculation and a computer analysis.) There are several ways of devising a satisfactory model, but probably the simplest way is to modify an existing nonlinear diode model appropriately. The following remarks, although generally applicable, are expressed specifically for PSpice use.

Suppose the diode emission parameter N is made very small, say 10^{-6} . In effect this reduces the voltage required for a given forward bias current by a factor of one million! If about a 50 mV change ordinarily causes an order of magnitude current change in a real diode then the modified diode requires only 50 nanovolts for the same change. Hence the forward-bias diode characteristic rises quite sharply, not abruptly but continuously, within a few nanovolts of the origin. (Keep in mind that this is a computational trick; the modified nonlinear model does not correspond to a real diode.) The reverse current is not the idealized zero value but it is quite small (and can be made smaller if the need arises). There is an additional 'catch' however that has to be considered to avoid computational problems. Because of the extraordinarily large increase in forward current with small forward-bias voltage it is possible for the computed current magnitude to exceed the storage size limitations of the analysis program or to become so large as to cause mathematical convergence problems. To avoid this another parameter in the computer diode model can be used; the nonlinear model includes a parasitic resistance (RS) in series with the 'real' diode, representing for example lead resistance. This can be used as a current limiting resistance; a milliohm will do nicely to limit current magnitudes to less than about 10 kiloamperes. (Again note this is just a computational trick to avoid computational problems.)

To realize this idealized diode computer model in PSpice declare a diode as

Dxxx from_node to_node DIDEAL

where DIDEAL is a reference to a special model. Then declare the special model with the statement

.MODEL DIDEAL D(N=1U, RS = .001)

PSpice will use the built-in default diode model with the parameters N and RS changed as prescribed. Examples of this usage appear below.

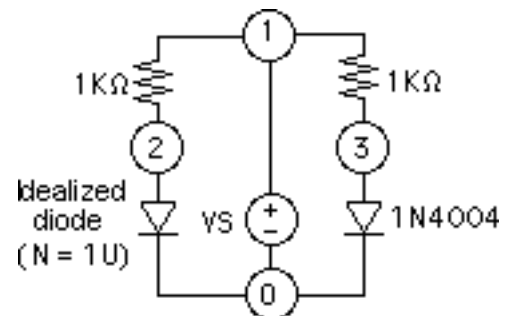
Diode Comparison Computation

The following computer analysis makes a comparison between a 1N4004 nonlinear PSpice model (closely approximating real diode characteristics) and a diode model idealized as described above.

```

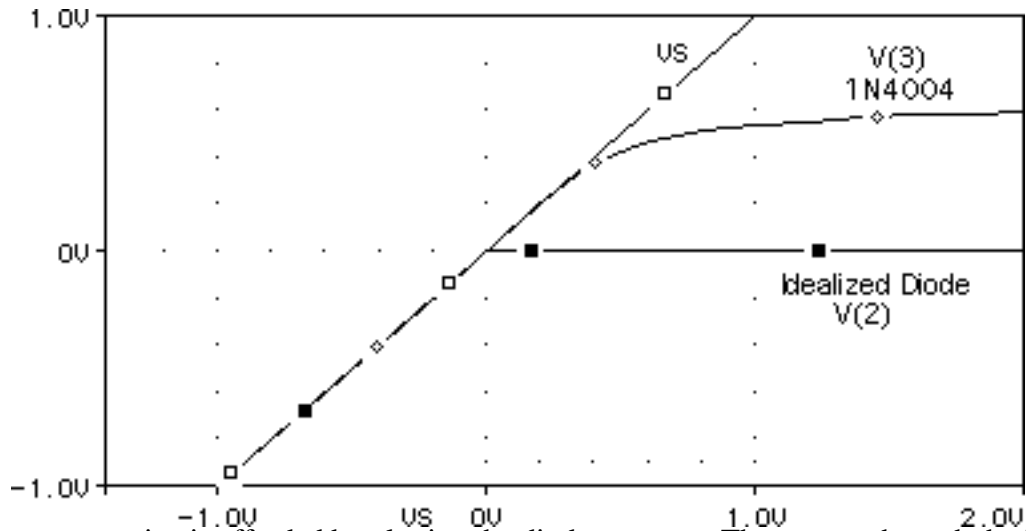
VS      1  0  DC0V
RS      1  2  1K
D1      2  0  DIDEAL
D2      3  0  D1N4004
RS1     1  3  1K
.MODEL  DIDEAL  D(N = 1U, RS = .001)
.MODEL  D1N4004 D(Is=14.11n N=1.984
+ Rs=33.89m Ikf=94.81 Xti=3 Eg=1.11
+ Cjo=25.89p M=.44 Vj=.3245 Fc=.5
+ Bv=600 Ibv=10u Tt=5.7u)
.DC VS -2 2 .2
.PROBE
.END

