INTERNATIONAL RECTIFIER

IOR

ZENER DIODES

HB-20B $3.00

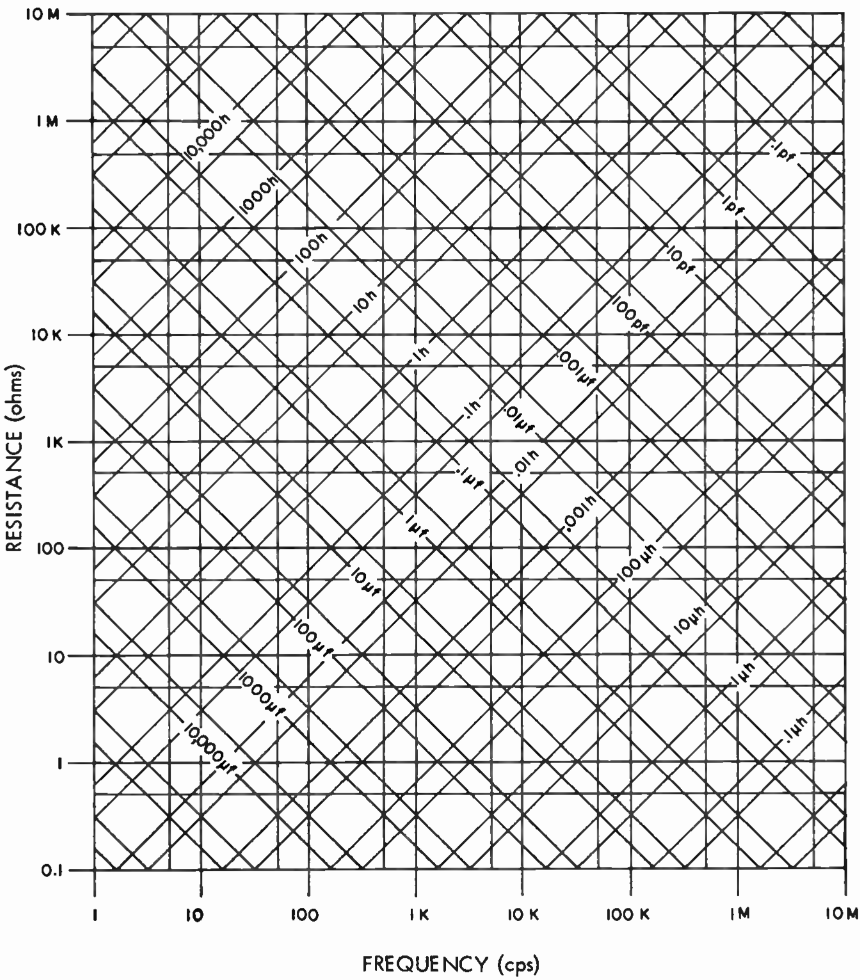
One of a series of technical handbooks published by:

INTERNATIONAL RECTIFIER IOR

Semiconductor Division, 233 Kansas Street, El Segundo, California 90245, Phone (213) 678-6281

APPENDIX C

REACTANCE CHART



INTERNATIONAL RECTIFIER

ZENER DIODES HANDBOOK

A theoretical discussion coupled with practical considerations and illustrated application data on the use of semiconductor voltage regulating devices.

Third Edition — Second Printing — July, 1969

CONTENTS

GROUP	0	I	II	III	IV	V	VI
Neg. Pos. Valence:	0	−7 +1	−6 +2	−5 +3	−4 +4	−3 +5	−2 +6
Formula Type		R_2O	RO	R_2O_3	RH_4 RO_2	RH_3 R_2O_5	RH_2 RO_3
Period 1:		1 H Hydrogen 1.00797					
Period 2:	2 He Helium 4.0026	3 Li Lithium 6.939	4 Be Beryllium 9.0122	5 B Boron 10.811	6 C Carbon 12.01115	7 N Nitrogen 14.0067	Oxy 15.9
Period 3:	10 Ne Neon 20.183	11 Na Sodium 22.9898	12 Mg Magnesium 24.312	13 Al Aluminum 26.9815	14 Si Silicon 28.086	15 P Phosphorus 30.9738	Sul 32.
Period 4:	18 Ar Argon 39.948	19 K Potassium 39.102 29 Cu Copper 63.54	20 Ca Calcium 40.08 30 Zn Zinc 65.37	21 Sc Scandium 44.956 31 Ga Gallium 69.72	22 Ti Titanium 47.90 32 Ge Germanium 72.59	23 V Vanadium 50.942 33 As Arsenic 74.9216	24 Cr Chromium 51.996 Selen 7
Period 5:	36 Kr Krypton 83.80	37 Rb Rubidium 85.47 47 Ag Silver 107.870	38 Sr Strontium 87.62 48 Cd Cadmium 112.40	39 Y Yttrium 88.905 49 In Indium 114.82	40 Zr Zirconium 91.22 50 Sn Tin 118.69	41 Nb(Columbium) Niobium 92.906 51 Sb Antimony 121.75	42 Mo Molybdenur 95.94 Tellur 127
Period 6:	54 Xe Xenon 131.30	55 Cs Cesium 132.905 79 Au Gold 196.967	56 Ba Barium 137.34 80 Hg Mercury 200.59	57 La Lanthanum 138.91 81 Tl Thallium 204.37	72 Hf Hafnium 178.49 82 Pb Lead 207.19	73 Ta Tantalum 180.948 83 Bi Bismuth 208.980	74 W(Wolfram) Tungsten 183.85 Polon (2
Period 7:	86 Rn Radon (222)	87 Fr Francium (223)	88 Ra Radium (226)	89 Ac Actinum (227)	104 ELEMENT 104 has been reported but the discovery has yet to be confirmed. (260)		

Periodic table of chemical elements

VII	VIII		
-1 +7			
RH R_2O_7			
9 F **Fluorine** 18.9984			
17 Cl **Chlorine** 35.453			
ganese 380 35 Br **Bromine** 79.909	26 Fe **Iron** 55.847	27 Co **Cobalt** 58.9332	28 Ni **Nickel** 58.71
netium 53 I **Iodine** 126.9044	44 Ru **Ruthenium** 101.07	45 Rh **Rhodium** 102.905	46 Pd **Palladium** 106.4
nium 2 85 At **Astatine** (210)	76 Os **Osmium** 190.2	77 Ir **Iridium** 192.2	78 Pt **Platinum** 195.09

ATOMIC WEIGHTS

The Atomic Weights are based on the isotope Carbon-12. In 1961, the International Union of Pure and Applied Chemistry officially approved use of the isotope C^{12} as the base for determining the relative atomic weights of elements. For radioactive elements, the atomic weight of the most stable known isotope is listed parenthetically.

ATOMIC NUMBERS

Atomic Numbers are used as the serial numbers of the elements. Atomic numbers are derived from the number of electrons in the orbital system of an atom of an element and the number of protons in the nucleus of the atom.

LANTHANIDE SERIES		ACTINIDE SERIES	
58 Ce 140.12	**Cerium**	90 Th 232.038	**Thorium**
59 Pr 140.907	**Praseodymium**	91 Pa (231)	**Protactinium**
60 Nd 144.24	**Neodymium**	92 U 238.03	**Uranium**
61 Pm (145)	**Promethium**	93 Np (237)	**Neptunium**
62 Sm 150.35	**Samarium**	94 Pu (242)	**Plutonium**
63 Eu 151.96	**Europium**	95 Am (243)	**Americium**
64 Gd 157.25	**Gadolinium**	96 Cm (247)	**Curium**
65 Tb 158.924	**Terbium**	97 Bk (249)	**Berkelium**
66 Dy 162.50	**Dysprosium**	98 Cf (251)	**Californium**
67 Ho 164.930	**Holmium**	99 Es (254)	**Einsteinium**
68 Er 167.26	**Erbium**	100 Fm (253)	**Fermium**
69 Tm 168.934	**Thulium**	101 Md (256)	**Mendelevium**
70 Yb 173.04	**Ytterbium**	102 No (254)	**Nobelium**
71 Lu 174.97	**Lutetium**	103 Lr (257)	**Lawrencium**

FOREWORD

While the term "semiconductors" is relatively new—one that has graduated from laboratory usage only in the last few years—semiconductors existed in the United States as far back as 1938. Some of the military equipment during and after World War II used selenium rectifiers as voltage regulators much the same as silicon Zener diodes are being employed today. The technical advances made in semiconductors since Shockley's theories on P-N junctions in 1949 brought about the development of silicon voltage regulators, among many other improved types of semiconductors.

Dr. Carl Zener was responsible for much of the initial groundwork on the breakdown mechanism theory. Thus, semiconductor devices designed to utilize the breakdown phenomenon came to be known as Zener diodes. With the appearance of new data, doubts arose as to the appropriateness of using the Zener theory to explain breakdown in silicon and germanium diodes. The avalanche mechanism, proposed by McKay and elaborated by others, is now believed to be more in accord with the facts. Since the term "Zener diode" has been firmly fixed in today's technical idiom, it has been used throughout this book to describe the device.

Significant achievements in semiconductor technology are followed inevitably by an acute need for thorough and accurate information on the subject. The engineering and publication staffs of International Rectifier continually prepare descriptive theoretical and application articles for publication in engineering and trade journals on each new development as it appears. The more pertinent papers on Zener theory and Zener diodes are listed, with acknowledgements, elsewhere in this handbook.

In the application section, emphasis is placed on illustrating some of the many applications for Zener diodes, rather than the design of the equipment with which they are associated. In many circuits, typical values and specific conditions are given. For more detailed design data, the reader can determine values by referring to the various textbooks on transistor circuitry and power supply design.

No attempt has been made here to cover all of the countless possible uses of this versatile component. Solutions to certain circuit problems through the use of Zener diodes are aimed at illustrating the capabilities of the devices, while familiarizing them to the design engineer . . . as well as stimulating imagination. It is hoped that this handbook will become an important addition to your reference library, and that its contents will prove valuable to you in the understanding and development of circuitry for the electronic equipment of tomorrow.

Introduction to Semiconductor Theory and Reverse Breakdown

The Zener diode (also called avalanche diode, voltage regulator, etc.) is an electronic solid-state device which utilizes certain features of the electrical characteristics of a rectifying junction. In some cases, the forward characteristics are employed – for example, in temperature compensation. Usually, it is the reverse characteristics and, in particular, the breakdown phenomena which are utilized.

In order to understand the underlying mechanism of reverse breakdown, it is desirable to review some of the fundamental principles of semiconductor theory. The discussion will be limited to the two most widely used semiconductors, germanium and silicon, and the reader interested in further details or in a more advanced treatment is referred to the bibliography given at the end of the chapter (cited literature references are indicated by superscript numbers).

Semiconductors

On the basis of electrical conductivity, solids can broadly be classified into metals, semiconductors, and insulators. To obtain a reasonably clear picture of semiconductors, we shall invade the field of solid-state physics and introduce some of the main concepts employed in discussing electrical semiconductivity.[1-8]

We begin with a consideration of electrons and their movement in semiconductors. It should be stated at the outset that while there are very many electrons in a piece of semiconductor material, most of them are strongly attached to their parent atoms and are not free to move and to conduct a current. Only a relatively few are loosely attached. These are the outermost electrons in an atom and the ones that usually determine its chemical valence.[9] For this reason, these "free" electrons are called *valence electrons*. Henceforth, we shall be dealing only with valence electrons.

Electric current is usually considered as a flow of electrons and one Ampere is, by definition, a flow of 6.24×10^{18} electrons per second. One of the most important concepts in semiconductor theory, however, is that an electric current can be ascribed not only to the flow of electrons but also to that of positively charged *holes*.

A positive hole, or more simply a hole, can be considered as the absence of an electron from a place where it normally would be found. In many cases, it is more convenient to describe a situation by means of holes and their movement; in other cases, the movement of electrons is more convenient; in still other cases, the movement of both holes and electrons must be taken into account.

The concept of holes and their movement is not so strange as it may appear at first glance. Consider, for instance, a crowded garage with only one vacant parking space, located at the rear (Fig. 1a). Suppose now that a car arrives at the entrance to the garage. The parking attendant will shift many of the cars towards the rear to accommodate the new car. In essence, however, he has brought the vacant space forwards towards the garage entrance (Fig. 1b). The hole has therefore moved, and it has moved in a direction opposite to the flow of the cars.

Applying the analogy of Fig. 1 to a semiconductor, the automobiles would be called electrons and the spaces would be called holes. Since in this case the electrons are in excess, they are the *majority carriers* (of current) and the holes are the *minority carriers*. A semiconductor of this type would be referred to as *N-type*, because most of the current is carried by negative charges (electrons).

When the opposite situation prevails and the holes are in excess, the semiconductor is referred to as *P-type*, because most of the current is carried by positive charges (holes). The electrons in this case become the minority carriers.

Electron and Hole Currents

Having introduced the concept of two types of charge carriers, let us consider what happens when a voltage is applied and they acquire a *drift velocity* to conduct a current. Consider a homogeneous piece of silicon at room temperature. Due to thermal agitation, the charge carriers (electrons, for simplicity) are moving randomly about, some with thermal velocities as high as 250,000 miles per hour! There is no observable current, however, because in the absence of an applied field there is no *net* drift of the electrons.*

When a voltage is applied across the silicon there will be a net drift of electrons toward the positive electrode and of holes toward the negative electrode. The current i_n due to n electrons per cm^3 moving with an average drift velocity v_n is simply given by

$$i_n = nev_n \qquad (2)$$

where e is the charge on an electron, 1.59×10^{-19} Coulombs.

If we consider the average electron drift velocity as being proportional to the field, E, we can write

$$v_n = \mu_n E \qquad (3)$$

where the constant of proportionality, μ_n, is called the *electron drift mobility*.

* Diffusion currents can, of course, flow when there is a gradient in the density of the carriers, and the direction of flow is from high to low concentration. The magnitude of the flow is given by Fick's First Law of Diffusion

$$f = -\frac{dn}{dx} D \qquad (1)$$

where n is the carrier density, x is the distance, and the constant of proportionality, D, is called the *diffusion coefficient*. As we shall see later, diffusion currents are very important in junctions of P and N material.

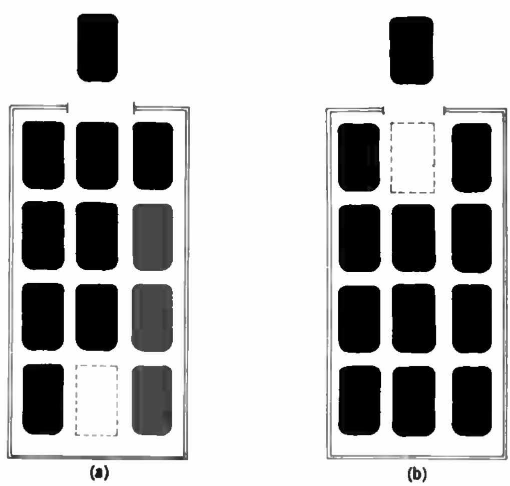


Fig. 1. Parking lot analogy to demonstrate movement of a "hole."

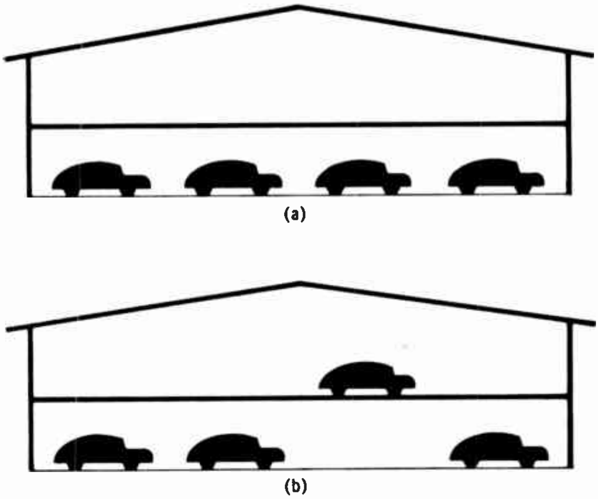


Fig. 2. Garage analogy to demonstrate energy levels. In (a), cars cannot move about. (After Shockley[3]).

Substituting Eq. (3) in (2) we find that the electron current is

$$i_n = ne\mu_n E \qquad (4)$$

Similar expressions apply to the holes, so that the hole current is given by

$$i_p = pe\mu_p E \qquad (5)$$

where p is the density of holes, and μ_p is the *hole drift mobility*.

The total current is given by the sum of the electron and hole currents, that is,

$$I = i_n + i_p = (ne\mu_n + pe\mu_p)E \qquad (6)$$

By analogy to Ohm's Law, the term in parenthesis is referred to as the *conductivity*, σ, (the reciprocal of the resistivity, ρ).

The motion of holes and of electrons should not be visualized as taking place in uninterrupted straight lines between the electrodes. Instead, the carriers move in a given direction for only a relatively short distance before colliding with the atoms of the crystal, to be deflected toward a different direction. Nevertheless, under the influence of the applied field there is a net flow toward the electrodes, and this manifests itself as a current.

When the applied electric field is very high (say 10^5 Volts/cm) a new effect takes place. What happens now is that the electron and hole carriers may gain enough kinetic energy during the time between collisions with the crystal lattice, that they are able to strip off electrons from the atoms. Every time that an electron is stripped off ("ionized"), a hole is left behind. We therefore speak of the creation of *electron-hole pairs*, and the process referred to here is called *ionization by collision*. The process is of importance, since it can lead to multiplication of carriers and a current avalanche. We shall discuss it further on under the section on reverse breakdown.

Energy Bands

Let us now interpret the behavior of electrons and holes in terms of energy band diagrams. Consider a garage with two stories; imagine further that the lower level is completely filled with cars and that the top level is completely empty (Fig. 2a). Obviously, there can be no movement of cars in either level.

Suppose now that we lift a car from the bottom to the top level (Fig. 2b).

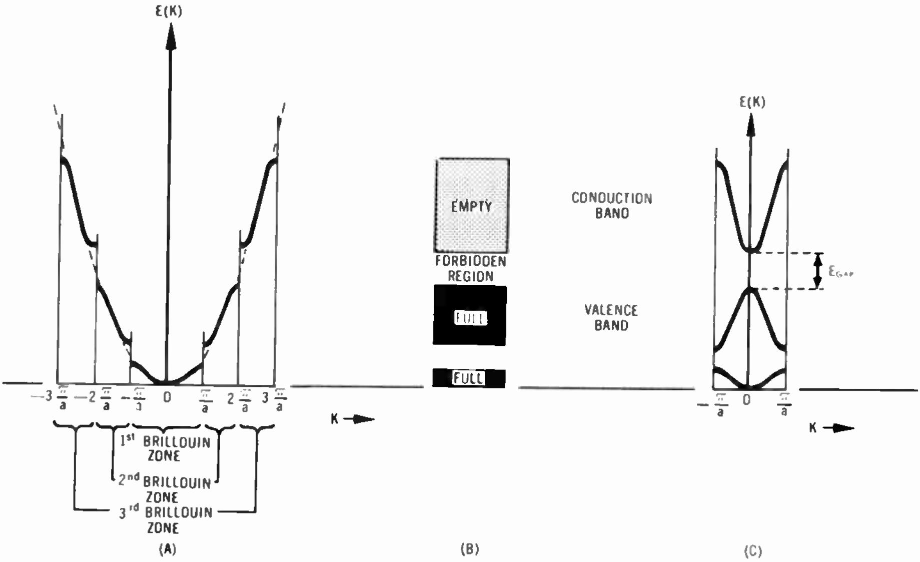


Fig. 3 (A) Energy of an electron as a function of wave number for the periodically varying field of a semiconductor lattice.[1] The gaps in the energy curve give rise to forbidden energy regions which are indicated schematically in (B). By translation of the energy curves into the first zone, (C) is obtained from (A).

This requires energy, hence the lone car on the top floor will have more potential energy than those on the first floor. Furthermore, it is now capable of being moved across the second floor level; likewise, the vacancy on the first floor level can also be moved about.