```

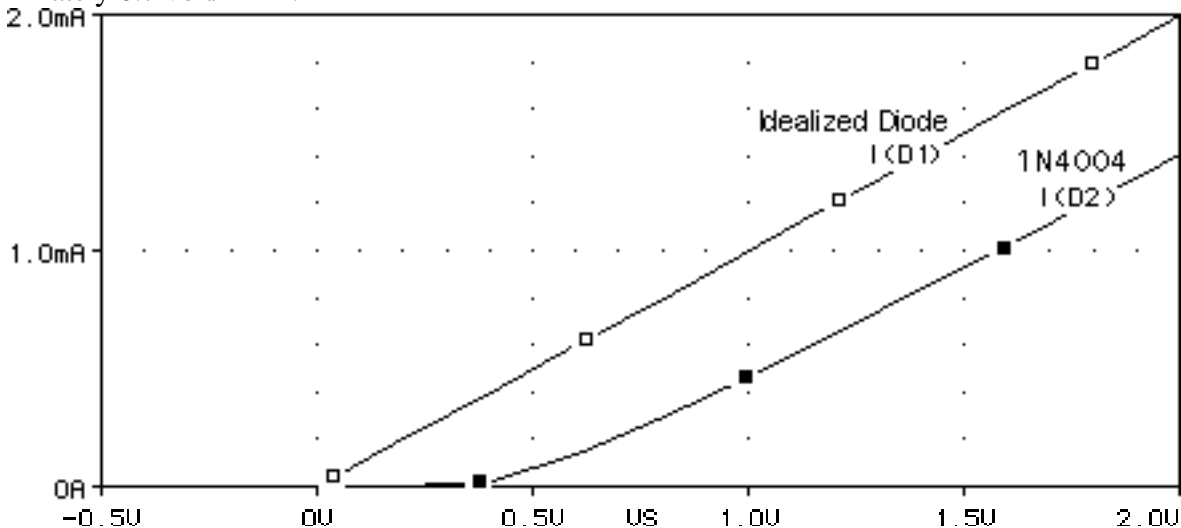


As voltage VS increases from an initial value of -2 volts the diodes eventually become forward-biased. The idealized diode 'breaks over' at essentially zero volts and the voltage drop across the diode does not change

noticeably thereafter. The 1N4004, on the other hand, must wait for the 'knee' to be passed before enough current flows to cause a significant voltage drop across the series resistor. The 1N4004 bias voltage then increases slightly as the diode current increases.



A different perspective is afforded by plotting the diode currents. The current through the idealized diode is determined by V_S and the series resistor since the forward-bias voltage is negligible; the slope corresponds as it should to a 1 K resistance. The 1N4004 characteristic shows some curvature. There is negligible current flow until the diode bias *1N4004/Idealized Diode Comparison exceeds the 'knee' of the characteristic at a few tenths of a volt. The diode current increases somewhat less rapidly than for the idealized diode because the forward voltage rises as current increases; there is hence less voltage drop across the series resistor. Eventually the source voltage becomes 'large' compared to changes in the diode voltage and the voltage drop across the resistor is essentially V_S . Note that the slope approaches 1 K as the source voltage increases. The current difference between the curves (at a given voltage) approaches approximately $0.7\text{ volt}/1\text{ K}$.



Incidentally note that the idealized diode should be a good approximation for the real diode in circuits in which the 0.7 volt (approx.) forward-bias voltage is small compared to other significant circuit voltages. (Actually an even better idealized model is easily formulated, as will be seen later, by adding a 0.7 V source in series with the idealized diode to shift the 'idealized' curve to the right.)

Reminder About Junction Dynamics

The discussion thus far has been of steady-state junction behavior. In general for a change of state to occur a corresponding charge redistribution must occur within the junction. Inevitably there is a finite transition period during which change occurs. How important this transition period is depends on whether the charge

redistribution delay is small compared to the time for which circuit voltages change significantly. If it is then the new internal steady state is reached before the external circuit can respond appreciably, and the internal redistribution delay can be ignored. In effect the device behavior can be described adequately as a sequence of steady-state conditions. This is in fact the case for a great many important applications, and in general it is presumed to be the case in subsequent discussion unless and until an explicit reconsideration is made somewhat later. Note that this does not exclude circuit dynamics from consideration, i.e., so long as significant reactive element time delays are much larger than the device time constants the latter can be neglected.

ZENER DIODE

Introduction to Zener Diodes

It is noted before that the reverse-bias saturation current of a general-purpose junction diode is so small that it is ordinarily masked by currents associated with high-resistance conducting paths across the junction. The leakage current also can be masked by currents associated with other phenomena occurring in what is a very complex physical junction environment. .

If the reverse-bias voltage magnitude is increased above a threshold (the specific value depends on the junction geometry, material parameters, and doping levels) one or the other (and possibly even both concurrently) of two new phenomena occur. These phenomena, which are different from the junction phenomena described before, establish a new mechanism of current flow, generically referred to as 'Zener breakdown', which masks the junction reverse-bias leakage current. Large current changes occur with very small changes in reverse-bias voltage, similar to forward-bias operation but for quite different reasons.

These phenomena occur in all semiconductor junction diodes. However the reverse-bias breakdown voltage characteristic can be reproduced with relative precision by carefully controlling doping and other manufacturing process parameters. For ordinary use diode breakdown is characterized simply by specification of a minimum reverse-bias breakdown voltage and current; the magnitude of the breakdown voltage is guaranteed to be no less and the current (for a stated voltage) no more than specified values. These diodes are not intended for operation in a reverse breakdown state, and are expected to maintain a low-conduction state when operated in reverse-bias within the specified breakdown voltage limit.

Diodes whose reverse-breakdown characteristics are controlled precisely during manufacture commonly are called Zener diodes. Zener diodes command a premium because of the special production controls and selection involved in their manufacture, and are intended specifically for operation in the reverse-breakdown mode primarily as inexpensive voltage references.

Zener Diode Breakdown Characteristics

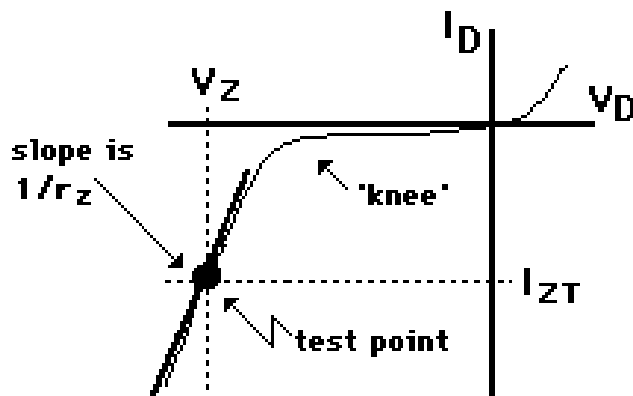
Ordinarily the reverse-bias blocking action of a PN junction allows only a small 'leakage' current to flow. However if a sufficiently large reverse-bias is applied other junction phenomena develop which dominate the leakage current, allowing in effect much larger reverse-bias currents. This is the 'breakdown' 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 phenomena 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.