The same situation occurs in silicon. We think of broad levels, called *energy bands*, such as is shown in Fig. 3b. They can be considered the solid state equivalent of the familiar discrete energy levels of extranuclear electrons in isolated atoms. Due to the great number of atoms in a solid, each energy band consists of numerous individual energy levels, so close together as to be almost continuous.[1]

The region between energy bands is "forbidden" in the sense that these energy values are non-existent, and there are physical and mathematical reasons for this phenomenon, which forms part of what is known as the Band Theory of Solids.[1-8] The width of the forbidden region is called the *energy gap*, E_g, and is usually expressed in electron volts.* For silicon at room temperature, E_g is about 1.1 ev; for germanium it is about 0.7 ev.

Returning to Fig. 3b, the first empty band is called the *conduction band* and the one just below it is called the *valence band*. It is customary in semiconductor work to deal only with these two energy bands.

Intrinsic Conduction

In germanium and silicon at room temperature, the valence band is almost completely full of electrons. A few electrons, however, have acquired enough thermal energy to move upstairs (energy-wise) into the conduction band.† This process is referred to as *thermal generation* of electron-hole pairs (See Fig. 4).

The process of generation is being continually balanced by an equal and opposite process, known as *recombination*. The latter corresponds to a drop across the forbidden energy gap of an

* 1 electron volt = 1.6×10^{-19} joules of energy.

† Physically, this means that a certain proportion of the covalent bonds between adjacent atoms have been broken as a result of thermal agitation. That is, a certain amount of electrons (and corresponding holes) have been freed to conduct a current.

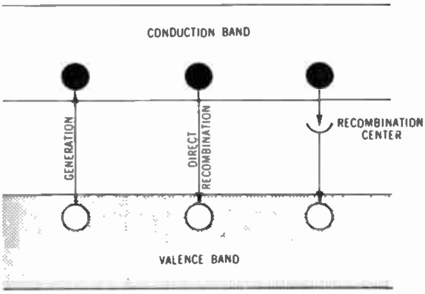


Fig. 4. Schematic representation of the generation and recombination of carriers.

electron from the conduction to the valence band.‡

The few electrons in the conduction band (and the corresponding holes in the valence band) allow only a very small current to flow. If we heat up the silicon, however, we notice a greater current. This is due to the fact that we are driving more electrons up into the conduction band and therefore creating more current carriers (n and p in Eq. 6). Contrary to what we notice in metals, therefore, silicon (over a certain range of temperature) has a *negative temperature coefficient of resistance.*

It can be shown that the product of the electron and hole concentrations in a given semiconductor is a function of temperature only and is given by

$$np = n_i^2 = AT^3 \exp\left[\frac{E_g}{kT}\right] \quad (7)$$

where A is a constant, T the absolute temperature, Eg is the energy of the forbidden gap at the temperature T, k is Boltzmann's constant, and n_i is known as the *intrinsic carrier concentration,* the concentration of holes or electrons in a pure crystal.

For germanium and silicon, n_i at room temperature is approximately 2.4×10^{13} and 1.5×10^{10} carriers per cm^3, respectively.[11] The practical significance of this is that it is much more difficult to prepare intrinsic silicon than intrinsic germanium, because the purity of the former has to be so much greater.

On the other hand, since the forbidden energy gap of silicon is greater than that of germanium, devices made from silicon can be operated at a higher temperature. As we shall see, the properties of devices made from silicon and germanium are largely determined and controlled by small amounts of doping impurities deliberately added to the pure semiconductor. At high temperatures, the thermally generated carriers are so numerous that they mask the effect of the carriers produced by the impurities, and the device ceases to operate satisfactorily.

Extrinsic Conduction

The type of conduction discussed so far, and illustrated by Fig. 5a, is known as *intrinsic conduction.* Another way of controlling the current is by the effect of very small amounts of impurity atoms introduced into the crystal lattice of the semiconductor. This type of conduction, known as impurity or *extrinsic*

‡ Physically, this means that an electron and a hole have met and the electron has fallen back into the bond between adjacent atoms. If the transition of the electron across the forbidden gap is a direct one, it is called *radiative recombination* because radiation is emitted (as in luminescence phenomena). More commonly, the transition occurs through an intermediate step, called a *recombination center.* Recombination centers arise from impurity atoms, dislocations, vacancies, and interstitials. They play a very important role in semiconductor devices because the net effect of recombination centers is to decrease the lifetime of excess carriers. An excellent review of recombination has been given by Bemski.[10]

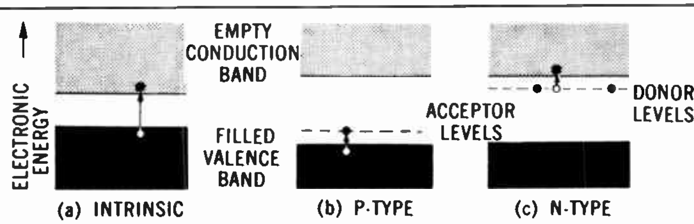


Fig. 5. Electronic energy diagrams for semiconductors. Black and white circles represent electrons and holes, respectively.

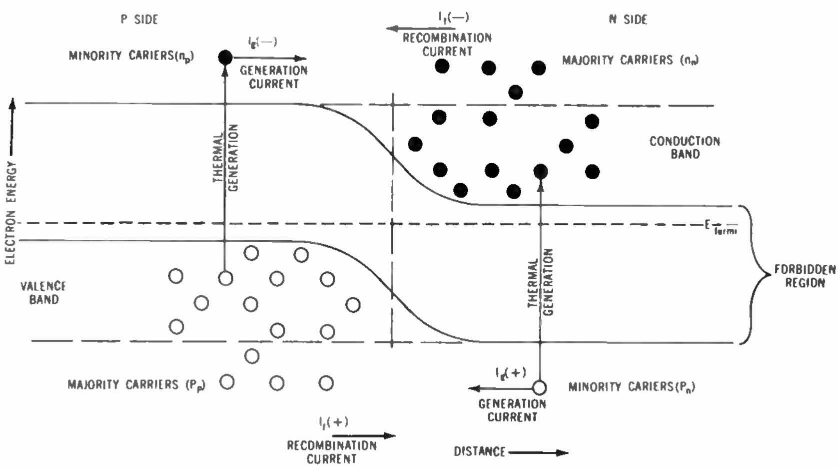


Fig. 6. Electron energy diagram for a P-N junction at thermal equilibrium and no applied bias. White and black circle represent holes and electrons, respectively.

conduction, makes use of intermediate, localized energy levels within the forbidden energy gap (Fig. 5, b and c).

If the impurity level lies just above the (bottom) valence band, we have what is known as an *acceptor level* and the semiconductor is called *P*-type (Fig. 5b). Consider, for instance, a sample of pure silicon doped with a very small amount of boron. Boron belongs to Group III of the Periodic Table of the Elements* and has only three valence electrons.[9] When a boron atom replaces a silicon atom in the crystal lattice it cannot supply the normal quota of four valence electrons that each silicon atom does. A positive hole carrier is therefore created around each boron atom. Since there is now an excess of holes, the current will be carried mainly by them.

Referring again to Fig. 5b, we see that the new energy level (indicated by broken lines to signify that it is not continuous, but confined to the individual boron atoms) can readily *accept* electrons from the nearby full valence band. The current is carried by holes left behind in the valence band when electrons move out. It can be visualized as a lateral movement of holes within the valence band, towards the negative electrode.

* Refer to the Periodic Table of the Chemical Elements on page iv.

Similarly, we can produce *N*-type silicon by doping with a pentavalent element such as antimony, which belongs to Group V of the Periodic Table. The antimony atoms donate five valence electrons instead of silicon's four, and the extra electrons occupy discrete energy levels, close to the upper conduction band (Fig. 5c). These levels are called *donor levels* because they readily donate electrons to the nearby band, where they are able to conduct a current.

Boron is called an accePtor and antimony a doNor. A useful mnemonic device is to recall that the *p* in acceptor and the *n* in donor are associated with positive and negative carriers, respectively.

P-N Junctions

The boundary between adjacent *P*- and *N*-material within the same single crystal is of great theoretical interest and of tremendous practical importance. Energy band diagrams can be calculated and drawn for *P-N* junctions, and the properties and characteristics of these junctions can generally be interpreted more easily by means of such diagrams.

Fig. 6 is a schematic diagram of the energy relationships of a *P-N* junction in, say, silicon. For simplicity, the hole concentration on the *P*-side is shown as approximately the same as the electron concentration on the *N*-side. Furthermore, the acceptor and donor levels have been omitted, but it should be understood that these levels exist and are responsible for the large number of holes in the (valence band of the) *P*-side and the large number of electrons in the (conduction band of the) *N*-side, respectively.

A study of Fig. 6 reveals that the *P*-side of the silicon is higher up on the electronic energy scale than the *N*-side. This indicates that, relative to the *N*-side, the *P*-side has a larger negative energy. At first sight, this appears contradictory but what happens is described below.

Imagine the *P* and *N* sides of the junction as though they were separated. Each side has no net charge because there is electrical neutrality. That is, in each side the total number of mobile carriers is balanced by an equal number of immobile ionized atoms.

Now, imagine the two sides as coming into contact. The *P*-side becomes negatively charged because the excess of positive holes that it possesses tends to diffuse over to the negative side (Eq. 1). This process leaves behind less positive holes than required for equilibrium, which is equivalent to making the *P*-side slightly negative.

Similarly, the excess electrons on the *N*-side tend to diffuse over to the *P*-side (according to Eq. 1), thereby leaving the *N*-side positively charged. Consequently, the *N*-side is shown on the diagram as being more positive (that is, having less electronic energy) than the *P*-side.

The diffusion of holes (majority carriers) away from the *P*-side, and of electrons (majority carriers) away from the *N*-side, eventually is balanced by the counteracting effect of electrical attraction of the charged surfaces for minority carriers. The latter, which have been thermally generated, tend to drift back across the *P-N* junction because there is a lower potential for them on the opposite side of the junction.*

We thus have two opposing effects in equilibrium and there is no net current: The forward, diffusion current is balanced by a reverse, generation current. This balance holds both for holes and electrons, so, in essence, we have four currents in equilibrium.

The forward, diffusion current of carriers is also called a recombination current, because it is determined by the recombination rate of holes and electrons. A hole, for instance, that leaves the *P*-region starts out as a majority carrier but when it reaches the *N*-region it automatically becomes a minority carrier. It then diffuses for an average *lifetime* τ_p before recombining with an electron and losing its identity.

In addition, the average distance the holes penetrate the *N*-region is called the *diffusion length*, L_p,† and it is related to the hole lifetime by the following relationship

$$L_p = \sqrt{D_p \tau_p} \qquad (8)$$

where D_p is our old friend, the hole diffusion constant (*Cf.* Eq. 1).

A similar situation holds for the movement of electrons into the *P*-region, and the diffusion length of electrons (in the *P*-region) is given by

$$L_n = \sqrt{D_n \tau_n} \qquad (9)$$

where D_n is the electron diffusion constant and τ_n is the electron lifetime.

Fig. 6 illustrates the situation for the *P-N* junction in equilibrium without any applied voltage. The overall current can be considered as being made up of two parts, one corresponding to drift and one resulting from diffusion. The drift effect is given by Eqs. 4 and 5. The diffusion effect is obtained from Eq. 1, where the particle flow can be converted to current by multiplying by the electronic charge, *e*, to yield the

* The electrons can be considered as sliding down the electronic potential hill. The holes, since they are oppositely charged, will easily slide up hill. A simple analogy is given by the motion of marbles and of air bubbles within a horizontal tube filled with water: Tipping the tube causes the marbles to roll downwards and the bubbles to float upwards. Similarly, if the cars in Fig. 1 had no brakes, tilting the ground would cause the cars to roll downhill and the spaces to move uphill in the opposite direction.

† Actually L_p is the average distance required for the hole concentration, P_n, to drop to $\frac{1}{e}$ of its original value.

following diffusion currents:

$$i_p = -e\, D_p \frac{dp}{dx} \qquad (10)$$

$$i_n = +e\, D_n \frac{dn}{dx} \qquad (11)$$

where i_p is the hole diffusion current and i_n is the electron diffusion current. The electron current has a positive sign because the negative electronic charge yields an equivalent current opposite in direction to the electron flow.

Combining Eqs. 4, 5, 10, and 11 we find that the overall hole and electron currents are given by

$$i_p = pe\mu_p E - e\, D_p \frac{dp}{dx} \qquad (12)$$

$$i_n = ne\mu_n E + e\, D_n \frac{dn}{dx} \qquad (13)$$

The electrostatic field E arises, as discussed above, from the charging action on opposite sides of the *P-N* junction as a result of diffusion processes.*

Let us now consider what takes place when an outside voltage V is applied to the *P-N* junction. In the forward direction, that is making the *P*-side positive, the situation is illustrated by Fig. 7. It can be seen that by making the *P*-side more positive with respect to the *N*-side, we have depressed the energy bands on the *P*-side and raised them on the *N*-side. Consequently, the potential barrier between the two sides has been decreased (by the amount V) and majority carriers are able to surmount the junction barrier. As shown in the figure, a large supply of majority carriers is available for diffusion, resulting in a large forward current.

The net forward current is actually determined, however, by the concentration of the minority carriers. Consider the diffusion of electrons into the *P*-region, where they become excess minority carriers. According to Eq. 11, the diffusion current is strongly dependent on the *gradient* of the electron density. In order to maintain a steep concentration gradient, the excess electrons which penetrate (are injected into) the *P*-region must be disposed of. This is accomplished through recombination.

* The charging action also gives rise to an electrostatic potential, φ_0, commonly known as the built-in potential. It is related to the field by the following equation

$$E = -\frac{d\varphi_0}{dx} \qquad (14)$$

Silicon has typical built-in potentials of about 0.5 to 0.8 Volts.

It can also be shown that φ_0 is related to the intrinsic carrier density n_i by the following expression:

$$\varphi_0 = \frac{kT}{e} \ln \frac{n_n p_p}{n_i^2} \qquad (15)$$

where the subscripts n and p refer to the concentrations of the carriers in the *N*- and *P*-areas outside the transition region of the junction. Since n_i increases very rapidly with temperature (Eq. 7), the built-in potentials tend to drop to zero as the temperature is raised.

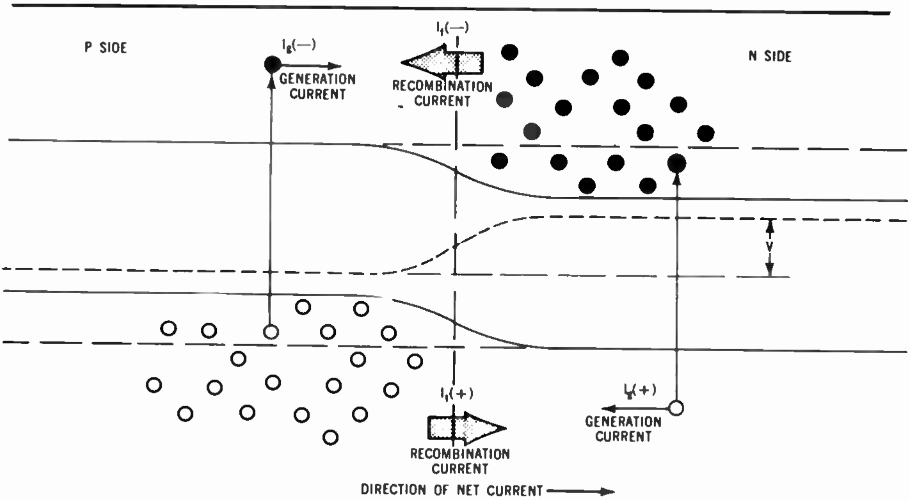


Fig. 7. Electron energy diagram for a P-N junction biased in the forward direction (P-side positive).

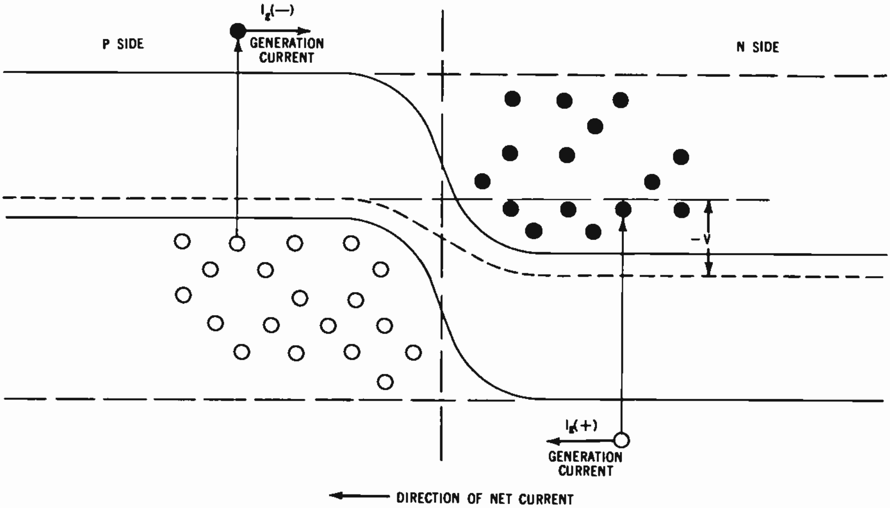


Fig. 8. Electron energy diagram for a P-N junction biased in the reverse direction (P-side negative).

On the average, each electron diffuses a length L_n into the *P*-region and survives for a lifetime τ_n. Then it recombines with a hole (which is the majority carrier) and both the electron and the hole are withdrawn from circulation. In order to preserve electrical neutrality, another hole (from the positive contact electrode) automatically replaces the hole that recombined. This last process leads to a flow of majority carriers (holes in the *P*-region) and accounts for the greater part of the current in the *P*-side (Fig. 9).

When a voltage is applied in the reverse direction (*P*-side made negative), the situation is illustrated by Fig. 8. It can be seen that the difference in height between the energy bands of the *P*- and *N*-sides has increased to the point where there is little or no diffusion of majority carriers across the junction. The only current now flowing is a very small one, due to thermally generated minority carriers which are attracted across the junction by the electrostatic field. This current gives rise to the reverse leakage current in a rectifier.

The current is again determined by minority carriers. Only those generated (on an average) within a diffusion length can reach the junction and drift over to the electrostatically charged opposite side. As shown by Eq. 16, the current is essentially independent of voltage and saturates at I_s, the value given by Eq. 17. For an ideal *P-N* junction operated under certain simplifying conditions, Shockley[12] has derived a very useful expression which enables the *I-V* characteristics to be predicted in terms of basic semiconductor properties. This expression is

$$I = I_n + I_p = I_s \left[\exp\left(\frac{eV}{kT}\right) - 1 \right] \qquad (16)$$

where I = junction current—Amp/cm^2
V = junction voltage—Volts

and the saturation current is

$$I_s = e \left[\frac{D_p p_n}{L_p} + \frac{D_n n_p}{L_n} \right] \qquad (17)$$

where, respectively, D_p and D_n are the hole and electron diffusion coefficients; p_n and n_p the concentration of holes and electrons in the *N*- and *P*-regions; and L_p and L_n the hole and electron diffusion lengths.

Eq. 16 indicates that the forward current (V positive) increases exponentially with voltage, whereas the reverse current (V negative) saturates at $-I_s$, corresponding to the thermal generation of minority carriers within a diffusion length of each side of the junction.

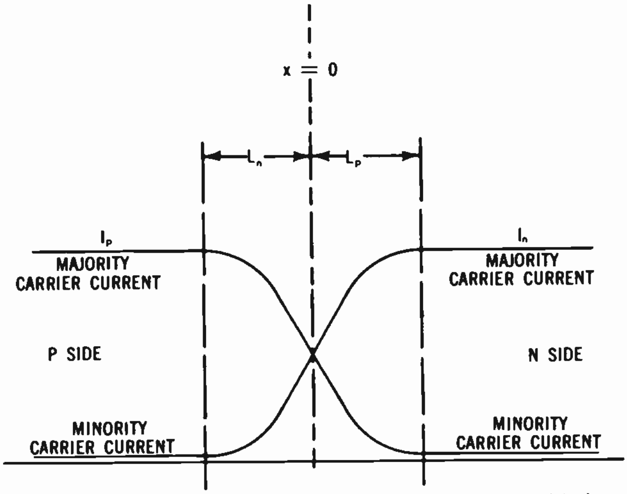


Fig. 9. Schematic representation of the distribution of electrons and hole currents across a P-N junction.

At low voltages, the hole current in the forward direction depends directly on the hole diffusion coefficient and the hole density in the *N*-material, but inversely as the diffusion length. A similar reasoning applies to the electron current.

Fig. 9 indicates how the electron and hole currents vary in the semiconductor with a biased *P-N* junction: Within the transition region of the junction, the currents are controlled by minority carriers; outside of that region, the drift of majority carriers predominates.

The ideal theory as given by Eq. 16 agrees fairly well with experimental data on germanium junctions operated at low levels and at room temperature. It does not apply to silicon at room temperature or to germanium at low temperature. Consequently, a number of modifications and extensions have been made to the simple theory.[13, 14]

In the case of silicon *P-N* junctions, it is found that the reverse leakage current is not as small or as constant as Eq. 17 predicts, and it is necessary to consider the effects of spontaneous carrier generation and recombination within the barrier region.*

Likewise, departures from Eq. 16 occur when the junctions are operated at high forward current densities, and modifications to the simplified theory are again required. The analysis becomes fairly difficult since non-linear differential equations are involved and various approximations have to be made.[13]

Reverse Breakdown

For our present discussion, the most important region in the voltage-current characteristics of the silicon rectifier is that of high reverse bias, leading to *reverse breakdown* and its utilization in numerous devices and applications. It is beyond the scope of this article to present a detailed discussion of the breakdown mechanism, but an account of some of the present-day thoughts on the matter is desirable, to provide a better understanding of Zener diodes.

Fig. 10 illustrates schematically the current-voltage relationships in a reverse-biased silicon diode. At low voltages, only a very small current flows, which, as we have seen, is due to the drift of thermally generated minority carriers. In an ideal diode, this current would saturate at a relatively low value, given by Eq. 17.

In practice, however, the current is

* This has been shown to be theoretically possible.[15] Since the generation current is dependent on the volume of semiconductor material in which generation occurs, and since the junction barrier width increases with applied reverse bias, so too does the reverse current of silicon diodes increase slightly with voltage.

not completely voltage-independent. The reasons for departure from the ideal case are attributed to surface effects, channels, inversion layers, and gross defects in the body of the junction.[13-17] Some recent data[18] appears to indicate that in silicon with short lifetime, bulk recombination centers are responsible for the large space-charge generated currents; as the lifetime is increased, however, the surface leakage currents become predominant.

Looking again at Fig. 10, we notice that as the reverse voltage is increased the leakage current remains essentially constant until a certain critical voltage, the *breakdown voltage*, is reached, at which point the current increases by many orders of magnitude. This is the breakdown region, and the current is known as the Zener current or the avalanche current, as explained further on below.

The breakdown point is of great interest not only in theory but also in practice, since the reverse voltage which can be imposed on rectifiers and other semiconductor devices is determined by it. In conventional diodes and rectifiers, it is desirable, if not mandatory, to operate considerably below the breakdown point; Zener or avalanche diodes, on the other hand, are designed to operate at the breakdown point.

Zener Breakdown

Just what causes the very rapid rise in current at the breakdown point? For

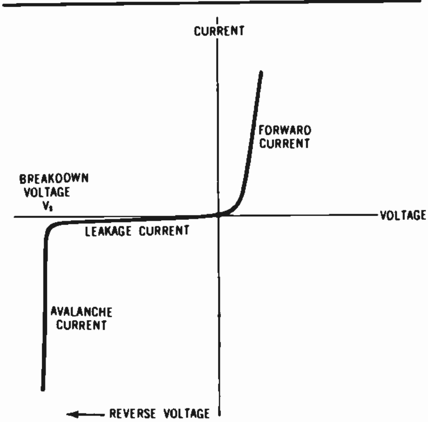


Fig. 10. Current-voltage characteristics of a typical P-N junction.

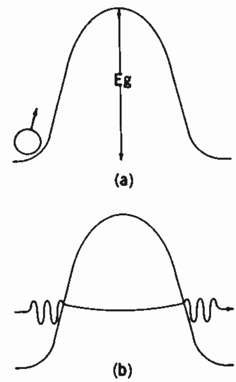


Fig. 11. (a) An electron with energy less than Eg cannot cross over the potential barrier, according to classical mechanics; (b) In quantum mechanics, the electron is considered as wave motion and there is a finite probability of it "tunneling" through the potential barrier.

some time this question was an interesting and challenging one, but numerous experiments and theoretical analyses which have been made in recent years have added considerably to our knowledge of the mechanism of reverse breakdown.

Two types of breakdown have been considered, the Zener type and the avalanche type. In the early days of semiconductor device development it was believed[19] that the Zener mechanism was responsible for breakdown in germanium and silicon diodes. Devices, therefore, designed to utilize the breakdown phenomenon came to be known as Zener diodes. At the present, it is believed that both mechanisms may be operative, with the avalanche process being the predominant one, especially at reverse voltages above about 6 Volts. Consequently, the term avalanche diode may be more appropriate but, rightly or wrongly, the designation Zener diode is likely to remain.

The Zener effect, so-called because it is based on a theory originally developed (for dielectrics) by C. Zener,[20] is actually a case of internal field emission. Under very intense fields (of the order of 3×10^5 Volts/cm), electrons can be made to traverse the forbidden energy gap by a process known as quantum-mechanical tunneling.*

If the energy gap is likened unto a hill (potential barrier), classical mechanics tells us that electrons on one side of the hill cannot reach the other side unless they have sufficient energy to climb over the hill.

In the new, quantum mechanics such a limitation does not exist, however, and there is a statistical probability of finding electrons on the other side (Fig. 11). This probability, P, is given by an expression of the following type

$$\log P = \text{constant}\ \frac{E_g^{3/2}}{E} \qquad (18)$$

where E_g is the energy gap and E is the applied field. When E is large, the probability approaches unity. Further calculations convert the probability into an actual current which is a function of the applied high electric field.

It is also possible to obtain equations relating the breakdown voltage to the resistivity of the semiconductor on each side of the junction.

In the case of abrupt silicon *P-N* junctions, the Zener breakdown voltage is given by

$$V_z = 39\rho_n + 8\rho_p \qquad (19)$$

where ρ_n and ρ_p are the resistivities in the N and P sides, respectively.

This equation, however, does not agree with experimental facts. For instance, Zener diodes made with 0.1 Ohm-cm *N*-type silicon (*P*-side resistivity negligible) have a Zener voltage of about 15 Volts instead of the 3.9 predicted by Eq. 19.

Avalanche Breakdown

As experimental data appeared indicating lack of agreement with the Zener theory, doubts arose as to its validity in explaining breakdown in silicon and germanium diodes. The avalanche mechanism proposed by McKay[22, 23] and elaborated by Wolff,[24] is now believed to be more in accord with the facts. It is similar to the process which occurs in an electrical discharge in gases, the theory of which was developed by Thompson.[25]

It will be recalled from our discussion of the reverse-biased *P-N* junction that the reverse current is due to the drift of minority carriers (holes in the *N*-side and electrons in the *P*-side) across the junction. In the avalanche process we visualize these carriers as being accelerated by the increasing voltage to higher and higher velocities. Eventually, they acquire sufficient energy** to be able to strip off (ionize) bound valence electrons from the silicon atoms.

It should be noted that when an electron is knocked off from the parent atom a positive hole is automatically created. Every ionizing carrier (hole or electron) therefore creates *two* additional carriers. These new carriers are now, in turn, capable of being accelerated by the high field and of creating additional electron-hole pairs! One can readily see that this cumulative process is one of rapid multiplication of carriers, leading to a tremendously rapid increase in reverse current.

An important parameter in the avalanche theory is the *ionization coefficient*, α, which can be considered as the number of ionizing collisions a given carrier makes in one cm path length. In other words, α is the number of electron-hole pairs created by a single carrier (electron or hole) as it travels one cm under the applied electric field.

By means of suitable experiments and making certain assumptions (for instance, that α is the same for electrons as for holes) it is possible to determine values of α as a function of the applied field. This has been done[22] for different types of junctions, and the values obtained range from about 800 at an electric field strength of 200 kV/cm (across the *P-N* junction), to about 60,000 at a field strength of 500 kV/cm. Above these voltages, the ionization coefficient appears to saturate.

The actual determination of α is done through measurements of the *multiplication ratio*, M, which is the ratio of

* This tunneling process appears to take place in the recently announced diodes developed by Esaki[21] and which have come to be known in this country as "tunnel diodes."

** The critical energy, or the "threshold for pair-production," by electrons in silicon is about 2.25 electron volts.[26] See also Refs. 24 and 28.

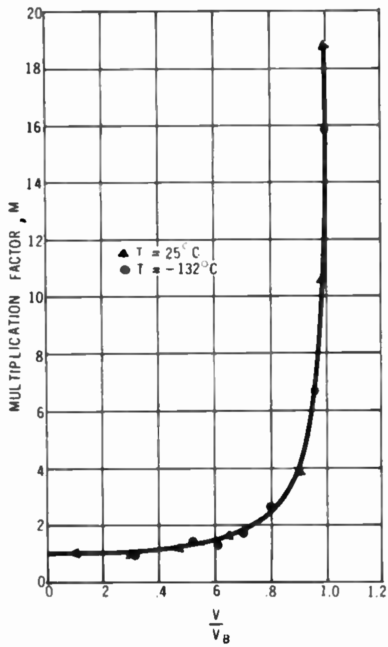


Fig. 12. Multiplication curves for a linear-gradient junction with the voltage scale normalized to the breakdown voltage for different temperatures. (After McKay[23]).

the current entering the junction. Multiplication, as we have seen, occurs even before we reach the breakdown voltage. By illuminating one side of the junction or bombarding it with alpha particles, one can inject carriers into it at voltages below the breakdown voltage and determine M from a study of the V-I characteristics.[22] A typical multiplication curve[23] is shown in Fig. 12.

It can be shown that the multiplication ratio M is related to the ionization coefficient α by means of the following expression

$$1-\frac{1}{M}=\int_{o}^{w}\alpha(E)dx \qquad (20)$$

where w is the width of the barrier. Solutions to the above equation can be obtained for different types of junctions. Thus, for a silicon step junction (alloy type) one obtains

$$\alpha=1.52\times10^{-7}N_I\frac{d\left(1-\frac{1}{M}\right)}{dE_m} \qquad (21)$$

where E_m is the maximum field, and N_I is the net impurity density (donors minus acceptors).

Values of the parameter α obtained from multiplication experiments agree quite well with those predicted theoretically according to the theory developed by Wolff.[24]

The above discussion assumed that α was the same for both electrons and holes. Subsequent experiments have indicated that this is not actually the case,[27-29] and the mathematical analysis becomes more involved. Interesting is the fact that in silicon the ionization coefficient for electrons is larger than for holes, whereas in germanium the opposite is true.

By using more refined methods for measuring the multiplication as a function of bias, Chynoweth[26, 29] discovered that α can be expressed quite simply as an exponential function of the field, E, or

$$\alpha=a\exp\left(-\frac{b}{E}\right) \qquad (22)$$

where a and b are constants. Since the above equation is not indicated by Wolff's theory,[24] it is suggested that a fresh theoretical approach is necessary. Quite recently, Maserjian[40] has assumed the above exponential dependence for α to be general and has derived expressions which predict avalanche breakdown for most P-N structures.

In alloyed type diodes, the breakdown voltage, V_B, can be adjusted by control of the resistivity (net impurity density). For a given resistivity, N-type silicon and germanium breakdown at a higher voltage than P-type. This is due to the fact that in these materials the electron mobility is greater than the hole mobility. In P-material the minority carrier current is by electrons and V_B will not have to be as high to reach avalanche conditions.

Although the avalanche theory is believed to be the predominant process, there is evidence that the Zener mechanism occurs at low voltages and in thin junctions.[22, 31] In silicon diodes with breakdown voltage below about six Volts, the breakdown point decreases with rise in temperature. This negative temperature coefficient is consistent with the decrease of the forbidden energy gap with temperature and a consequent lower Zener field.

On the other hand, silicon junctions with breakdown above about six Volts have a positive temperature coefficient, because the mobility of the carriers decreases with rise in temperature and higher voltages are required to produce ionization by collision.

Although theory would predict a "sharp" breakdown, in practice one often finds a rounding of the current-voltage characteristics, or what is known as a "soft" breakdown or a "soft knee," particularly at low voltages. The causative factors are not yet completely understood, but they have been ascribed to surface conditions, bulk effects,[32] and thermal effects.[33]

It has also been suggested that the effect of temperature rise due to current flow must be considered in order to explain the shape of the *V-I* curve in the avalanche region.[34] In addition, the *V-I* curve itself is not smooth but has a number of discontinuities, which are believed due to the existence of localized breakdown regions, called *microplasmas*.[36]

Apparently, the actual breakdown occurs only in isolated discrete spots, where the current density is obviously very high. The emission of visible light has been observed[35-37] in diodes operating in the avalanche region, and it is believed that the light originates in these spots which may be due to dislocations.[38] The effect is best observed in silicon solar cells, where the junction lies parallel and very close to the surface. Photons with energy as high as

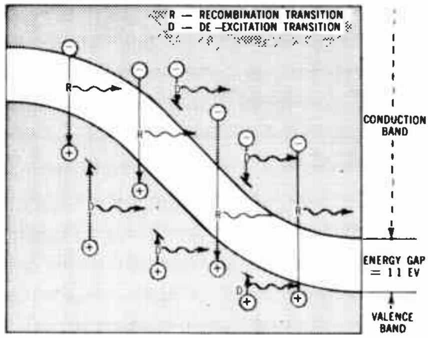


Fig. 13. Schematic representation of the two types of processes responsible for light emission. (After Chynoweth[39]).

3.2 electron volts (blue light) have been detected.

The spectrum of the light emitted by junctions in which avalanche is taking place indicates that two concurrent mechanisms are probably responsible for the light emission:[38] (a) *de-excitation radiation* — where carriers lose energy (and emit photons with energies up to the threshold value of 2.3 ev), but remain within the same energy band; (b) *recombination radiation* — where highly energetic electrons in the conduction band drop into the valence band and recombine with a hole (emitting photons of energies above 1.1 ev, the energy gap of silicon). Fig. 13 is a schematic representation of the two processes.[39]

Considerable research is still going on pertaining to the fascinating and important phenomenon of reverse breakdown. These efforts are leading to a greater understanding of the behavior of electrons and holes in semiconductors, and to the design and development of improved semiconductor devices.

REFERENCES

1. C. A. Escoffery, "First principles of semiconductors," *Elec. Eng.* 76, 142 (1957); Reprinted in *Engineering Handbook*, published by the International Rectifier Corporation, El Segundo, California, 1959.
2. Ruth F. Schwarz, "Introduction to semiconductor theory," *Elec. Mfg.* 63, 107 (1959).
3. W. Shockley, *Electrons and Holes in Semiconductors*, (D. Van Nostrand Co., Inc., New York, 1950).
4. G. Goudet and C. Meuleau, (translated by G. King), *Semiconductors: Their Theory and Practice* (Macdonald and Evans Ltd., London, 1957).
5. E. Spenke, (translated by Jenny *et al*), *Electronic Semiconductors* (McGraw-Hill Book Co., Inc., New York, 1958).
6. A. J. Dekker, *Solid State Physics* (Prentice-Hall, Englewood Cliffs, N. J., 1957).
7. W. C. Dunlap, Jr., *An Introduction to Semiconductors* (John Wiley and Sons, Inc., New York, 1957).
8. N. B. Hannay, ed., *Semiconductors* (Reinhold Pub. Corp., New York, 1959).

9. C. A. Escoffery, "New frontiers for semiconductors," *Electronics* 32, 43 (1959).

10. G. Bemski, "Recombination in semiconductors," *Proc. I.R.E.* 46, 990 (1958).

11. E. M. Conwell, "Properties of silicon and germanium: II," *Proc. I.R.E.* 46, 1281 (1958).

12. W. Shockley, "The theory of *p-n* junctions in semiconductor and *p-n* junction transistors," *Bell Sys. Tech. J.* 28, 435 (1949).

13. J. L. Moll, "The evolution of the theory for the voltage-current characteristic of *p-n* junctions," *Proc. I.R.E.* 46, 1076 (1958).

14. H. W. Henkels, "Germanium and silicon rectifiers," *Proc. I.R.E.* 46, 1086 (1958).

15. W. Shockley and W. T. Read, Jr., "Statistics of recombinations of holes and electrons," *Phys. Rev.* 87, 835 (1952).

16. For reverse current due to surface defects, see:
   (a) A. C. McWhorter and R. H. Kingston, "Channels and excess reverse current in grown germanium *p-n* junction diodes," *Proc. I.R.E.* 42, 1376 (1954).
   (b) A. R. F. Plummer, "Observations on the growth of excess current in germanium *p-n* junctions," *Proc. Phys. Soc. B.* 69, 539 (1956).
   (c) M. Cutler and H. M. Bath, "Surface leakage current in silicon fused-junction diodes," *Proc. I.R.E.* 45, 39 (1957).
   (d) W. T. Eriksen, H. Satz, and G. A. deMars, "Excess surface currents in germanium and silicon diodes," *J. Appl. Phys.* 28, 133 (1957).
   (e) Reference 18.

17. For reverse current due to bulk effects see:
   (a) E. M. Pell, "Reverse current and carrier lifetime as a function of temperature in germanium junction diodes," *J. Appl. Phys.* 26, 658 (1955).
   (b) E. M. Pell and G. M. Roe "Reverse current and carrier lifetime as a function of temperature in silicon junction diodes," *J. Appl. Phys.* 27, 768 (1956).
   (c) C. T. Sah, R. N. Noyce, and W. Shockley, "Carrier generation and recombination in *p-n* junctions and *p-n* junction characteristics," *Proc. I.R.E.* 45, 1228 (1957).
   (d) H. Kleinknecht and K. Seiler, "Einkristalle und pn shichtkristalle aus silizium," *Zeits. f. Physik* 139, 599 (1957).
   (e) Reference 18.

18. D. J. Sandiford, "Heat treatment centers and bulk currents in silicon *p-n* junctions," *J. Appl. Phys.* 30, 1981 (1959).

19. K. B. McAfee, E. J. Ryder, W. Shockley, and M. Sparks, "Observations of Zener current in germanium *p-n* junctions," *Phys. Rev.* 83, 650 (1951).

20. C. Zener, "A theory of electrical breakdown of solid dielectrics," *Proc. Roy. Soc.* (London) 145A, 523 (1934).

21. L. Esaki, "New phenomenon in narrow germanium *p-n junctions*," *Phys. Rev.* 109, 603 (1958).

22. K. G. McKay and K. B. McAfee, "Electron multiplication in silicon and germanium," *Phys. Rev.* 91, 1079 (1953).

23. K. G. McKay, "Avalanche breakdown in silicon," *Phys. Rev.* 94, 877 (1954).

24. P. A. Wolff, "Theory of electron multiplication in silicon and germanium," *Phys. Rev.* 95, 1415 (1954).

25. L. B. Loeb, *Fundamental Processes of Electrical Discharge in Gases* (John Wiley and Sons, Inc., New York; 1939), p. 372 ff.

26. A. G. Chynoweth and K. G. McKay, "Threshold energy for electron-hole pair production by electrons in silicon," *Phys. Rev.* 108, 29 (1957).

27. S. L. Miller, "Avalanche breakdown in germanium," *Phys. Rev.* 99, 1234 (1955).

28. S. L. Miller, "Ionization rates for holes and electrons in silicon," *Phys. Rev.* 105, 1246 (1957).

29. A. G. Chynoweth, "Ionization rates for electrons and holes in silicon," *Phys. Rev.* 109, 1537 (1958).

30. M. B. Prince, "Drift mobilities in semiconductors, I. Germanium," *Phys. Rev.* 92, 681 (1953); M. B. Prince, "Drift mobilities in semiconductors, II - Silicon," *Phys. Rev.* 93, 1204 (1954); G. W. Ludwig and R. L. Watters, "Drift and conductivity mobility in silicon," *Phys. Rev.* 101, 1699 (1956).