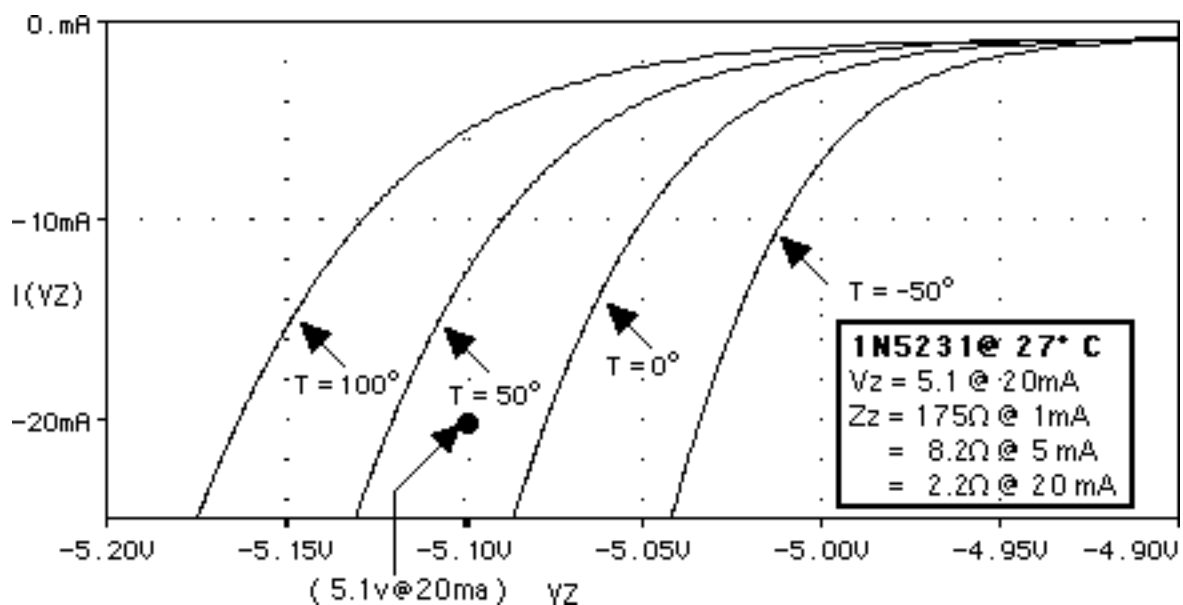
While the breakdown characteristics for the two phenomena are not exactly the same they are close enough so that the distinction may be suppressed in general for purposes of circuit design. Thus although the Zener effect originally referred to the quantum mechanical phenomena the label Zener diode is 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. The test current 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 expanded 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.

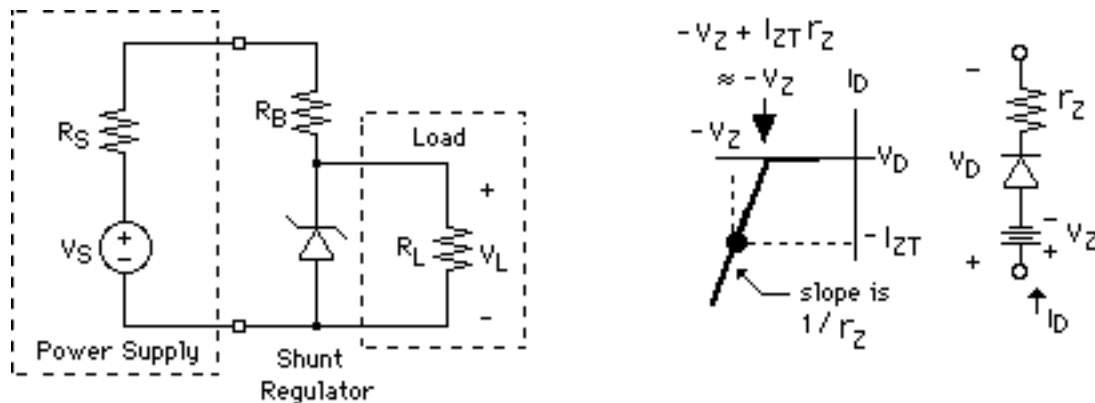
The following figure displays PSpice computations of breakdown characteristics using a nonlinear model of the 1N5231 Zener diode (Manufacturer's specifications at 27°C are: $V_Z = 5.1 @ 20mA$, $r_z = 175 @ 1mA$, $r_z = 8.2 @ 5mA$, $r_z = 2.2 @ 20mA$). The breakdown characteristics are plotted for a broad temperature range; note the small voltage sensitivity to temperature variation for a given diode current. Note also the narrow range of voltages at a given temperature that correspond to large current changes.



Since a forward-biased diode already provides large current changes for small changes in bias voltage why a special interest in breakdown operation? The simple answer is that the breakdown voltage can be manufactured to precise specifications over a very large range of voltages. For example, in a nominal 5 volt range the specified breakdown voltages for several diodes are: 1N5230@4.7v, 1N5231@5.1v, 1N5232@5.6v, and 1N5233@6.0v. When operated in the breakdown region Zener diodes are widely used as reliable and inexpensive voltage references.

Zener Diode 'Shunt' Regulator

A 'shunt voltage regulator' provides an informative relatively uncomplicated circuit application of a Zener diode; the circuit diagram is drawn below left. V_S and R_S represent the Thevenin equivalent circuit as seen looking back into the terminals of a power supply, and R_B and the Zener diode serve as a control devices to regulate the voltage across the load R_L . Note the standard icon used to represent the Zener diode.



Absent the regulating elements ($R_B \rightarrow 0$ and Zener diode removed) an increase in load current, for example by reducing R_L , produces a larger 'internal' power supply voltage drop across R_S , and consequently a lower power supply terminal voltage. This variability of the power supply terminal voltage is measured by the 'regulation' of the power supply, defined formally as the change in terminal voltage between 'no load' and 'full load' conditions, divided by the 'no load' voltage. It is essentially a measure of the effect of the internal resistance of the supply on the terminal voltage, as the load current changes from one to the other extreme of its specified operating range.

To obtain an improved regulation the power supply will be made to provide a (essentially) constant current, large enough to provide at least the maximum load current needed. When a smaller current is to be provided to the load the excess part of the constant power supply current will be diverted. Both these actions are obtained by adding the Zener diode and ballast resistor R_B to modify the load as seen by the supply. Insofar as the power supply is concerned its load current consists of the current drawn through R_L , plus the added current drawn through the Zener diode. The total power supply current thus is larger than what it would be absent the Zener diode and R_B .

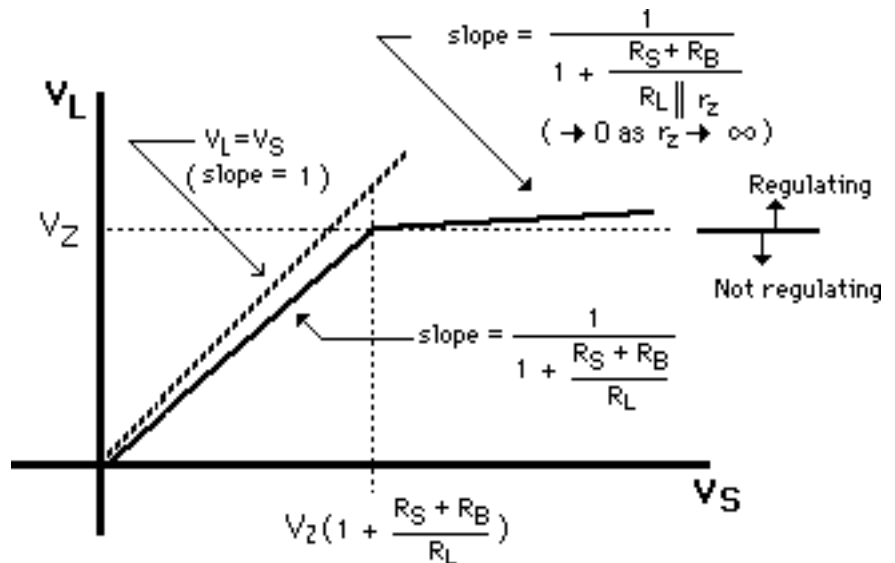
To the extent that the Zener voltage is constant (assuming operation in the breakdown region) the circuit draws a constant supply current, dividing that current between the Zener diode branch and the actual load. (Because the Zener resistance r_Z is not exactly zero the Zener voltage increases slightly as the diode current increases and this causes the supply current to decrease slightly.) The essential idea, as noted before, is to shunt 'excess' current through the Zener diode when R_L is a maximum, and then decrease the amount of this shunted current as load current demand increases. Since the power supply itself tends to 'see' a fixed current, its terminal voltage changes little. The Zener diode provides an approximation to a constant voltage source over a large current range, and R_B provides a corrective voltage drop between the supply voltage (generally larger than the Zener voltage) and the Zener regulated load voltage.

Because the diode is a nonlinear circuit element an analytical examination would be an involved one. Hence, as a simplifying measure more than adequate to provide an appreciation of the regulating action, the Zener breakdown characteristics are approximated as shown to the right of the circuit diagram above. The piecewise linear representation of the Zener characteristic, which is applicable over the range of normal

operation of the Zener diode (i.e., in reverse bias breakdown), is used for approximate calculations. (Do not confuse the idealized diode used in the model with the Zener diode whose characteristic is being modeled; the Zener diode characteristic is being approximated over a limited range of operation by a combination of idealized circuit element models.

The regulating action is observable in the curves obtained by an analysis of the PWL circuit using the piecewise-linear Zener diode approximation; these are drawn below.

The unity slope reference (dashed) line is the voltage transfer characteristic of the power supply alone, i.e., when the regulating elements are removed so that $V_L = V_S$. This provides a reference against which to display the effect of regulation. The other (solid line) curve describes the circuit with the regulating elements inserted. When the power supply voltage is low (so that $V_L < V_Z$) the Zener diode is in reverse-bias, but not yet in breakdown; the diode state is more or less open-circuit and has negligible effect on circuit operation. R_B and R_L form a resistive voltage divider, making V_L somewhat smaller than V_S , and so the regulation curve is a line segment of slope somewhat less than 1.



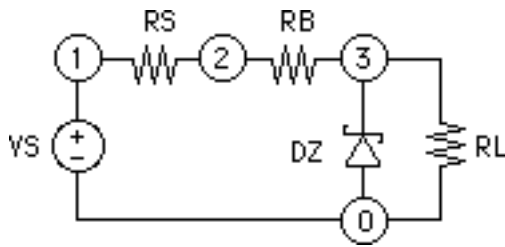
As V_S is increased the breakdown voltage V_Z of idealized diode model is reached; because of the voltage-divider action V_S will be somewhat larger than V_Z when this occurs (see figure). This causes the 'constant' Zener voltage (with some variation because of the finite Zener resistance) to appear across the load. Note that quality of the regulation is measured by $R_L || r_Z \ll R_S + R_B$.