31. A. G. Chynoweth and K. G. McKay, "Internal field emission in silicon *p-n* junctions," *Phys. Rev.* 106, 418 (1957).

32. E. M. Pell, "Influence of electric field in diffusion region upon breakdown in germanium *p-n* junctions," *J. Appl. Phys.* 28, 459 (1957).

33. A. W. Matz, "Thermal turnover in germanium *p-n* junctions," *Proc. I.E.E.* (B) 104, 555 (1957).

34. B. Senitzky and P. D. Radin, "Effect of internal heating on the breakdown characteristics of silicon *p-n* junctions," *J. Appl. Phys.* 30, 1945 (1959).

35. R. Newman, "Visible light from a silicon *p-n* junction," *Phys. Rev.* 100, 700 (1955).

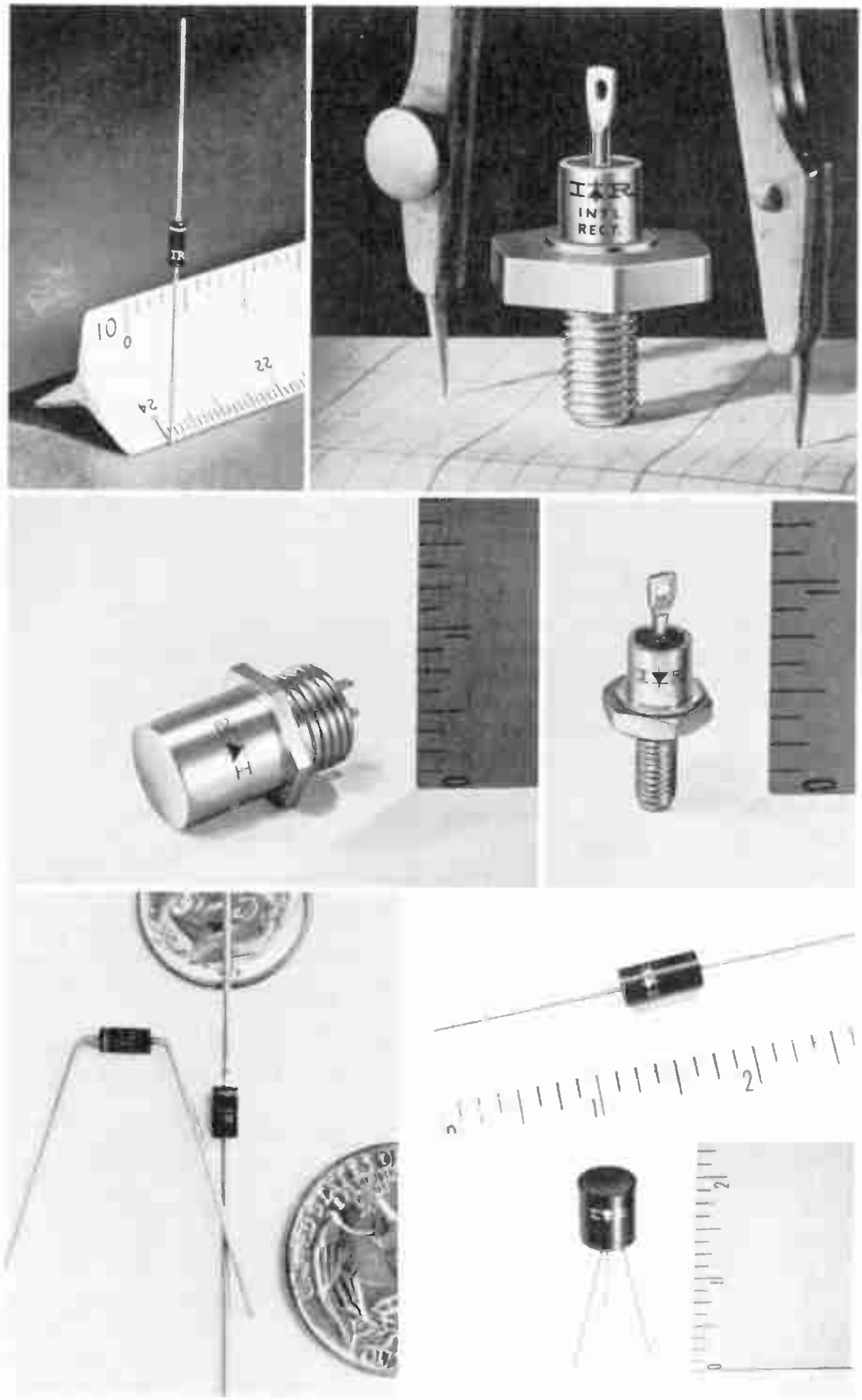
36. A. G. Chynoweth and K. G. McKay, "Photon emission from avalanche breakdown in silicon," *Phys. Rev.* 102, 369 (1956).

37. D. J. Rose, "Microplasmas in silicon," *Phys. Rev.* 105, 413 (1957).

38. A. G. Chynoweth and G. L. Pearson, "Effect of dislocations on breakdown in silicon *p-n* junctions," *J. Appl. Phys.* 29, 1103 (1958).

39. A. G. Chynoweth, "Electrical breakdown in *p-n* junctions," *Bell Lab. Record* 37, 47 (1958).

40. J. Maserjian, "Determination of avalanche breakdown in *p-n* junctions," *J. Appl. Phys.* 30, 1613 (1959).



International Rectifier manufactures a Zener diode type and rating for every voltage regulation and control application.

Silicon Zener Diode Regulators

Practical Considerations for Applying Zener Diodes to Electronic Circuitry

To all external appearances, the Zener diode looks much the same as other silicon rectifying devices and is available in axial lead and stud mount packages. Further, the device is capable of converting alternating current (AC) into pulsating direct current (DC), as is the silicon rectifier.

Inside the Zener diode package is a different story, however. As pointed out in Chapter 1, these units are "cultured" to exhibit a particular effect. If the diode is biased in the forward direction (positive to anode), considerable current will flow when the 0.6 Volt "barrier potential" is exceeded. When a source of low voltage is applied to the diode in the reverse direction (positive to cathode), the junction back resistance remains quite high and junction current is in terms of microamperes. As the reverse potential is increased, the junction reaches a critical point and the diode avalanches. This breakdown is not destructive, as long as the diode's dissipation capabilities are not exceeded, and the device may be cycled in and out of the region as often as is necessary.

The effect just described is a function of silicon resistivity in the lightly doped p-n junction and to a less extent by the gradient in the impurity concentration near the junction. These conditions can be carefully controlled in manufacture and it is possible for International Rectifier to supply regulators in JEDEC steps from 2.4 to 200 Volts.[1]

When avalanche breakdown is reached, the normally high back resistance drops to a low value and the junction current increases rapidly, limited principally by circuit resistance. As the voltage is increased beyond the breakdown point, the diode current increases proportionately, but the junction voltage remains essentially constant as shown in Fig. 1.

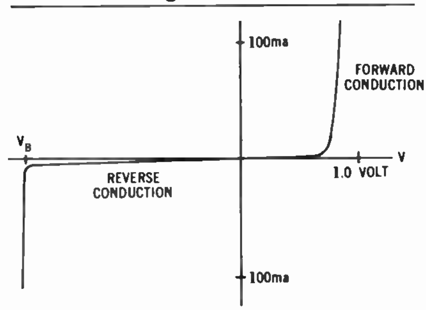


Fig. 1. Current and voltage characteristics for a typical Zener diode regulator.

Using the Zener Diode

Fig. 2 is the familiar circuit for a Zener diode regulator. The diode, a shunt element, draws current through resistor R, which is also in series with the load. The total current through R is

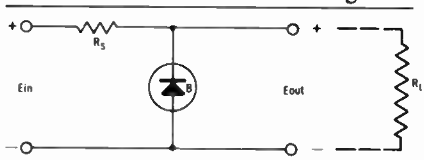


Fig. 2. A Zener diode in a shunt regulator configuration.

the sum of the Zener and load currents.

If E_{in} increases, the current through both the Zener element and R_L will increase. Simultaneously, however, the diode resistance decreases and the junction current increases, thereby adjusting the voltage drop across R_L and maintaining the original voltage across the diode regulator. The ability to maintain this voltage is determined by the temperature coefficient and the Zener diode impedance.

Load variations have a similar effect on the diode regulator. As the load increases, or decreases, the Zener shunt element will draw less or more current respectively. The net result is substantially a constant output voltage across the diode regulator.

Design Considerations

The operation of the simple regulator can be developed into several useful equations and from the derivations, the component values and circuit conditions can be calculated.

The basic circuit equation is:

$$V_{in} = R_s(I_z + I_L) + V_z$$

where, for this and the following equations;

V_{in} = supply voltage
V_z = Zener voltage
V_L = output voltage
I_z = Zener diode current
I_L = load current
R_s = series resistor
R_L = load resistance
P_d = Zener diode dissipation
P_{dmax} = maximum dissipation

From this, the following is derived:

$$R_s = \frac{V_{in} - V_z}{I_z + I_L}$$

$$I_z = \frac{V_{in} - V_z}{R_s} - I_L$$

$$P_d = \left(\frac{V_{in} - V_z}{R_s} - I_L\right) V_z$$

The following equations can be used to calculate the approximate circuit values:

For constant V_{in} and variable I_L;

$$R_s = \frac{V_{in} - V_z}{I_{Lmax} + 0.1\ I_{Lmax}}$$

$$P_{dmax} = \left(\frac{V_{in} - V_z}{R_s} - I_L\right) V_z$$

For constant I_L and variable V_{in};

$$R_s = \frac{V_{inmin} - V_z}{I_L + 0.1\ I_L}$$

$$P_{dmax} = \left(\frac{V_{inmax} - V_z}{R_s} - I_L\right) V_z$$

For variable V_{in} and variable I_L;

$$R_s = \frac{V_{inmin} - V_z}{I_{Lmax} + 0.1\ I_{Lmax}}$$

$$P_{dmax} = \left(\frac{V_{inmax} - V_z}{R_s} - I_{Lmin}\right) V_z$$

It should be pointed out that these equations include a tolerance factor to guarantee that the diode will regulate at I_{Lmax}. For conservative designs, an empirical factor of 10% of I_{Lmax} should be used as I_{zmin}.

When .1 I_{Lmax} is used as I_{zmin}, the value of R_s may not be optimum but will be as large as possible consistent with desired regulation and minimum device dissipation.

It is assumed in the equations that the resistance of the voltage source (V_{in}) and the dynamic impedance of the diode are a negligible part of R_s. If the impedance of these two components is an appreciable portion of R_s they must be taken into consideration.

From the equations just given, it can be seen that resistor R should be so selected that the current in the diode does not exceed I_{zmax} or the maximum device dissipation with small load currents. In conservative designs I_z is chosen for approximately 20% of I_{zmax}. When connected as a shunt regulator, the diode must absorb current variations between the I_z and I_{zmax} limits.

It should be noted that although I_z is the recommended operating point, any current beyond Zener breakdown may be arbitrarily selected. A family of reverse current curves for typical Zener diodes is shown in Fig. 3. Note particularly that the 4.7 and 5.6 Volt diodes have a round knee at low currents, while the higher voltage units exhibit a sharper breakdown. For this reason, if operation at very low currents is desired, diodes exhibiting a V_z above approximately 7 Volts should be chosen.

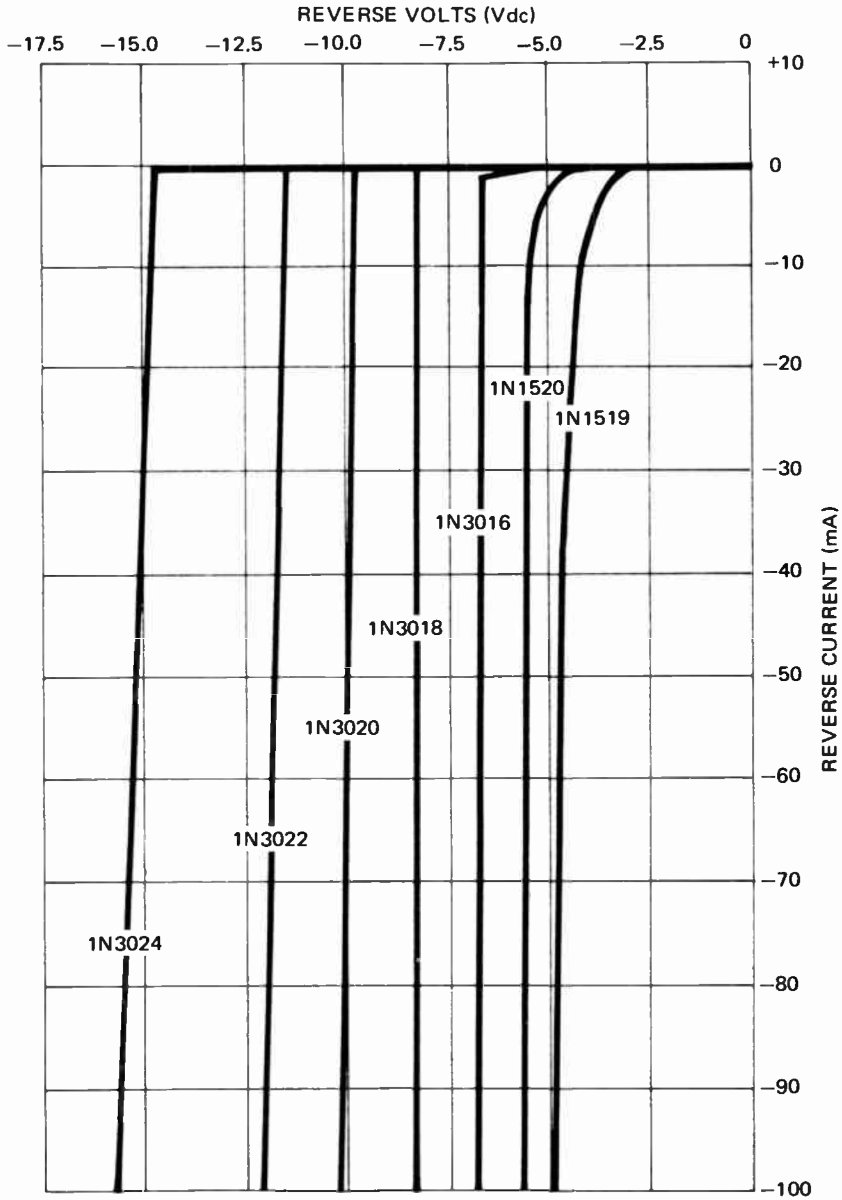


Fig. 3. Reverse current curves of typical Zener diodes.

Power Handling Ability

The maximum permissible diode current is limited by the temperature rise of the junction and by the heat sink provided to dissipate this heat. As the junction temperature rises, the dynamic resistance will also increase. It is this thermally induced resistance which limits the maximum operating temperature of the diode.

Heat generated internally at the crystal junction combines with the ambient temperature to determine the total junction temperature. The power handling ability of most diodes is rated at an operating temperature of 25°C, or

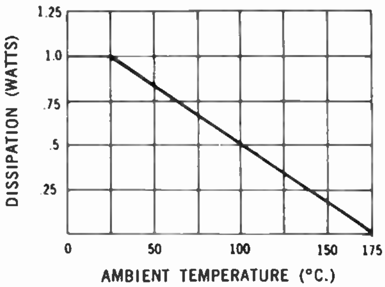


Fig. 4. Typical temperature derating curve for a 1 Watt diode.

approximately room temperature. As the ambient temperature increases beyond this point, the diode can no longer dissipate as much heat and must be derated if operating near I_{zmax}. A typical derating curve for a one-watt Zener diode is shown in Fig. 4.

Dynamic Resistance

The dynamic resistance (R_z) in the avalanche region is the basic parameter for establishing the regulating ability of the silicon voltage regulator. It is one of the most important factors to be considered when selecting a diode to be used in a regulator application. The dynamic resistance is determined by measuring the AC voltage developed across the regulating diode when operating with an AC signal impressed on the DC corresponding to I_z. Additional information on the test procedure is given in Chapter 5, AC Applications.

The dynamic resistance is an expression of the change in voltage for a small change in current about its DC operating point. All International Rectifier Zener diodes are measured by using a value of AC which is 10% of I_z. For a diode having a nominal I_z rating of 30 mA, a 3 mA AC signal would be used.

Since V_z is relatively constant, the diode resistance becomes a function of dc flowing through the diode. Under conditions of rated current, it is in the order of a fraction of an Ohm for low voltage diodes to several hundred Ohms for high voltage types. As R_z is dependent upon the operating current, those diodes capable of high dissipation (hence higher current) offer a much lower resistance.

A change in dynamic resistance can be observed with variations in ambient temperature, the higher the temperature, the greater the resistance. This change is linear with temperature, increasing by approximately 30% for a 100°C temperature rise.

Temperature Coefficient

The resistivity of silicon varies with temperature and therefore a temperature -coefficient is inherent in the Zener diode. Fig. 5 shows a typical curve of Zener voltage versus temperature, in detail. It can be seen that the voltage is dependent to a degree upon the operating and ambient temperature.

As explained in Chapter 1, the breakdown voltage influences the temperature coefficient. In the true Zener region the temperature coefficient is negative. This is to say — an increase in temperature will cause a decrease in the Zener voltage. In the avalanche region (above 5 to 6 Volts), the temperature coefficient is noticeably positive. This transition region is not well defined, and 5.1 Volt diodes may exhibit a negative or a positive tempera-

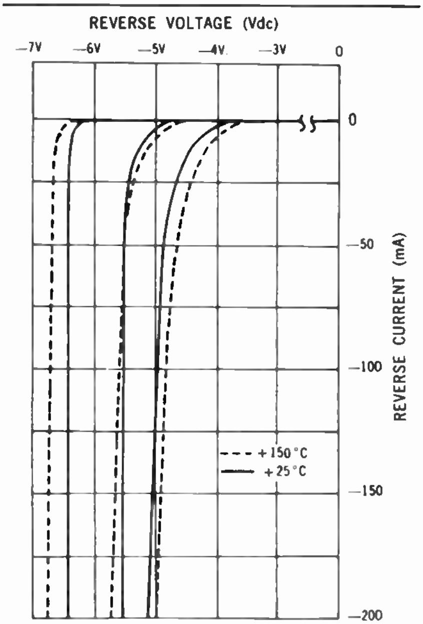


Fig. 5. Temperature coefficient curves for several diodes in the 5 to 6 volt range.

ture coefficient, depending on the junction current.

In certain applications, the circuit designer may take advantage of these coefficients to provide compensation for another circuit element tending to drift in an opposite direction. However, most applications call for a stable reference in the presence of temperature changes. A discussion of zero temperature coefficient diodes is, therefore, in order.

Zero Temperature Coefficient

In the region between 4.7 and 6.2 Volts the temperature coefficient may be made positive or negative by control of the reverse current. A point of zero temperature coefficient is shown in Fig. 5 (point a), at a value of 45 mA. It should be understood that this point of exact zero coefficient, for diodes in the 5.1 Volt region, holds true for a specific current only.

The temperature coefficient is a constant for a given regulator, and is related to its breakdown voltage as illustrated in the curve, Fig. 6. This graph indicates that the coefficient, although approaching 0.1% per degree Celsius (Centigrade) at the higher voltages, passes through zero at about 5 Volts and then becomes negative for lower voltages, reaching −0.04%/°C at approximately 3.5 Volts.

The zero temperature coefficient

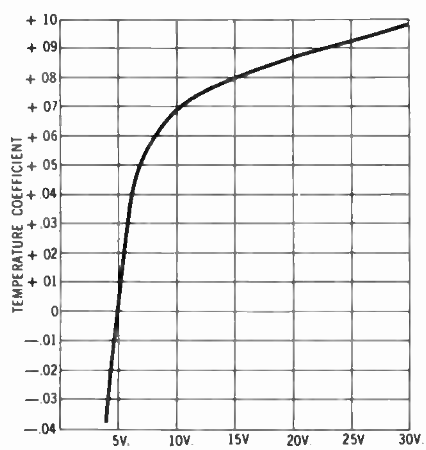


Fig. 6. Zener diode temperature coefficient curve for a typical 5 Volt unit.