An important implicit requirement not always recognized explicitly is that the regulating action depends on the assumed operation of the Zener diode in its breakdown region. This is not something that happens automatically; it must be designed to be so by proper choice of element values. It means, for example, enough current must be drawn by the Zener diode to maintain proper operation even in the 'worst case' situation when the maximum load current has been siphoned off from the supply current, i.e., for the minimum diode current. Hence at full-load current one should design the circuit to provide a minimum Zener 'keep-alive' current of roughly $0.1 I_{ZT}$. On the other hand when the load draws the minimum current the increased current through the Zener (the source current will not change much) should not exceed the rated I_{ZT} . Between these operating requirements, and of course knowing the (nominal) Zener diode voltage, an appropriate value of R_B can be determined.

This calculation is particularly noteworthy here because it is rather different from the more familiar case of solving for specific element values common in introductory circuits courses. Two extreme ('worst-case') conditions are involved in the form of inequalities, not equalities. The result of the calculation is not the value of the resistance to use but rather inequalities that specify a range of acceptable resistances, greater than some value but less than another. The details of the circuit behavior will depend to some extent on the choice made.

Question: What would it mean if, as is not impossible, the upper bound on a mathematically acceptable resistance value is less than the lower bound?

For comparison to the calculated regulation characteristics a PSpice computed set of characteristics using a nonlinear diode model is shown following the circuit diagram and netlist.



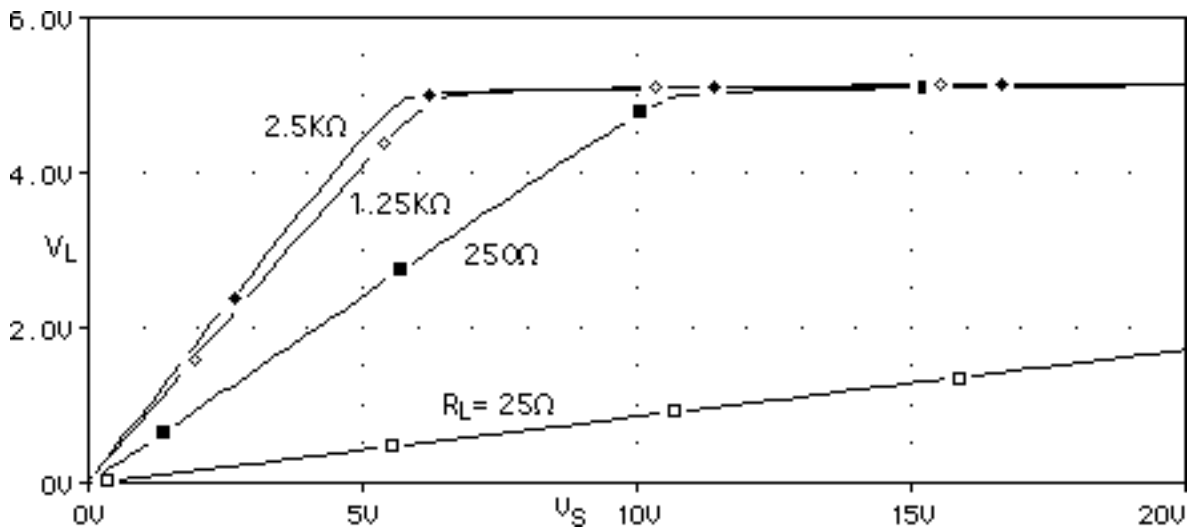
* Zener Regulator

```

VS      1  0  DC      0V
RS      1  2      220
RB      2  3      47
DZ      0  3  D1N5231
RL      3  0  RMOD 250
.MODEL D1N5231 D(Is=1.004f Rs=.5875
+ Ikf=0 N=1 Xti=3 Eg=1.11 Cjo=160p
+ M=.5484 Vj=.75 Fc=.5 Isr=1.8n Nr=2
+ Bv=5.1 +Ibv=27.721m Nbv=1.1779
+ Ibvl=1.1646m Nbv1=21.894 Tbv1=176.47u)
.MODEL RMOD RES(R= 1)
.STEP   RES RMOD(R) LIST .1 1 5 10
.DC VS 0 20 .2
.PROBE
.END

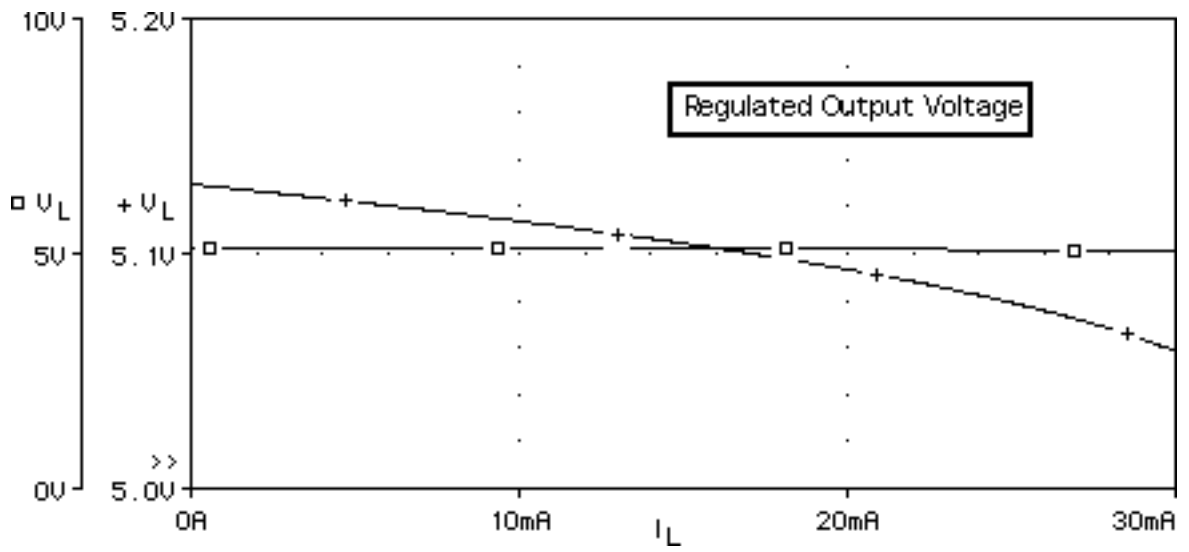
```