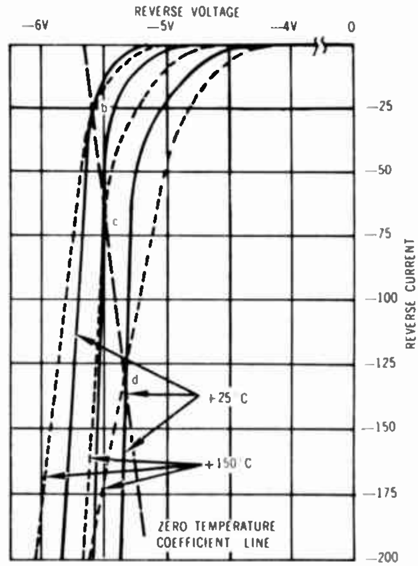


Fig. 7. Coefficient curves for a group of diodes in the 5.1 Volt region showing the zero temperature coefficient time.

characteristic is not limited to diodes of exactly 5 Volts breakdown only, but can be found at various operating current points in the voltage range from 4.7 to 6.2 Volts. As shown in Fig. 7, points b, c, and d, are intersections that provide zero temperature coefficients at a definite reverse current. Below 4.5 Volts, the high current value necessary to achieve the zero coefficient condition is prohibitive for it approaches the maximum I_z of the device. Above 6.2 Volts, the zero coefficient intersection has already reached the saturation leakage current level.

In order to illustrate more clearly how reverse current is the determining factor in achieving this essentially zero temperature coefficient effect, a diode rated at 5.1 Volts was selected and its thermal stability measured at a number of different reverse currents. The results of the measurements are plotted graphically in Fig. 8. This family of curves is, in effect, an expansion of the crossover region at point c in Fig. 7 and shows the deviation from an initial voltage drop at +25°C for various amounts of reverse current. For example, a diode having a drop of 5.1

Volts at $+25°C$ at a reverse current of 50 mA would exhibit a negative change in voltage as the temperature is increased. In this case a 50° increase in temperature would cause the voltage to drop from its initial value to 5.092 Volts at $+75°C$, or a net change of minus 9 millivolts. On the other hand, 80 mA of current would cause the voltage to increase by approximately 7 millivolts for the same temperature rise.

Of specific interest is the fact that, at this current, the voltage will first change in the positive direction and then at about $+75°C$ will reverse, pass through zero, and then go predominantly negative. At $+150°C$ the net change will be almost exactly equal and opposite to the positive deviation at the lower temperature. This particular diode at 80 mA would therefore have a stability rating of approximately $\pm0.003\%/°C$ from $+25°C$ through $+150°C$.

If operation over a limited temperature range is expected; say $+25°C$ through $+75°C$, it would then be advantageous to operate at an intermediate current value; in this case 57.5 mA where a temperature coefficient of slightly over $\pm0.001\%/°C$ can be expected.

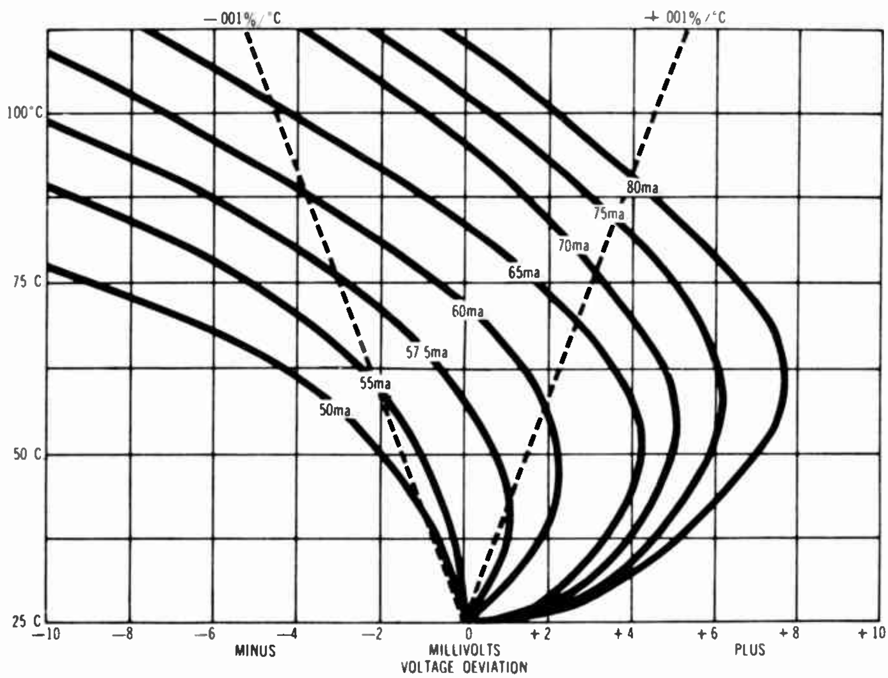


Fig. 8. An expansion in the region of c (Fig. 7) showing that reverse current determines temperature coefficient.

1. Electronics Industries Association, "JEDEC Suggested Standard No. 2", Washington D.C. July 1966.

Silicon Temperature Compensated Reference Diodes

The Solid State Equivalent of the Standard Cell

The Silicon Temperature Compensated Reference Diode (commonly known as the TC Zener) is a multiple junction form of a Zener regulator diode. Its primary characteristic is that of high stability of breakdown voltage with reference to temperature variations. As such it fills a specific area of need for the designer of precision equipment such as computers, ultra stable power supplies, or in any application requiring standard cell accuracy.

Theory of Operation

Among the various Zener diodes, there are some which lend themselves readily to temperature coefficient compensation. One of these is the 5.6 Volt Zener. It has a positive temperature coefficient which yields a change in Zener voltage of approximately +2.0 millivolts per degree Celsius (Centigrade). A forward biased silicon diode on the other hand will have a negative temperature coefficient of approximately −2.0 mV/°C. By controlling the temperature coefficients of each diode chip through close control on silicon resistivity and the process of junction formation, it is possible to manufacture junctions which will very nearly compensate one another. If the dies are then placed electrically in series as shown in Fig. 1, and if the voltage sensed is the total voltage drop of Zener die and forward biased die (stabistor), then temperature variations will have little effect on output voltages. This is the 6.2 Volt temperature compensated Zener diode.

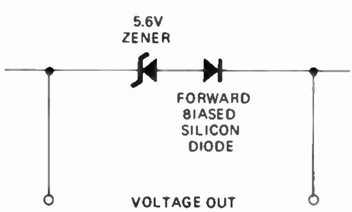


Fig. 1. Typical TC Reference Diode Schematic.

There are also two other standard reference voltages, 8.4 and 9.0 Volts. These devices consist of three dies in series with one Zener junction and two stabistors. The 8.4 Volt unit consists of a 7.0 Volt Zener die and two stabistors. The 9.0 Volt unit consists of a 7.6 Volt Zener die and two stabistors. Close control in resistivity of the silicon used to manufacture the stabistors for these devices accounts for the small difference between their output voltages.

TC Zener Parameters

Since the TC Zener is electrically nothing more than a low temperature coefficient Zener diode with high breakdown voltage in the "Forward Biased" mode (The stabistor junction is reverse biased), most of what has already been said on Zener regulators applies to this device. It has dynamic impedance, reverse leakage, capacitance, etc. The two parameters that set this device apart from Zener regulators are low temperature coefficient and qualities of time stability. It is these two areas that must be explored in specifying a TC Zener.

Temperature Coefficient

The original terms used to describe temperature coefficient, %/°C, indicated a percentage change in breakdown voltage for each degree of ambient temperature change. Common values used in specifying the maximum limits are ±.01%/°C, .005%/°C, .002%/°C, .001%/°C, .0005%/°C. Advances in technology have recently brought this figure down to .0002%/°C. These values are now most commonly spoken in terms of parts per million (See Table I). The graph in

TABLE I

.01%/°C	= 100PPM/°C
.005%/°C	= 50PPM/°C
.002%/°C	= 20PPM/°C
.001%/°C	= 10PPM/°C
.0005%/°C	= 5PPM/°C
.0002%/°C	= 2PPM/°C

Fig. 2 shows two ways in which the specification on the 1N821 may be interpreted for testing. The terminology of %/°C was initially accepted by the industry as implying a linear relationship between temperature and the change of Zener voltage. It soon became apparent that manufacturers did not mean to imply a linear relationship; in fact they could not supply diodes with linear characteristics over the wide ranges of temperature where these low values of TC were tested. Thus evolved the concept of the "Hourglass Envelope." It was hoped that while exact linearity could not be guaranteed, the TC characteristic would at least fall within the confines of the "Hourglass," a shape formed by intercepting a reference point (25°C) with the positive and negative slopes of the maximum %/°C change specified on the device as shown in Fig. 2, ".01%/°C envelope" (implied). Many devices fell into this envelope, but some did not, even though the maximum change indicated by measuring the change of voltage from a reference point to the temperature extremes showed the device to be within the maximum limit.

The plots of voltage change vs. temperature shown in Fig. 3 are not uncommon for a TC Zener and indicate why there was a problem of guaranteeing the Hourglass Envelope. The diode which exceeds the envelope limits could not be detected without testing intermediate temperature test points between 25°C and 100°C. This type of testing proved extremely costly, so the "Box Method" was evolved.

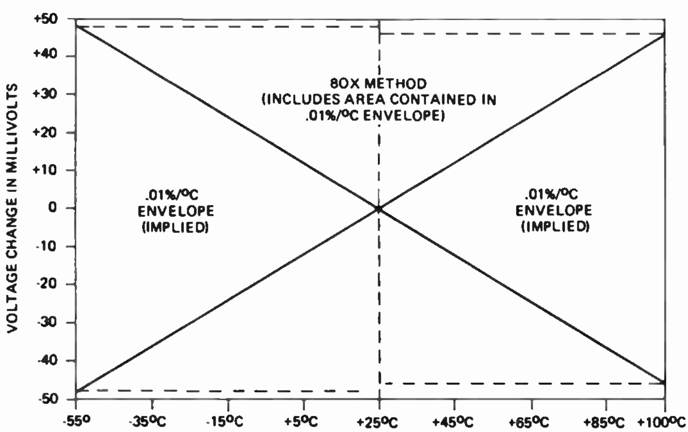


Fig. 2. Temperature Coefficient maximum limits for 1N821.

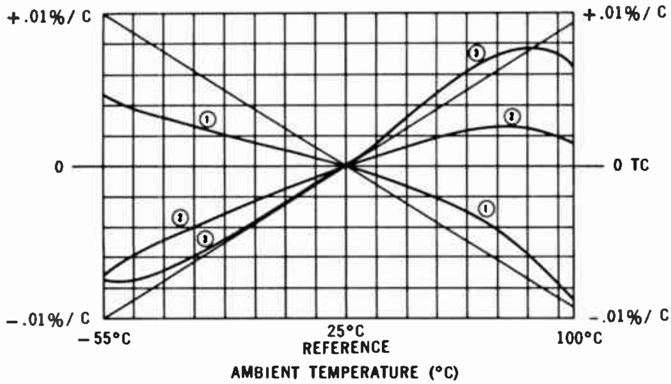


Fig. 3. The Hourglass Envelope for a .01%/°C Diode with Some Typical TC Curves

The original concept of the Box Method (still used on the 1N821 series) stated that the change of voltage from the reference point at the specified temperatures would not exceed the limits of Max. Δv_z = Rated TC in $\frac{\%/°C}{100} \times V_z$ @ 25°C × number of degrees Celsius between test points. The 25°C reference point assures that better than 90% of the devices (recently stated as 99% by one manufacturer) will in fact fall within the limits of the box. As shown in Fig. 4, it is still possible for some devices to exceed the box limits.

The Characteristic of Temperature Coefficient

A problem in the TC specification arises because of the characteristic of the temperature coefficient of a TC Zener. The TC of a reverse biased Zener junction having a given positive coefficient tends to become less positive as ambient temperature rises through the region of 40°C to 75°C. The specific temperature and the magnitude of change will vary from one diode to another but it will always occur to some degree. A forward biased junction also has a tendency to have a change of slope of its temperature coefficient (always a negative TC)

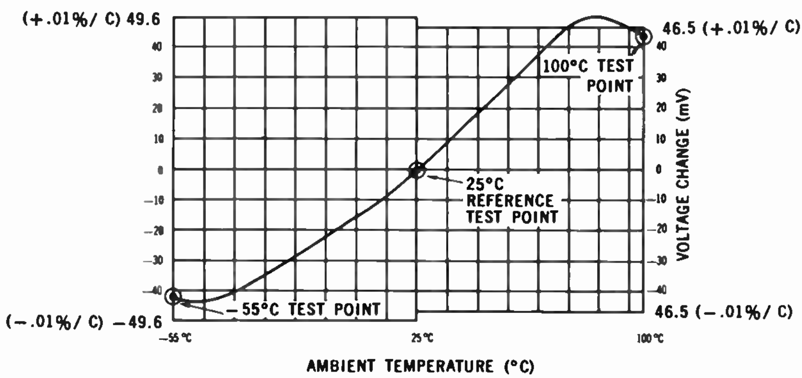


Fig. 4. 1N821 TC Characteristic Plotted in the Original "Box Method" Test Limit

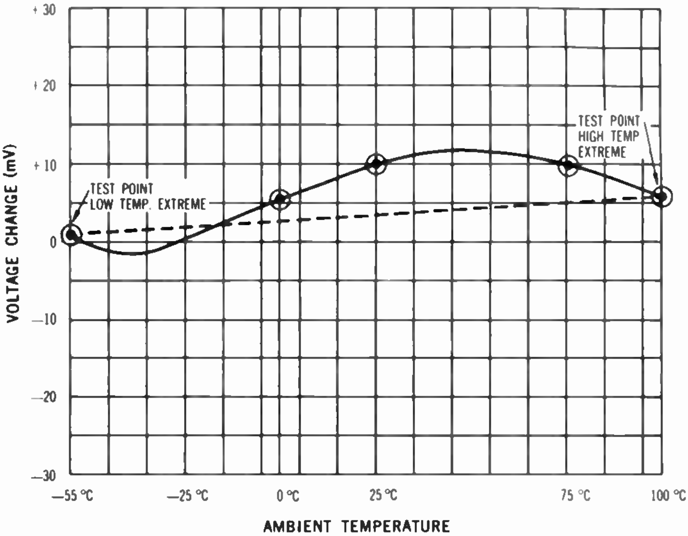


Fig. 5. A Typical TC Characteristic Showing the Hyperbolic Change of Slope

as temperature rises through this region. The combination of the two TC characteristics changing will cause a hyperbolic curve as shown in Fig. 5. This change of slope will continue as the junction temperature goes beyond the 75°C point and in many cases the slope will increase. If the diode in Fig. 5 was a 1N938A (10PPM/°C, −55°C to 100°C), measurement in accordance with the JEDEC registration would yield a good diode. This is because the indicated change from one extreme to the other was only five millivolts, whereas the maximum permissible change over the range of temperature would be 14 millivolts. Yet the voltage actually changed along a slope of 0.2 millivolts/°C from −25° to +25°C yielding an actual TC in that region of nearly 25 PPM/°C. Therefore, it should be obvious that buying the lowest TC diode available, tested over the widest range does not ensure the best TC over all temperature ranges.

The best possible approach to purchasing a commercially available device is to select a diode which was tested over the anticipated operating range. A common range of temperature found in many solid state circuits is 0° to 75°C. The 1N939 would be the best 9 Volt TC diode for this range of ambient temperature, not the 1N939A or the 1N939B, since it is tested over just that range of temperature, not from −55°C to 100°C or 150°C as are the 1N939A and 1N939B. In addition, the 1N939 would be the least expensive diode of the three.

Where high precision is required, the best approach to TC is to specify the anticipated temperature range and purchase a special device.

Correlation Between Test and Use

There are other pitfalls in temperature coefficient that the designer should consider. Most manufacturers, in an effort to perform exact TC testing, have been making measurements in oil baths precisely controlled to the various temperatures. This procedure, while fully accurate when the diode is mounted in this type of ambient, is not accurate when the device is mounted on a printed circuit board in air. The reason for this is the junction to ambient thermal impedance of the device. When the diode is immersed in an oil bath with the oil being agitated at a very fast rate the junction to ambient thermal impedance is very low. Placing the same diode in an ambient of relatively still air with a ½″ or ¾″ lead length between diode body and the circuit board tie point will yield a junction temperature of, for instance, 10°C higher than that of the same diode immersed in agitated oil. This can create a very different appearance in the TC of the Reference Diode under actual operation compared to the TC measured by the manufacturer. Consider a TC characteristic with a hyperbolic shape as shown in Fig. 5. Suppose the slope occurring in region beyond 75°C were greater than that which is shown. Since the junction temperature, not the ambient temperature, determines the change in the diode voltage, the diode would exhibit a change commensurate with an additional 10°C above that experienced in the oil ambient. This would put it in an area where the diode has a high TC relative to that which was measured by the manufacturer. To offset this condition somewhat it is necessary to mount the diode with as short a lead length as possible (¼″ to ⅜″) but under that condition care must be taken not to damage the diode during the soldering operation. For an even lower junction to ambient thermal impedance the diode may be mounted with ⅜″ lead spacing to turret terminals which will substantially reduce the thermal impedance. The spacing must be ⅜″ or slightly less in order to obtain the best use of the turret terminal since this is the area where the maximum heat transfer takes place from the lead to the ambient. Another alternative is to operate the diode in an air ambient 10°C to 15°C below the manufacturer's maximum tested temperature. This will assure operation within the tested limits.

If extremely small changes of voltage in the Reference diode are necessary for proper operation and the TC requirement is considerably better than that available commercially, there is a way of achieving this. There are small ovens available which exactly fit the DO-7 package. They are set to operate at either 80°C or 125°C. The internal temperature will rise only 1°C for every 15° external ambient temperature increase up to the area of the set temperature. This will reduce the *effective* TC to a value of less than 10% of the manufacturer's measured value making the 0.5PPM/°C (ambient) device possible. The ovens operate from a 24 Volt DC source, require very little power and are small in size. They are well worth the cost where extremely low TC's are necessary.

Effect of Varying Current on TC

Another consideration in design, and a very important one, is the constant current source used to operate the Reference Diode. An important consideration with compensated Zeners is the constant current source used. The designer should not *allow* the constant current source to vary during operation. This will create a change in the reference voltage proportional (ΔI_z) (Z_z). The change created by allowing the current to vary ±100 microamps in a 1N827 with a 15 Ohm Z_z would create a change of ±1.5 millivolts for a total swing of 3 millivolts or ⅓ of the maximum allowed for a temperature range of −55°C to 100°C. Therefore, where special attention is required for high stability of reference voltage, selecting the best TC Zener must go hand in hand with providing a stable source of Zener current. Usually, manufacturers do not state what happens if the Zener current is 7.0 milliamps or 8.0 milliamps instead of 7.5 as specified. As a general rule, there will be a shift of 10% to 20% in the TC on the .01%/°C, .005%/°C, and .002%/°C devices as the Zener current is shifted away from the 7.5 milliamp test current by 0.5 milliamps. (Due to the critical balance of compensation existing in the better diodes, greater changes may be anticipated for these devices.) Obviously this will not apply to all diodes but can be applied as a good rule of thumb.

There is a phenomenon associated with the TC Zener which makes it possible to improve the TC by deliberately varying the Zener current. Diodes having a positive TC at 7.5 milliamps will have a better TC if the Zener current is set somewhere below 7.5 milliamps. It is sometimes possible to reduce the Zener current 6.5 milliamps and produce a 10PPM/°C device at 1.0 milliamp. On a diode with a negative TC at 7.5 milliamps, increasing the Zener current will produce a lower TC. While this is obviously not a tool for the major Reference Zener purchaser due to the high expense involved in testing large quantities of diodes, it can be a very good rule for the experimenting design engineer or technician. This rule allows the use of 50 and 100 PPM/°C devices which may be available, saving the time required to have a special low TC Zener made up at high cost to continue the experiment. At the successful completion of the experiment with the circuit tried and proven, the better TC Reference can be specified and manufactured in quantity for production runs of the circuit.

Time Stability

Another factor which is sometimes taken into consideration when designing high stability power supplies and reference sources is time stability.

After the designer selects the best diode and has a quality constant current supply, he must view the possibility that Zener voltage may vary with time. In order to make it possible to control this parameter, compensated Zener manufacturers have introduced lines of time stability tested and guaranteed devices. Some commonly specified values of time stability are 5 PPM/1000 hrs., 10 PPM, 20 PPM, 25 PPM, 50 PPM, and 100 PPM. Since these devices must be tested using elaborate instrumentation, the manufacturing costs in test are extremely high. Therefore, the cost to the purchaser is necessarily high.

Stability Instruments Possible

A typical plot of time stability on the 1N827 is shown in Fig. 6. Many of the small variations in voltage might be attributed to small variations in diode temperature, minute changes in Zener current, and various other instrument drift related problems. This is one of the prime limitations on supplying devices in better time stability ratings, instrumentation stability.

The Compensated Zener Compared to the Standard Cell

Through the years, the primary source for standard voltage values has been standard cells.

The standard cell is a "wet cell" battery having a voltage of about 1.019V. It serves the industry by providing a constant voltage output so long as it is handled properly and not overloaded. The standard cell is sensitive to both current being drawn out of the cell and reverse bias forcing current back into the cell. Its temperature coefficient is about .002%/°C and its operating temperature range is from 4°C to 40°C.

This short tangent into the characteristics of the standard cell clearly illustrates a key roll of the compensated

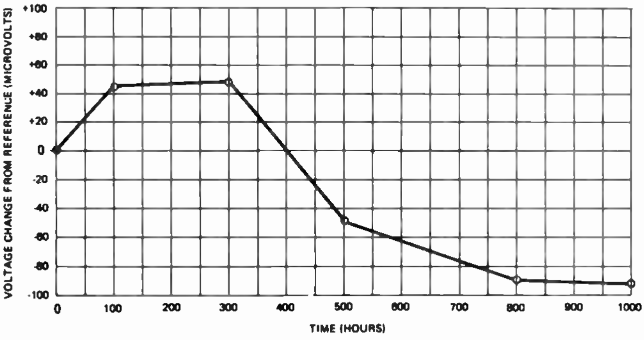


Fig. 6 Typical long term stability – selected 1N827 (2) 40° C.

Zener in industry. The compensated Zener can be applied as a working standard cell for rough handling and wide ranges of ambient temperature. The standard cell provides a laboratory standard with consistent time stability characteristics while the compensated Zener is used in different AC voltmeters, digital voltmeters, secondary voltage standards, etc.

The time stability of the compensated Zener has now reached a point where the true "Solid State Standard Cell" is a reality. These devices are being selected for stability characteristics at least as good as the normal type of standard cells. As technology improves, more of these devices will be available and costs will reach the level where the solid state standard cell will replace the old style standard in many laboratory applications.

Applications

Since a source of constant current is necessary to obtain the best voltage stability from the compensated Zener, some techniques for attaining constant current are shown in Fig. 7.

Four simple methods of obtaining a constant current that will be virtually unaffected by temperature are depicted in Fig. 7. Circuit (a) is usable only if a source of well-regulated high voltage is available. It must have sufficient reserve current available to supply the additional 10 mA required by the reference element. This power supply should be almost impervious to temperature shifts.

In cases where no such separate regulated source is available, or where isolation is required, an independent supply must be designed using two series-connected 1N937A's as "pre-regulators," as shown in circuit (b). Since these are also highly stable devices themselves, the end reference element is assured of an extremely constant current over a wide range of temperatures. The chief disadvantage is the cost of the additional 1N937A's.

Less costly Zener diodes may be used as the pre-regulator in place of the 1N937A's as diagrammed in circuit

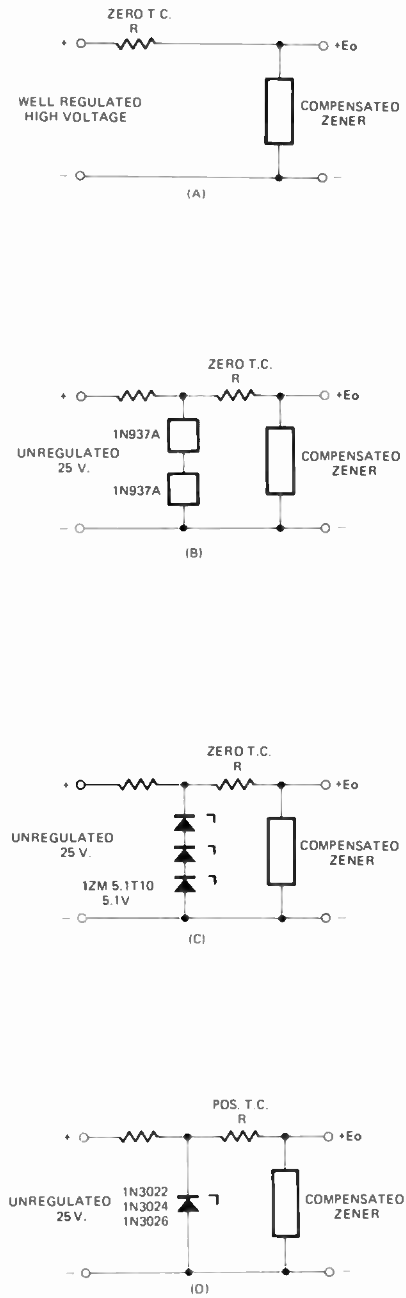


Fig. 7. Typical bias current systems. For most applications (c) would be the "preferred circuit."

(c) of Fig. 7. The three 5.1 Volt diodes operating near their zero temperature coefficient current point provide excellent stability.

A fourth method involves only the use of a single Zener diode in the range of 12 to 20 Volts as shown in (d). Such a diode, however, will suffer from an inherently positive temperature coefficient; i.e., its voltage will rise with temperature which increases the current through the reference element. This disadvantage can be overcome by compensating with resistor R, which also has a positive coefficient. As the voltage across the Zener regulator increases with temperature, the resistor will tend to increase in resistance so as to maintain a relatively constant current through the compensated Zener.

In all the circuits shown in Fig. 7, resistor R is selected to provide the 7.5 mA reference current.

Differential Amplifier

Power supply regulators quite often employ a differential amplifier as a sensing device. A typical circuit is shown in Fig. 8.

T.C.D. is a reference diode which provides a reference voltage maintaining a precise 6.2 Volts at the base of Q_1.

The voltage applied to the base of Q_2 varies due to variations in voltage across the resistor voltage divider R_1, potentiometer R_2, and R_3, thus varying the current through resistor R_4 which is common to the two emitters. The resulting change in voltage across R_4 causes a change in the output of transistor Q_1, and this change is fed to a regulator for the power supply.

Digital-to-Analog Converter

The circuit shown in Fig. 9 is a digital-to-analog converter and uses a reference diode as a voltage regulator. The accuracy of the circuit is dependent on very low tolerance resistors and the ability of the regulating diode to maintain a constant voltage E_Z.

The inputs to the converter are connected to binary storage elements such as flip-flops. If the flip-flops are high ("1" level) then the diodes CR_1, CR_2, CR_3, and CR_4 are turned on. The current flowing through CR_1 is E_Z/R; through CR_2 is $E_Z/2R$ and so on. Therefore, the output voltage (or current through R_L) is proportional to the binary number stored in the flip-flops. The maximum change in output current for which the reference diode must compensate will be $2E_Z/R$.

The voltage temperature coefficient of the reference diode will be degraded somewhat due to variations of I_Z. However, the designer by choice of com-

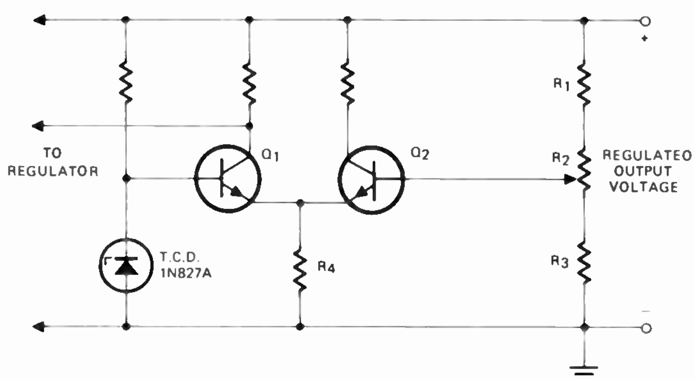


Fig. 8. Typical Circuit Design.

ponent values can limit the magnitude of the variations and still obtain the benefit of a much lower temperature coefficient than is provided by a device which is not temperature compensated.

Another very common application for the compensated Zener is use as a voltage regulator diode. In this application Zener current is not fixed and it varies over a wide range. Even though the specified current is not maintained the diode is still capable of compensating itself. While compensation may not be .01%/°C for the 1N935, it could be a .015%/°C. A normal 9.0 Volt uncompensated regulator has a .065%/°C temperature coefficient and the compensated Zener represents a marked improvement over the uncompensated.

DC Constant Voltage Supply

The power supply shown in Fig. 10 depends upon the compensated Zener for stability in the output voltage. The series regulator transistor base bias is a function of the differential between the divided output voltage and the difference voltage. When a difference exists, it is amplified and used to change the bias on the series regulator.

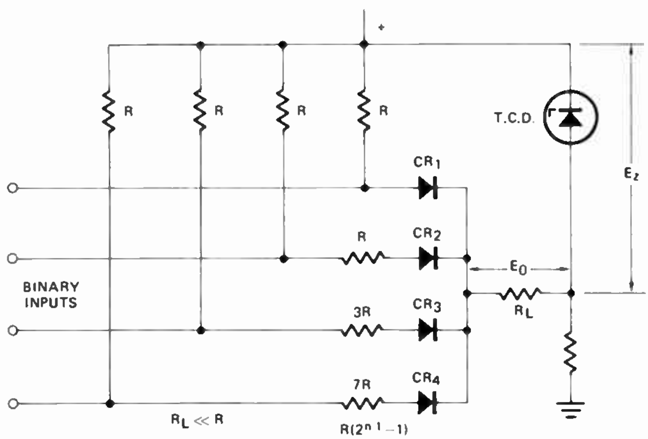


Fig. 9. Digital-to-Analog Converter.

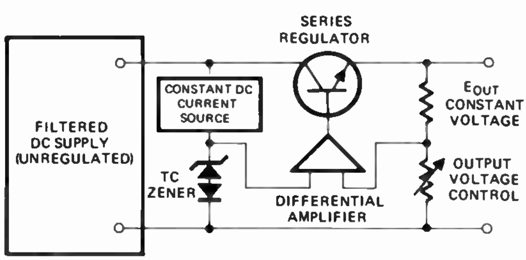


Fig. 10. Constant DC Voltage Supply Utilizing TC Zener for Output Voltage Stability Control

One of the controlled processes in the manufacture of Zener diodes.

Thermal Considerations

Important Considerations when Operating Zener Diodes at Elevated Temperatures

Any circuit element possessing resistance to current flow will generate heat. The quantity of heat developed will depend on several factors, among them the amount of current and the resistance of the element. Even more important, in the case of Zener diodes, is the ability of the device to dissipate as much of this heat as possible through convection, conduction, and radiation.

Heat, or more specifically, the junction temperature rise, limits the power handling ability of the element. Power, of course, is a product of the junction current and voltage. The maximum current that can be handled by a particular diode style (referred to as I_Z max.) is different for each Zener voltage. Zener diodes are usually operated at 20% of I_Z max., but in some applications the junction temperature is allowed to rise near the maximum limit of 175°C. If the device is elevated to excessively high temperatures, additional thermally induced junction resistance will be exhibited by the diode, destroying its value as a regulator. If there is no current limiting circuit element, the diode may enter the thermal runaway region and destroy itself.

Ambient Temperature

The heat produced internally is added to the environmental temperature to determine the total junction

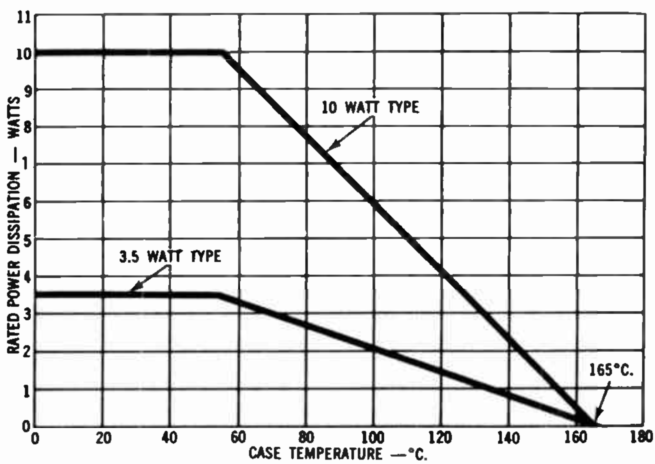


Fig. 1. Derating of Power Dissipation vs. Case Temperature for 10 Watt and 50 Watt Devices.

temperature. In most cases the power rating of International Rectifier Zener diodes is based on an ambient temperature between 25°C and 65°C, depending on type. Above this point the Zener diode must be derated or additional steps must be taken to draw heat from the junction. As an example, if a 2 x 2 inch cooling fin were soldered to a 1 Watt Zener diode, the device dissipation would be increased to approximately 3 Watts. In this case the fin acts as an increase in the heat radiating surface and lowers the junction temperature by conducting heat from the diode and then radiating it to the surrounding air by convection. Such a device is called a "heat sink" or heat exchanger.

Even though the diode is not operating near I_Z max., the use of cooling devices will reduce the degree of voltage variations due to internal temperature changes.

A typical derating curve for a 10 and 50 Watt diode is shown in Fig. 1. Derating curves such as this one are included for each Zener diode in International Rectifier's catalogs and bulletins.

Stud Mount Cooling

Although the axial lead type Zener diode does not normally require a heat sink, the larger stud-mounted diodes must be clamped to a metallic dissipator or heat sink in order to realize the full capabilities of the device. In this case the use of a mounting stud accomplishes a two-fold purpose. It provides an electrical connection to the anode (or in some devices – the cathode). In addition it transfers the thermal losses from the junction to the heat sink or cooling media.

In this discussion the electrical connection is of no particular importance. However, the removal of heat from the junction is of primary concern, since it plays such an important part in the proper operation of the regulator.

Thermal Resistance

Even though the diode is bolted to a metal plate or chassis, there are always points of thermal resistance in the "circuit." Fig. 2 shows an electrical analogy of the thermal circuit. The heat flow may be thought of as current, while the thermal resistance reacts much the same as in an electrical counterpart. Whenever current flows through the thermal resistance, "voltage," in the form of heat will develop. As shown by Fig. 2, there are three major points of thermal resistance: the diode assembly, the contact or mounting resistance, and the heat sink itself.

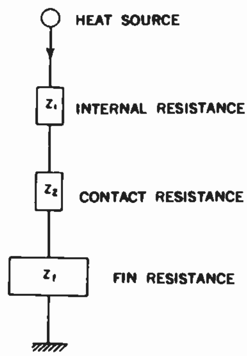


Fig. 2. Simplified electrical circuit analogue.

As far as the diode construction is concerned, the base material should be a good conductor of heat. Copper is used for the base of all stud mounted Zener diodes made by International Rectifier. In addition, the point of mounting contact must be flat and smooth. Accurate machining and cleaning removes this consideration .

From a design standpoint, there are several considerations that the engineer should be aware of. The mounting surface, like the diode, should be smooth and flat to insure maximum contact area. If heat is passed from the diode base to the heat exchanger at only one or two points of contact, the junction will operate at a substantially higher temperature than normal, and the effectiveness of the heat exchanger is lost.

The stud-mounted (DO-4 Outline) Zener diode, which has approximately 0.138 square inches of contact area, should be secured to a fin or chassis that is at least 1/32 inches thick. An area of 2 square inches (each side) will be adequate for the 3.5 Watt unit, while 4 square inches should be used in conjunction with a 10 Watt device. Painting the surfaces black will increase the heat dissipating quality.

Stud Pressure

For maximum transfer of heat, the two mating surfaces should be clamped together as tightly as practicable, considering the case and stud stresses. Clamping is usually accomplished by means of a threaded stud. The nut should be tightened with a torque wrench, at 12 to 15 inch-pounds, to obtain maximum pressure between the two mating surfaces. The pressure is critical as excess tension can strip the threads, rupture the stud at the case, or even distort the diode base. This can, in turn, stress the crystal wafer and cause changes in the electron characteristics of the diode or even crack the crystal, and destroy the diode.

Increasing Thermal Conduction

Good engineering practice also dictates use of a low thermal resistance lubricant between the diode base and the mounting surface. Even when both surfaces appear to be smoothly machined, microscopic examination reveals that there are actually a large number of point contacts. Between each point of contact is a tiny area of dead air, which constitutes a thermal barrier. To eliminate this condition, a small amount of low thermal resistance lubricant should be applied to both mating surfaces before tightening. A silicone grease, similar to Dow Corning DC-200, is recommended.

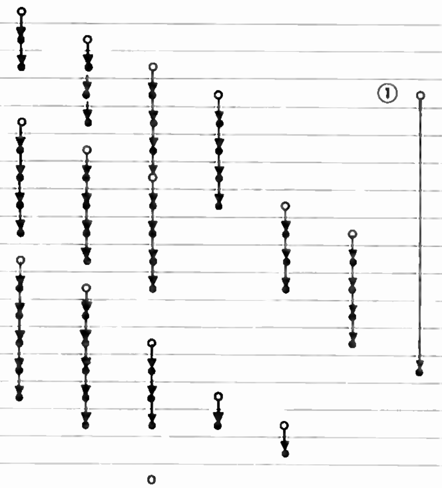
In certain applications it is necessary to insulate the diode case from the mounting surface, while maintaining high thermal conductivity. In these instances, thin washers of high dielectric strength materials are usually employed. Commonly used materials are mica (.003" tk), Mylar (.001 to .003" tk), or a mica bonded glass silicone in the form of insulating washers. A light coating of silicone grease is usually applied to the surfaces to maintain a lower thermal resistance, as decribed earlier. The table shown in Fig. 3 lists some typical values of thermal resistances for the mounting methods just described.

Occasionally, if the heat exchanger material is aluminum, an anodized surface will provide a high electrical resistance, while maintaining a low thermal resistance. Great care must be exercised, however, to avoid rupturing

INSULATING MATERIAL	°C/WATT
Metal to metal with DC-200 grease	2.0
Metal to Metal Dry	2.6
.001 Mylar with DC-200	5.2
.001 Mylar Dry	6.4
.003 Mylar with DC-200	5.6
.003 Mylar Dry	8.8
.003 Mica with DC-200	5.6
.003 Mica Dry	8.0

Fig. 3. Typical contact resistance for International Rectifier Zener diodes with 10-32 NF 2A (7/16 hex size) threads.

CATHODIC END	Anodic Index (mV)	Compatible Couples
1 Gold	0	
2 Rhodium	10	
3 Silver	15	
4 Nickel	30	
5 Copper	35	
6 Yellow Brass	40	
7 Bronze	45	
8 18% Cr. S.S.	50	
9 Tin & Chromium	60	
10 Tin-Lead Solders	65	
11 Lead	70	
12 Stainless Steels	80	
13 Cast Iron	85	
14 Steel (carbon)	85	
15 Aluminum	90	
16 Cadmium	95	
17 Zinc	125	
18 Magnesium	175	
ANODIC END		



① Aluminum is considered compatible with nickel, though not related in galvanic series.

Fig. 4. Galvanic series of metals.

the thin anodic film which would destroy the electrical insulating characteristics. In most applications the washer will provide the most rugged and dependable method of insulating the diode.

From an examination of the table (Fig. 3) it can be seen that adequate derating must be applied when diodes are electrically insulated from the mounting surface.

For a properly engineered product, the final test to prove design adequacy is to measure the Zener diode base temperature with the device operating at the maximum expected load and at maximum ambient temperature conditions. This will insure that operation is well within the manufacturer's current and temperature ratings.