Regulation computations are shown for several choices of load resistance (and so load current). Detailed comparison between the computed and analytical curves is left as an exercise, e.g., calculate the theoretical value of the source voltage at which regulation begins and compare to the computed value. Thus using $V_Z = 5.2\text{v}$, $R_L = 250\ \Omega$, $R_S = 220\ \Omega$, and $R_B = 47\ \Omega$, the PWL model regulation is calculated to start at 10.6 volts.



As a matter of some interest note that there is no regulating action for $0 < V_S < 20$ when $R_L = 25\ \Omega$; not enough supply current is available over and above the required load current to enable Zener diode breakdown to occur.

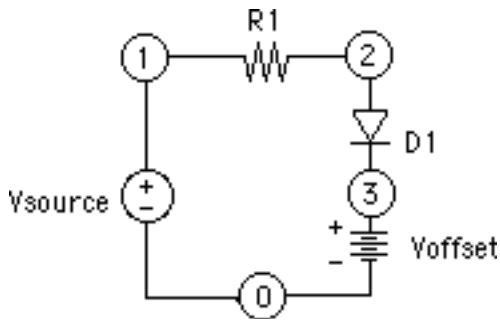
The regulated regulation curve also was computed, and is drawn below. (Replace R_L by a 'stepped' current source and fix V_S .) Note that the unregulated regulation curve would be $V_L = V_S - (R_S + R_L)I_L$; for the values used for the computation this is $V_L = 15 - 0.267I_L$ (units are volts, mA). Thus, for example, V_L @ $I_L = 30\ \text{mA}$ would be 7 volts unregulated. The data are shown using two scales, one to provide detail on the computed load voltage values and the other to provide perspective on the overall effect of the regulation. The 'cost' of the regulation in this illustration is associated with the 10 volt (approx.) drop across $R_S + R_L$.



PROBLEMS/ILLUSTRATIONS

(1) Diode Clipping

A computer analysis of the diode circuit drawn below, left, serves to further illustrate diode behavior. The (Thevenin equivalent) source presents a triangular waveform across the series combination of a diode and DC offset voltage source; refer to the accompanying PSpice netlist for parameter values used. The computer analysis uses the 1N4004 nonlinear PSpice diode model; this diode is a general-purpose low-power rectifier diode. However an idealized diode model can be used profitably both to anticipate and to evaluate the computational results. A formal PWL analysis is not difficult, but it is also not difficult to recognize directly that $V(2)$ will be equal to the source voltage when it is less than the offset voltage, and equal to the offset voltage otherwise. The computational results are plotted below. Evaluate the plot in terms of expectations from a PWL analysis. Sketch a plot of the anticipated diode currents, and then compute the plot for comparison.