Galvanic Action

Occasionally overlooked is the fact that the two mating surfaces should be chemically compatible to prevent formation of high resistance films due to galvanic action of dissimilar metals. This condition dictates that the diode and the heat sink material be as close as possible in the galvanic series to prevent destructive action when moist or corrosive atmospheres are present. The table in Fig. 4 lists the galvanic series of metals to assist in selecting compatible materials. The copper diode base and stud usually has an electroplated finish of nickel, tin, or silver; indicating that the heat exchanger surface should be close to those in the galvanic table. It is interesting to note that an "active" nickel surface on a diode may be safely used against a tin plated surface. Also it can be seen that a "passive" nickel surface is compatible with a silver plated diode. Further, bases with tin, nickel, or silver plating should not be used directly against aluminum surfaces where moisture or corrosive atmospheres exist, unless the aluminum contact surface is treated in some manner to make it chemically compatible. Aluminum may be electroplated with nickel, silver, or tin; or an irridite surface (chromate process) will provide a smooth and passive surface against which the diode may be clamped.

AC Applications

Zener Diodes May Be Used In AC Circuits As Clipping, Limiting, Regulating and Switching Devices

The design engineer may be inclined to associate the Zener diode, and Zener diode applications, with control and regulation of DC power supplies. Although the majority of Zener diodes are used in this type of circuitry, they are equally applicable in AC, audio, rf, and control systems, as will be pointed out in this and succeeding chapters.

When supplied with alternating current and connected as a shunt regulator (see Fig. 1), the Zener diode is capable of limiting both the positive and negative parts of an AC cycle (Photo 1).

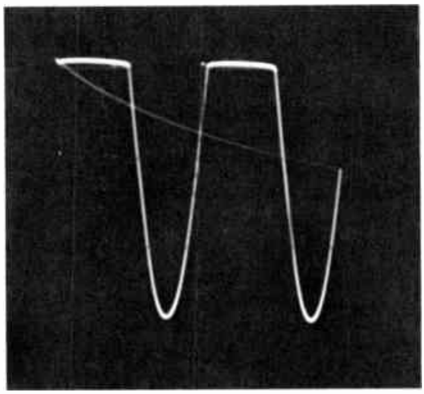


Photo 1. AC signal applied to the Zener diode regulator in Fig. 1. In this oscillogram the negative peak of the AC has not reached V_z, the avalanche breakdown voltage.

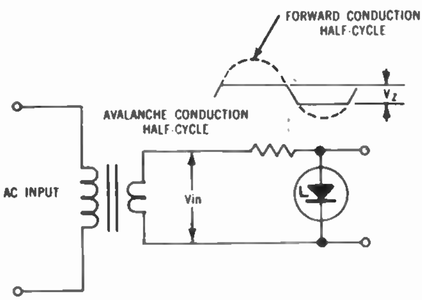


Fig. 1. Basic AC regulator circuit.

The diode conducts almost immediately after the signal passes through zero and into the positive segment. On the negative half cycle, the diode does not conduct until the applied voltage reaches V_z (Photo 2). The net result is a rather non-symmetrical square wave. This effect can be minimized with higher values of V_{in}, but can never be completely eliminated unless two shunt connected diodes are employed in a back-to-back configuration (Photo 3).

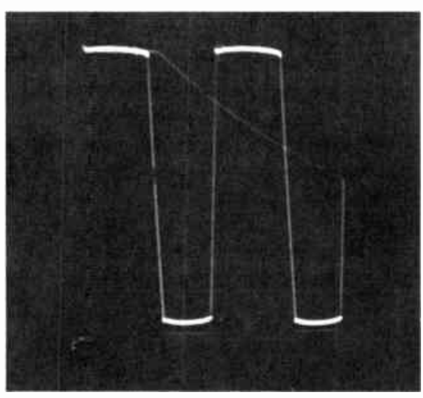


Photo 2. In this oscillogram the applied voltage has exceeded V_z. Note the width at the zero conduction point compared with the width at V_z.

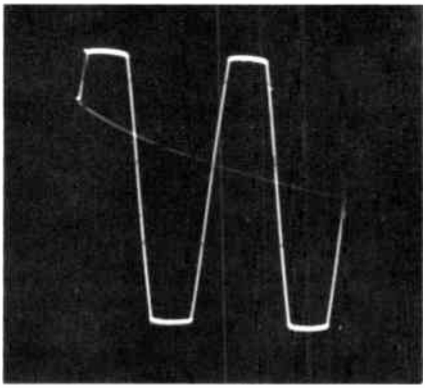


Photo 3. Two Zener diodes connected "back-to-back" (anode to anode) prevent excessive current flow during the forward conduction cycle. The regulator in Fig. 2 will exhibit this waveform.

Filament Regulators

The fact that the Zener diode can be used to limit AC as well as DC provides an interesting application as a filament voltage regulator. It is occasionally necessary to stabilize the AC potential applied to a tube filament. Before Zener diodes were commonly available, control of filament potential was accomplished exclusively by using saturable core reactors, ballast tubes, and elaborate rectifier/regulator circuits.

Filament regulation of a variable frequency oscillator is a logical application for the versatile Zener diode. Undesirable frequency shifts might occur with line voltage variations between 100 and 130 Volts. These extremes would cause a filament voltage change of 5.4 to 7.0 Volts.

An AC filament regulator, for six Volt tubes, may be supplied from a 12.6 Volt filament buss, as shown in Fig. 2. Although 6.3 Volt Zener diodes are available, the regulation is more important than the exact filament voltage. A Zener diode on the low side of the nominal filament voltage might increase tube life.

The back-to-back configuration must be employed to avoid excessive current flow during the forward conduction half-cycle. Connected as shown, the filament voltage is equal to the Zener voltage of Z, plus the forward voltage drop of Z_z on one half cycle.

When dealing with AC, the ratio of average to peak Zener current (ρ) must be taken into consideration. Although this figure varies with diodes and applied voltage, a figure of 0.6 will be suitable for most operating conditions.

In computing values for use in Fig. 2, the minimum value for I_z must be known. To insure proper regulation at minimum line voltage, select a value which is 10% of the minimum load current. For the purposes of illustration it may be assumed the filament current is a linear function and would vary between 260 and 330 mA for the above conditions. Thus an I_{zmin} of 26 mA is established. The minimum current through R_s will then be 286 mA, which is the sum of the minimum filament and Zener diode cur-

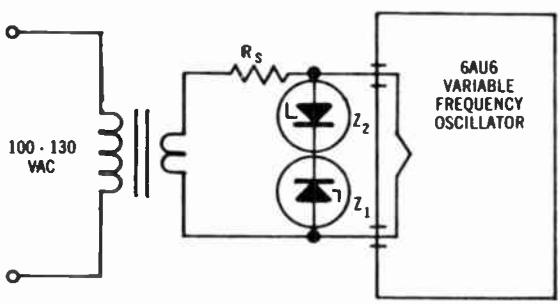


Fig. 2. An AC regulator application for filament temperature stabilization. The component values are worked out in the text.

rent. The peak current through this resistor at minimum line voltage is:

$$I_{Rspeak} = \frac{I_{Rs}}{\rho} \text{ or } \frac{.286}{.6} = .477 \text{ A}$$

where; I_{Rspeak} = peak current in R_s

ρ = ratio of average to peak current (0.6)

The resistance of R_s is:

$$R_s = \frac{V_{inmin}(1.414) - V_z}{I_{Rspeak}}$$

where; V_{inmin} = min. supply voltage (10.8)

V_z = Zener voltage

I_{Rspeak} = peak current in R_s

solving;

$$R_s = \frac{10.8(1.414) - 5.6}{.447 \text{ Amp.}}$$

$$R_s = 20.3 \text{ Ohms}$$

The maximum Zener current will occur when the line voltage is highest. Since the diode is conducting heavily, ρ must once again be considered. The maximum Zener current can be determined from the following equation:

$$I_{zmax} = \frac{V_{inmax}(1.414) - V_z}{R_s} (\rho)$$

where; V_{inmax} = maximum input voltage (14.0V)

V_z = Zener voltage (5.6V)

R_s = series resistor (20.3 Ohms)

ρ = ratio of average to peak current (0.6)

solving;

$$\frac{19.8 - 5.6}{20.3} 0.6$$

$$I_{zmax} = 0.42 \text{ Amp.}$$

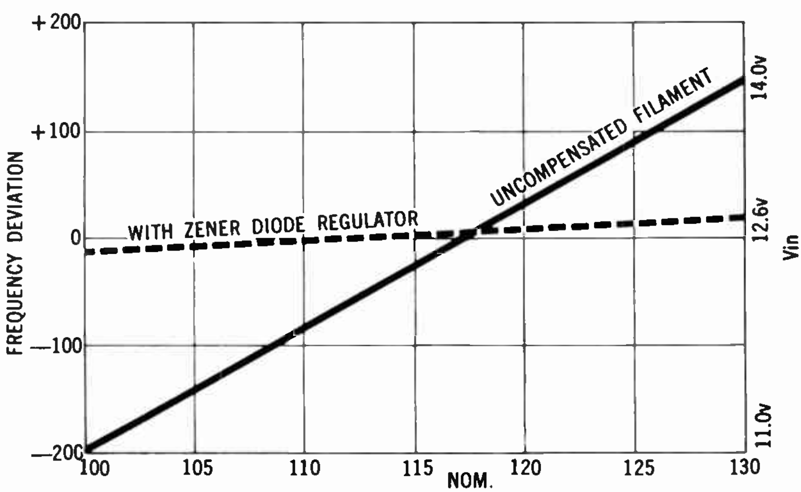


Fig. 3. Line variations plotted versus frequency shift.

Each diode will dissipate half the heat generated in the circuit. Therefore the power dissipating ability of each diode is equal to:

$$P_z = \frac{V_z \times I_{zmax}}{2}$$

where; P_z = power dissipated by each diode

I_{zmax} = maximum Zener current

solving; $P_z = \frac{2.35}{2} = 1.18$ Watts

Therefore the next larger size diode, a 3.5 Watt package, would be used.

The power dissipated by resistor R_s may be computed from the following:

$$P_{Rs} = R_s(I_{zmax} + I_{Lmin})^2$$
$$P_{Rs} = 20.3 \times (0.68)^2$$
$$P_{Rs} = 9.38 \text{ Watts}$$

Thus in Fig. 2, two International Rectifier 3.5 Watt Zener diodes would be used [1N1590 (5.6 V.)] in conjunction with a 20 Ohm, 10 Watt dropping resistor.

To illustrate the improvement in frequency stability, the curve of frequency versus line voltage was plotted and is reproduced in Fig. 3. It can be seen that the frequency shift is less than 50 Hertz in either direction from the mean frequency.

The regulator could also be placed in the primary circuit as shown in Fig. 4. Connected in this manner the diode is able to regulate each of the secondary windings. This circuit, however, requires higher voltage units and relatively higher power ratings, and is therefore more costly.

The filament regulation technique which has just been described in detail can also be applied to mobile installations where the voltage variation is more extreme than is usually encountered with ac power lines. The equations given in Chapter 2 would be applicable.

Where power consumption is a prime consideration on AC power circuits, an inductor or capacitor can be used as a ballast device. Such a circuit is shown in Fig. 5. The capacitor is selected to have a reactance substantially the same as a calculated value for R_s at the power line frequency — 60 or 400 Hertz.

Simple Regulated Power Supply

Simultaneous voltage regulation and DC rectification is achieved in the power supply shown in Fig. 6. This device could be used to supply regulated DC for a preamplifier and similar devices where extreme regulation is not a design necessity. The Zener diodes are selected to act as clippers in connection with R, thereby limiting the voltage to which capacitor C is charged. This provides control of a larger amount of power than is ordinarily handled by Zener diodes of this rating. At the same time the diodes act as rectifiers in this circuit.

Simple Oscilloscope Calibrator

Often a stable calibrating voltage source is required that is independent

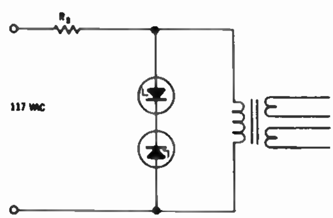


Fig. 4. A primary AC regulator.

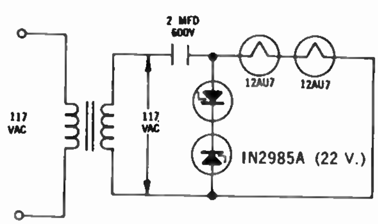


Fig. 5. An inductor or oil-filled capacitor can be used as a ballast in place of R_s.

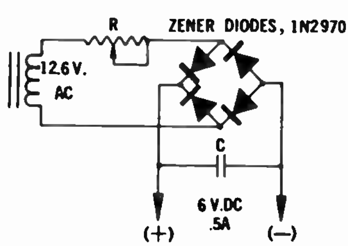


Fig. 6. A voltage controlled supply.

of line voltage variations. The circuit shown in Fig. 7 can be easily incorporated in an oscilloscope for production testing. A selected 10 Volt diode is used to provide a calibration of one Volt per division. Operation of the circuit is self explanatory, however it should be pointed out that the square wave will be non-symmetrical. Improvement in symmetry may be obtained by having an AC source voltage equal to 10 times Zener voltage. The appropriate limiting resistor and a Zener with a voltage in excess of 7.5 Volts will yield a relatively symmetrical waveform.

Saturable Reactor Override

An interesting application involves surge protection with Zener diodes. The load was a large oven, with heating elements fed from a three-phase source, through a saturable reactor. The reactor provided the control necessary for maintaining the oven at a preset temperature. However, a mode of limiting the initial surge of current into the cold heating elements was necessary to protect the components from burning out.

The system used involved the sensing of the line currents with current transformers, and employing the almost rectangular reverse voltage vs. current characteristics of a Zener diode to provide the required protection.

Note in Fig. 8 that the output of the three current transformers are connected in delta; then rectified to provide a DC output current which is proportional to the vector sum of the three ac line currents.

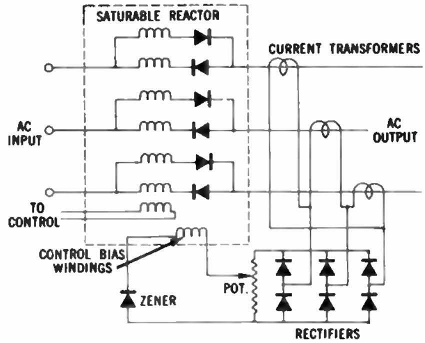


Fig. 8. Schematic diagram of control circuit.

By providing a fixed load resistance, an output voltage is obtained which is also proportional to the AC line current. The Zener diode connected across a portion of this fixed load will act as a switch and will perform the function of limiting the AC by means of the saturable reactor, or provide a warning system by energizing a relay.

In this particular application, the secondary current ratings of the current transformer were selected at ¼ Ampere. The DC output from the three-phase bridge rectifier in the circuit would be approximately ½ Ampere at full load.

The value of the potentiometer was selected at 35 Ohms to provide a full

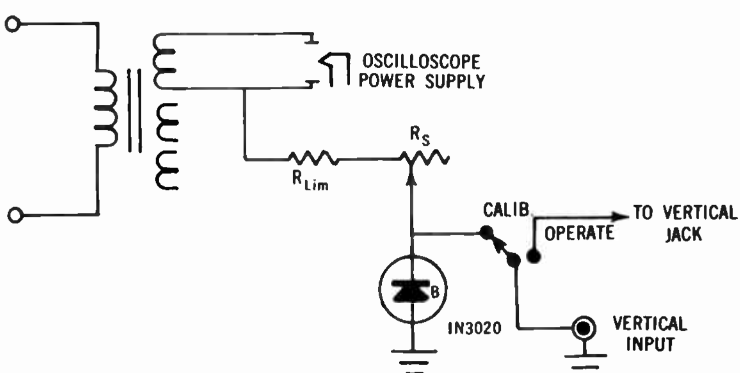


Fig. 7. A simple oscilloscope voltage calibrator.

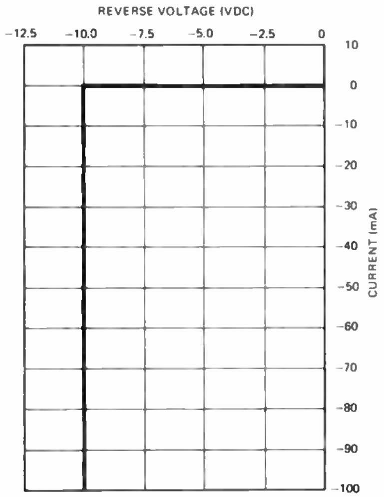


Fig. 9. 1N3020 Diode Breakdown Characteristics.

load voltage of 17.5 Volts. The rating of the potentiometer was specified as 25 Watts for a wide safety margin in operation. The rectifier requirements can be determined from the output voltage and current values. It can be seen that six International Rectifier silicon diodes such as the 10D Series can be utilized.

To provide the optimum characteristics from the Zener diode, a voltage of from 8 to 18 Volts should be selected. In this application an International Rectifier type 1N3020 diode was chosen, which provides a V_z at approximately 10 Volts. An almost rectangular curve is exhibited, as shown in Fig. 9. Note that at voltages below the breakdown point, the leakage current is minute.

The current through the Zener diode flows through a separate bias winding on the saturable reactor. This bias acts to cause the reactor to absorb more voltage, thereby tending to reduce the amount of load current. Thus, the AC line current will remain at the set maximum value until outside conditions cause a change in the operation. The potentiometer provides a simple adjustment of the current limiting point, with this system providing a minimum current limiting point at approximately 57% of full load current. By selection of the potentiometer value, the range of adjustment can be made to suit the application.

Testing Zener Diodes

One of the most important AC tests is to determine Zener diode impedance, for Z_z is directly related to the regulating ability of the diode. In this respect the Zener diode is comparable to the capacitor; as its reactance decreases so does the change in potential across the terminals.

The value of Z_z varies with junction current and diode size, as examination of the Zener diode electrical specifications will reveal. As an example, it can be seen that the 27 Volt, ¾ Watt diode has a dynamic resistance of 82 Ohms while the 1 Watt device has only 35 Ohms.

It becomes immediately obvious, that for repetitive results dynamic resistance measurements must be made with a specific set of conditions. International Rectifier measures production diodes in a test device similar to Fig. 10. The diode current is set at 20% of I_z max.

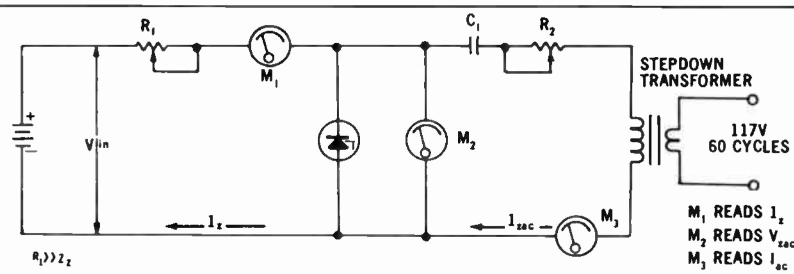


Fig. 10. A test jig for rapidly determining Zener diode dynamic resistance.

by adjustment of potentiometer R1, in conjunction with V_{in}, and diode current is indicated by M1.

A source of AC is also required for impedance measurements. This voltage, from the 60 Hertz source appears across the divider consisting of R2 and Z_z. Variable resistor R2 is set to provide a current value which is 10% of I_z as indicated by M3.

When these conditions have been established, the AC developed across the junction can be read on M2, a "Millivolter" or another sensitive AC voltmeter. When V_{zac} and I_{zac} are known, the dynamic resistance may be easily computed by using the equation;

$$Z_z = \frac{V_{zac}}{I_{zac}}$$

Another informative test setup is shown in Fig. 11. Connected in this manner the oscilloscope will display the dynamic operating characteristics of the diode.

The presentation is useful in schools and training programs, for both the forward and reverse characteristics may be effectively demonstrated (Photo 4).

The horizontal axis of the oscilloscope represents the voltage developed across the diode junction, while junction current is shown on the vertical axis. If desired, the forward conduction trace can be eliminated by connecting a gate diode in series with the 22 Ohm resistor (Photo 5).