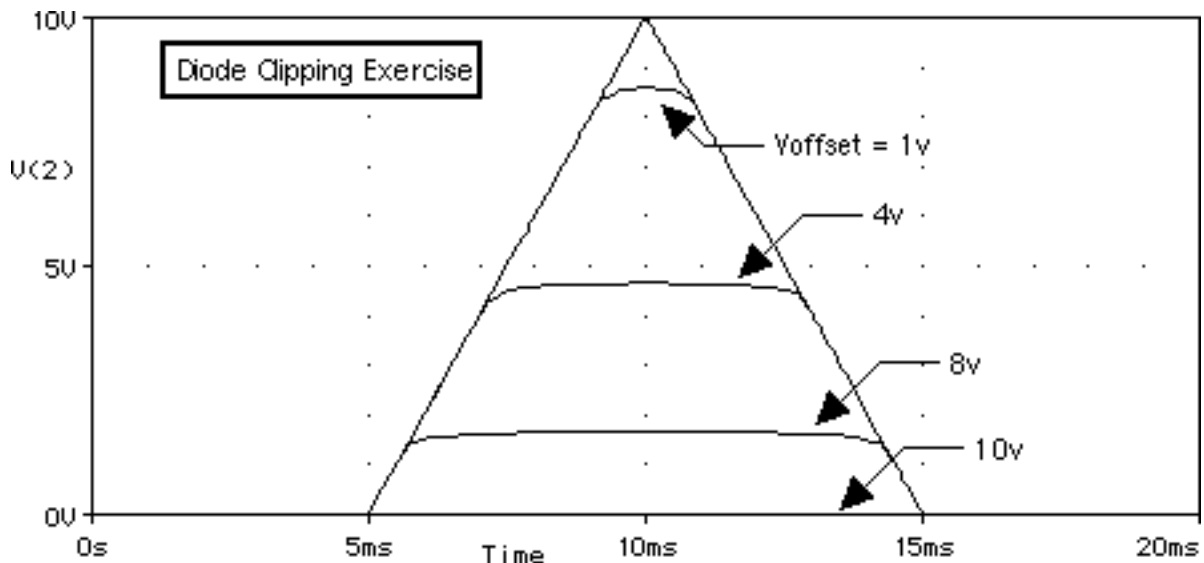


*Diode Clipping Illustration

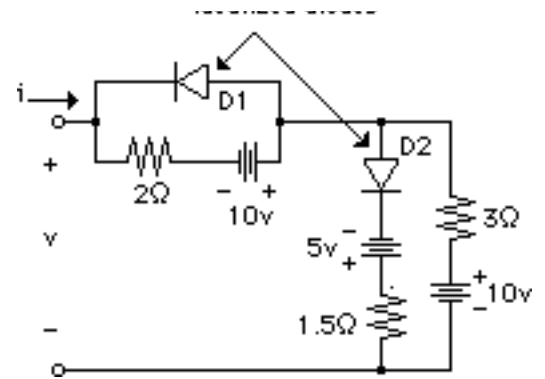
```

VS 1 0 PWL(0m 0 5m 0 10m 10
    + 15m 0 20m 0)
R1 1 2 1
D1 2 3 D1N004
VOFF 3 0 DC0
.MODEL D1N4004 D(Is=14.11n N=1.984
    + Rs=33.89m Ikf=94.81 Xti=3 Eg=1.11
    + Cjo=25.89p M=.44 Vj=.3245 Fc=.5
    + Bv=600 Ibv=10u Tt=5.7u)
.STEP VOFF LIST 1 4 8 10
.TRAN .2M 25M
.PROBE
.END
    
```

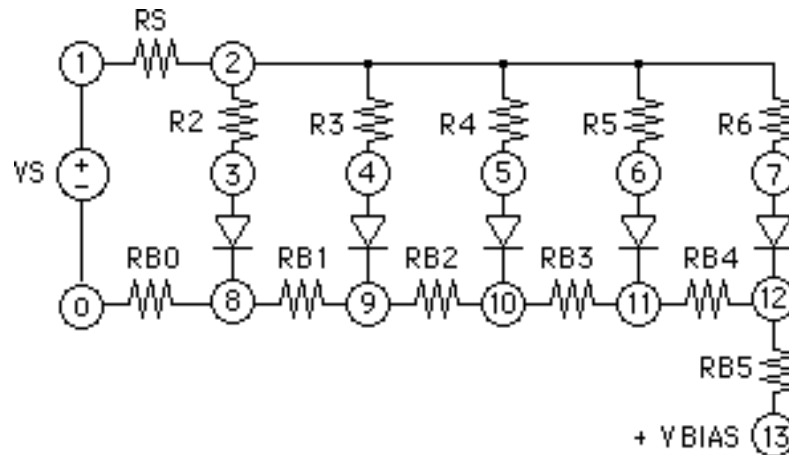
Note: The idealized diode model has a zero threshold ('knee') voltage while the 1N4004 threshold is 0.5 - 0.7 volts or so. Account for this in interpreting the plot.



(2) Analyze the circuit drawn to the right assuming D1 and D2 are idealized diodes. Then obtain a computer analysis of the circuit assuming first that D1 and D2 are 1N4004 diodes (see netlist above for PSpice model specification), and second that D1 and D2 are PSpice idealized diodes (use DIDEAL model). Note: this is the same circuit used in Problem 2, PWL ANALYSIS.



(3) The circuit diagram shows an experimental realization of the theoretical Square Root transfer function described previously; $V(2)^2$ VS for $0 \leq VS < 25$ volts. Note that the theoretical value will agree with an experimental value for $VS = 0$ volts. A second point of agreement can be realized by a calibration step; adjust VBIAS so that experimental and theoretical agreement is obtained for $VS = 25$ volts.



After verifying that the PSpice netlist following describes the circuit, compute the transfer characteristic. (Plot $V(2)^2$ vs. VS to judge the efficacy of the design; this should (ideally) plot as a line.)

*Square Root Characteristic

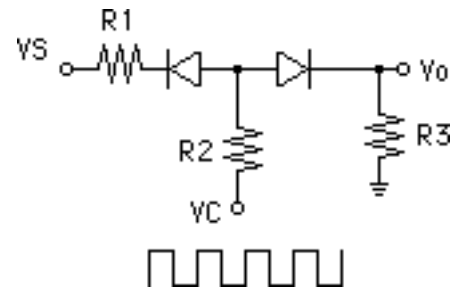
VS	1	0	DC	0V
RS	1	2		6.8K
R2	2	3		3.3K
R3	2	4		3.3K
R4	2	5		3.3K
R5	2	6		3.3K
R6	2	7		3.3K
D1	3	8		D1N4004
D2	4	9		D1N4004
D3	5	10		D1N4004
D4	6	11		D1N4004
D5	7	12		D1N4004
RB0	0	8		100
RB1	8	9		100
RB2	9	10		100
RB3	10	11		100
RB4	11	12		100
RB5	12	13		1K

VBIAS 13 0 11.5
 * Note: VBIAS chosen by trial and error
 * so that $V(2)^2$ is 5 volts when VS is
 * 25 volts.

```
.MODEL DIDEAL D(N=1U, RS=.001)
.MODEL D1N4004 D(Is=14.11n
+N=1.984 Rs=33.89m Ikf=94.81 Xti=3
+Eg=1.11 Cjo=25.89p M=.44 Vj=.3245
+Fc=.5 Bv=600 Ibv=10u Tt=5.7u)
.DC VS 0 25 0.1
.PROBE
.END
```

(4) The incremental resistance at a specified steady-state forward-biased operating point of a diode is the inverse of the slope at that operating point. Using the theoretical expression for the diode characteristic show that the incremental resistance at current I is $(NkT/q)/I$. Relate this calculation to the plot used before to determine the emission parameter.

(5) The large difference between forward- and reverse-bias conduction of a diode is the basis for the following (simplified) switching circuit illustration. A square-wave control voltage VC is applied as shown to the right. The back-to-back diodes both are reverse-biased, unless the voltage at the common node of the diodes is greater than VS. Using idealized diode approximations show that



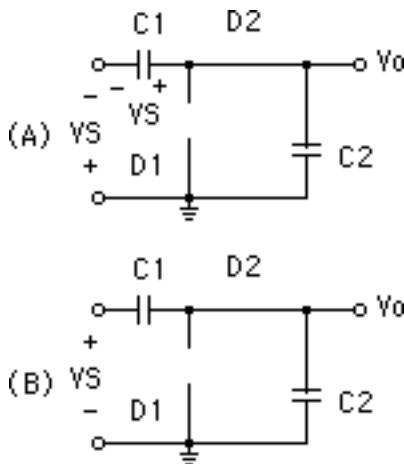
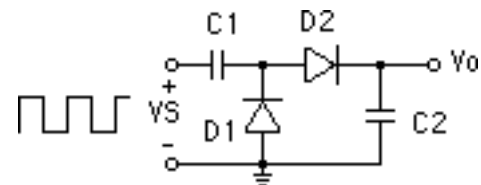
$$V_o = \frac{V_S}{1 + \frac{R_1}{R_2 || R_3}} + \frac{V_C}{1 + \frac{R_2}{R_1 || R_3}}$$

Assume VS is a triangular waveform of amplitude 4v, and VC is a square wave of amplitude 10v. Use R1=1K, R2=R3=47K. Use 1N4148 switching diodes rather than idealized diodes:

.MODEL D1N4148 D(Is=0.1p Rs=16 CJO=2p Tt=12n Bv=100 Ibv=0.1p)

Compute and plot Vo. Keeping in mind the back-to-back diode connection evaluate the influence of the threshold voltage for the diodes?

(6) The large difference between forward- and reverse-bias conduction of a diode is the basis for the (simplified) voltage doubler circuit shown. A bipolar signal (square-wave amplitude $\pm VS$ for simplicity) is applied as shown.



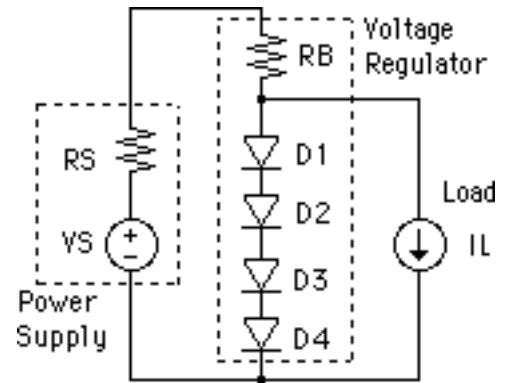
Assume, for simplicity, idealized diodes. During the negative peak (Fig.A) D1 is forward-biased and provides a path to charge C1 to VS. However D2 will be reverse-biased, isolating C2. During the following positive peak D1 will be reverse-biased, but D2 will be forward-biased, and part of the charge on C1 will transfer to C2.

Any charge transferred from C1 to C2 during one cycle will be restored on the next negative peak, and the process will repeat. Note that there is no path for C2 to discharge, and this circuit will continually pump charge into C2, continually increasing Vo. Equilibrium will be reached when C2 is finally charged (ideally) to 2VS (and C1 charges to VS) because then D2 will not be forward-biased on the positive peak.

Compute and plot the circuit performance for a $\pm 5v$ square wave with a 1 millisecond period. Use 1N4148 switching diodes (see above for .MODEL description). C1 = 1 μ F, C2 = 3 μ F.

(7) Diodes often are used as a kind of voltage reference in integrated circuits. Although there are other much more preferable means to serve the purpose used in this illustration nevertheless it is instructive to examine the principles involved.

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 which increases with increasing load current, so that the 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 diode 'tree'. The idea is to use the diode forward-bias property that large diode current changes involve exponentially smaller diode voltage changes.



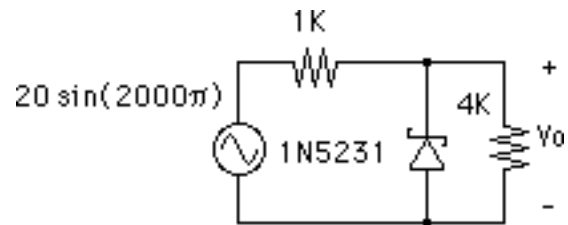
The circuit is designed to draw supply current greater than the maximum required by the load. This current then divides, part drawn off by the load and the remainder shunted through the diodes. As the load current demand changes the current division ratio changes. However provided the minimum current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

Compute the load voltage for the circuit as the load current varies from 0 to 20ma. Use $V_S = 10\text{v}$, $R_S = 200\ \Omega$, $R_B = 150\ \Omega$. Note that for the maximum 20 ma the terminal voltage will have dropped by 7 v from its no-load value. Use 1N4004 diodes

```
.MODEL D1N4004 D(Is=14.11n
+N=1.984 Rs=33.89m Ikf=94.81 Xti=3
+Eg=1.11 Cjo=25.89p M=.44 Vj=.3245
+Fc=.5 Bv=600 Ibv=10u Tt=5.7u)
```

Also compare the source current to the diode current.

(8) Use a PWL diode model to determine V_o for the circuit drawn below. Compare with a PSpice computation.



(9) Given a poorly regulated power supply whose Thevenin circuit consists of a 15 volt DC voltage source in series with $47\ \Omega$. The nominal 'test point' breakdown voltage of a 1N5231 Zener diode, read from the manufacturer's specifications, is 5.1 volts @ 20 milliamperes; the Zener resistance r_z is $17\ \Omega$.

Design a Zener shunt regulator as described above for load currents varying from a maximum of 15 ma down to a minimum of 2 milliamperes (nominal values). From the discussion note that this requires satisfying two conditions one of which sets an upper limit on the value of R_B allowed, with the other condition setting a lower limit. Select an appropriate resistance value, and compute the circuit performance using the PSpice diode model.

Note: use a current source to sweep load current.