This display can also be used for production testing and sorting by employing an oscilloscope with DC coupled x and y amplifiers. The horizontal base line can be calibrated in Volts/cm., while the vertical axis would represent 10 or 100 mA/cm. Further, a screen overlay could be used to indicate tolerance. Defective and intermittent diodes become immediately obvious, using this technique.

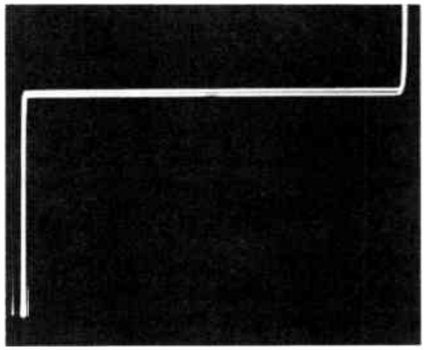


Photo 4. Oscillographic display for the test setup shown in Fig. 11. Note the sharp break at the Zener point and the near-vertical current increase.

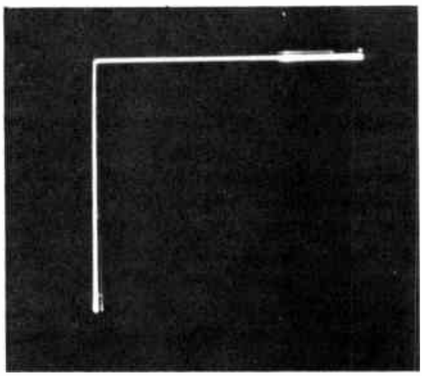


Photo 5. For demonstration, the forward conduction line can be eliminated by connecting a silicon rectifier in series with the 22 Ohm resistor in Fig. 11.

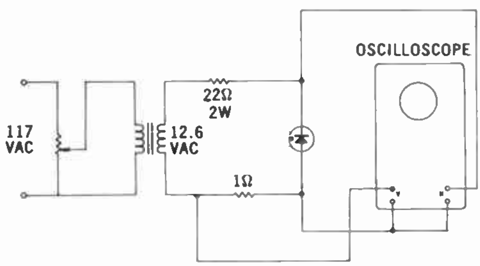


Fig. 11. A dynamic curve display for testing and demonstrating Zener diode characteristics.

Zener Diode Regulated Utility Power Supply described in Chapter 6.

DC Applications

For Power Supply Applications, the Zener Diode is an Excellent Reference to which Input and Output Variations may be Compared and Corrected

The small size, ruggedness, and dependability of Zener diodes make them ideal devices for compact lightweight regulated power supplies. They can perform functions of referencing and regulating DC voltages on an equal with vacuum tubes, and in many applications — such as high current, low voltage supplies — the Zener diode is superior to vacuum tubes and saturable core reactors.

The versatility of the Zener diode cannot be overemphasized. These solid state regulators can be used to control almost any load situation that might be encountered, and with a minimum of components.

Basic Regulator

In its simplest form the Zener diode regulator consists of a source of DC, a series resistor, and a shunt connected diode as shown in Fig. 1.

The value of R_s is determined by the load requirements. If R_s is too large, the diode will be unable to regulate at large values of I_L. Conversely if R_s is too small the diode dissipation rating may be exceeded at low I_L values. The optimum value for R_s can be calculated from the equation;

$$R_s = \frac{V_{inmin} - V_z}{I_{Lmax} + 0.1\, I_{Lmax}}$$

where; V_{in} = supply voltage
V_z = Zener voltage
I_L = load current

note: $0.1I_L$ is an empirical figure to insure diode regulation at high load currents.

When R_s is known, the maximum diode dissipation can be calculated by applying the following equation;

$$P_d = \left(\frac{V_{inmax} - V_z}{R_s} - I_L\right) V_z$$

where: P_d = Zener diode dissipation
R_s = series resistor in Ohms

Filter Effect

The Zener diode will respond to ripple in much the same manner as it does with slower voltage variations. It will tend to average out these fast changes also. Thus the Zener diode, by virtue of its low dynamic impedance, reacts much the same as a filter capacitor. The ripple factor, or ripple gain, can be stated as;

$$\%\ \text{Ripple} = \frac{e_o}{e_{in}} \times 100\%$$

where: e_{in} = input voltage at ripple F.
e_o = output voltage at ripple F.

Ripple factor is measured at 120 Hertz,

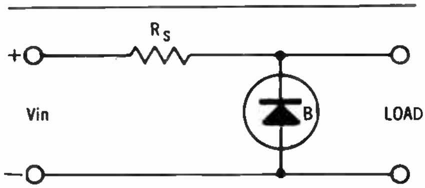


Fig. 1. Basic Zener diode regulator circuit.

the lowest frequency usually encountered in power supplies.

Additional power supply filtering can be obtained by connecting a Zener diode (with a value of V_z equal to the ripple trough) across the load. This will provide a large reduction of ripple impressed on the DC. In most circuit applications the Zener diode will be as effective as several thousand microfarads of capacity.

Several of the circuits discussed in this chapter may be driven by simple half-wave rectifiers in conjunction with less than 100 mfd. filter capacitors, due to this filter effect.

Regulation

A perfect regulator would provide a constant output with any change in input voltage or load current. Unfortunately this utopian condition cannot exist in practical applications, although Zener diode regulated supplies come close to the ideal situation.

The ability of the regulator to absorb input voltage variations is called the Input Regulator Factor and can be expressed as;

$$\gamma = \frac{\Delta V_o V_{in}}{\Delta V_{in} V_o} \text{ with } I_L \text{ constant}$$

where:

γ = Input regulation factor

Δ = change in voltage

Thus a change in input voltage should produce a minimum of output voltage variation.

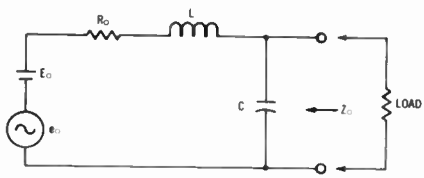
Changes in the load current will also disturb the regulation. Any increase in load current will produce a voltage drop across series elements and a reduction in V_o. In a regulated supply the change in output voltage is compensated by one of the circuit elements. The ability to compensate changes in I_L can be stated as;

$$\sigma = \frac{\Delta V_o R_L}{\Delta R_L V_o}$$

where: σ = load regulation factor

R_L = load resistance

This, and the preceding factor, can be termed a figure of merit for they express the regulating ability of the supply. For best performance these figures of merit should be minimized to as low a value as possible.



WHERE

E_o = OPEN CIRCUIT OUTPUT VOLTAGE
e_o = OPEN CIRCUIT RIPPLE VOLTAGE
R_o = STATIC RESISTANCE
L = APPEARANT INDUCTANCE
C = APPEARANT (AND REAL) CAPACITY
Z_o = DYNAMIC OUTPUT IMPEDENCE

Fig. 2. Equivalent circuit of output impedance.

Output Impedauce

The output impedance of a regulated power supply should be low; ideally it would be zero. In actual practice the load "sees" the dynamic output impedance as a combination of L, C, and R_o as illustrated in the equivalent circuit Fig. 2. The DC output resistance, R_o, is the ratio of change in output voltage to change in output current with the input voltage held constant, and is expressed by the equation;

$$R_o = \frac{\Delta V_o}{\Delta I_L}$$

The dynamic output impedance, Z_o, is the ratio of the AC components of the output voltage and current when the load is varied sinusoidally with the input voltage held constant. The output impedance can be determined by applying an audio signal to the output terminals and measuring the voltage developed across the power supply when the signal current is known.[1-2] Such a test circuit is shown in Fig. 3. The components C and R prevent the source ripple from masking the output impedance measurements.

Power supply output impedance can be expected to increase with decreases in the load current, for impedance reduction is a function of the current passing tube (or transistor) transconductance which is determined by the current flow. A typical output impedance curve is shown in Fig. 4, for two I_L conditions.

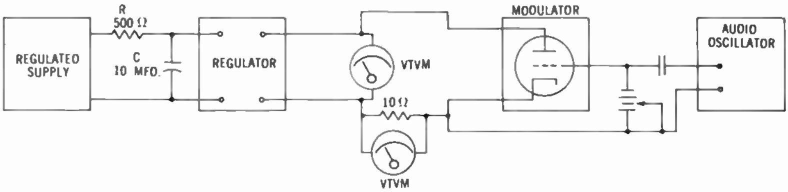


Fig. 3. Test circuit for measuring output impedance of a voltage regulator.

Obtaining Other Voltages

Occasionally it is necessary to regulate a voltage not normally obtainable from Zener diodes. As an example, it is quite permissible to series connect several Zener diodes to achieve a desired regulation level, as shown in Fig. 5. The diodes need not have equal breakdown voltages since the arrangement is self equalizing. However, the power handling ability of each diode should be the same. In addition, the current ranges should be similar or the loads so arranged to avoid damaging any of the diodes. Consideration must also be given to the fact that in the series configuration dynamic impedance values are additive.

The Zener diode can also be used as a series voltage dropping device. In the configuration shown in Figure 6, the Zener serves the designer by dropping the input voltage from 28 Volts to 22.4

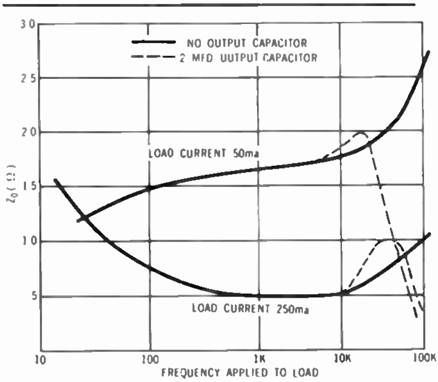


Fig. 4. A typical output impedance curve for a regulated power supply.

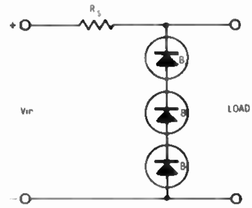


Fig. 5. Zener diodes can be series connected to obtain higher values of V_z.

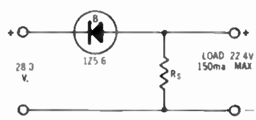


Fig. 6. A series configuration can be used when small voltage drops are required.

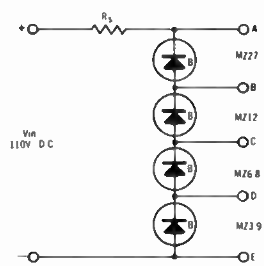


Fig. 7. A Zener diode voltage divider system.

Volts. The Zener does not provide any regulation in this configuration as it is able to do in the shunt configuration. This configuration is best used where only a small drop in voltage is required, since the diode would be a 5.6 Volt, low temperature coefficient type. The entire load current, plus R_s current, must flow through the series element, therefore care must be taken to make sure that the load current is never excessive.

A group of Zener diodes may be used as a divider to obtain several regulated voltages simultaneously. Such a configuration is shown in Fig. 7. This circuit could be used as a meter calibration source for checking linearity at several points on the meter scales. Four diodes are used, as shown, to supply 10 possible voltages as follows,

Voltage	*Terminals*
3.9	D-E
6.8	C-D
10.7	C-E
12.0	B-C
18.8	B-D
22.7	B-E
27.0	A-B
39.0	A-C
45.8	A-D
49.7	A-E

In applications where it is necessary to deliver a regulated voltage lower than normally available with Zener diodes, the circuit shown in Fig. 8 may be used. Two Zener diodes are used and the regulated difference potential is utilized (Fig. 8a). Temperature compensation in this configuration would yield about 1.5mV/°C in the output voltage.

Another technique used to provide low values of regulated voltage is to use a number of forward biased diodes in much the same manner as a Zener diode (See Figure 8b). However, this technique is limited to 0.7 Volt steps in the case of silicon diodes and 0.3 Volt steps for germanium diodes. It does however allow the design of a grounded point in the reference supply.

When a source of well regulated, adjustable voltage is required, the circuit shown in Fig. 9 may be employed. The dynamic regulation is excellent for the first diode tends to act as a preregulator. Any combination of diodes may be used to achieve the desired range of output voltage.

Shunt Transistor Regulators

The voltage control ability of the Zener diode can be increased considerably if the diode is used to control the operating point of a transistor (Fig.

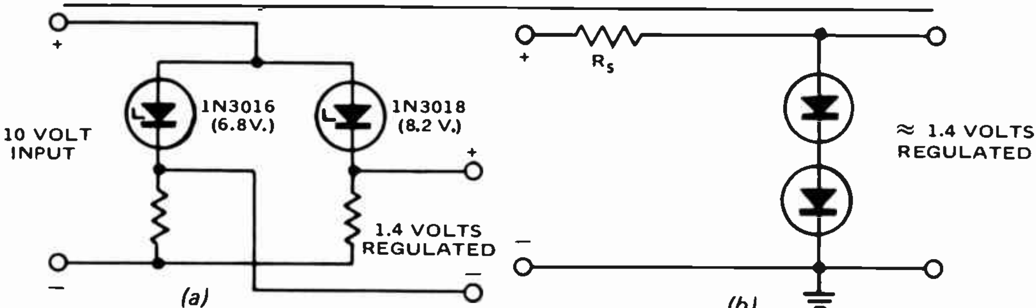


Fig. 8. Low power, low voltage supplies.

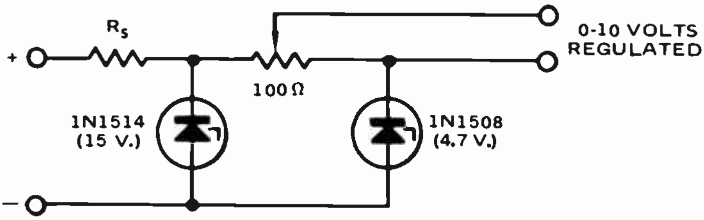


Fig. 9. A well regulated adjustable supply.

10), or a group of transistors. In the shunt configuration only the transistor base current flows through the Zener diode. In addition to increasing the power handling ability, the regulating factor will be improved by the current gain of the transistor.

A cascade shunt regulator is shown in Fig. 11. The Zener diode controls the base potential of Q1, which functions as an emitter follower and current amplifier. The voltage E_{CB} of Q1 determines the bias on Q2, the shunt regulator. This circuit would be applied where an extremely large current variation is encountered. Several shunt connected transistors would be used for Q2.

The transistor shunt regulator can be adapted to supply voltages lower or higher than V_z. Fig. 12 illustrates a shunt regulated supply for outputs greater than the Zener voltage. Neglecting R_s, or assuming V_{in} to be at the junction of R_s and R1, the output voltage will be determined by the ratio of R1/R2;

$$V_o = \frac{R1+R2}{R2}$$

If, as an example, R1 and R2 are the same value (as in Fig. 12), the output voltage will be twice V_z. Resistor R3 compensates for variations in the supply to the regulator. The exact value can be determined empirically by substituting a potentiometer for R3. Vary the input voltage, while adjusting R3 for minimum V_o variation. Excessive resistance at R3 will cause overcompensation, i.e. V_o will drop as V_{in} increases. Current amplification causes capacitor C1 to appear as a large electrolytic capacitor across the output terminals. The ripple will be less than 10 mV, when the regulator is supplied by a full-wave rectifier with 20 mfd. capacity.

When voltages lower than V_z are required, the circuit in Fig. 13 is quite useful. The transistor collector/emitter potential is regulated at V_z. The setting of potentiometer R2 determines the regulated output voltage. This potentiometer should be as low resistance

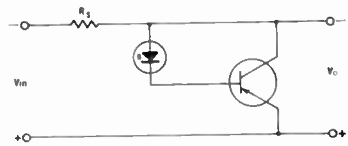


Fig. 10. In this shunt regulator circuit a 1 Watt Zener diode can control 10 Watt load.

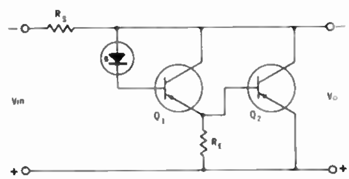


Fig. 11. A cascade shunt regulator.

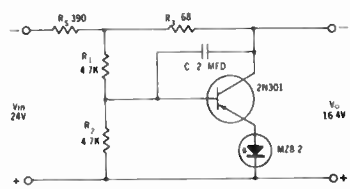


Fig. 12. A shunt regulator for voltages higher than V_z.

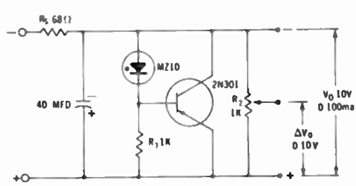


Fig. 13. Shunt regulator for voltages lower than V_z.

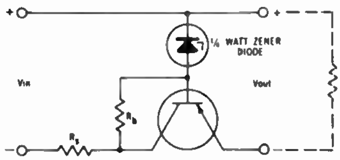


Fig. 14. A transistor regulator may be used to regulate V_o. Base bias is held constant by the Zener diode.

as possible compatible with the load requirement, to minimize voltage variations due to load changes.

Series Transistor Regulators

Although the Zener diode finds application in high current power supplies, its power handling ability is limited by the available dissipation. It would be quite impossible to use the circuit shown in Fig. 1 to control—say 1,000 Watts—unless the diode was of rather immense proportions.

In these situations, it is customary to use a transistor (or group of transistors) as the series element in conjunction with a Zener diode controlled reference. Such a circuit is shown in Fig. 14. The source, V_{in}, is applied to the Zener diode through R_s and the base bias resistor R_B, establishing a reference base voltage with respect to the positive terminal. In effect the transistor functions as an emitter follower. Thus the emitter voltage is held within a few tenths of a Volt of the base potential, which is determined by the Zener diode.

The transistor acts as a series element to absorb voltage variations. Since the entire load flows through the transistor it must be able to dissipate the power absorbed by the junction. The power handling ability of the supply shown will be determined entirely by the number of transistors used and the ability to remove heat from the junction(s). Zener current is reduced to a small fraction of the original (Fig. 1) by employing the transistor.

Typical values for Fig. 14 might be R = 22 Ohms, R_b = 1K, in conjunction with a 4.7 Volt, ¼ Watt Zener diode. To check regulation a fixed load was attached to the output terminals and the input voltage varied between 0 and 30 Volts. The resulting curve is shown in dashed line in Fig. 15. As a check of output regulation, the input was set at 16 Volts and the value of I_L varied betwen 10 and 130 mA. The output regulation is shown by the solid line in Fig. 15.

To illustrate the improvement by using low dynamic resistance diodes, a 3.5 Watt International Rectifier 1N1589 (4.7V.) diode was substituted for the ¼ Watt unit. The increase in regulation was approximately 10:1.

This circuit can deliver a variable regulated output by installing a 1K potentiometer across the Zener diode and connecting the movable arm to the base of the transistor. The regulation factor will be reduced slightly due to the shunting effect of the potentiometer.

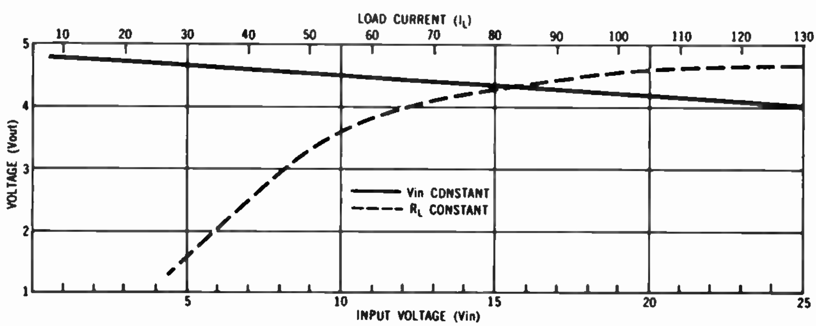


Fig. 15. Regulation with changes in I_L and V_{in} for the series regulator, Fig. 14.

Somewhat better regulation will be obtained if various Zener diodes are switched into the circuit for different output voltages.

High Current Regulator

The design of a high current supply will be dictated by the load requirements. The unregulated source must be able to deliver between 3 and 6 Volts more than the required output with full load conditions. In all cases the dissipation ability of the current passing transistors must kept in mind.

The circuit shown in Fig. 16 illustrates this point. It consists of two cascade current amplifiers controlling four parallel connected current passing transistors.

If the output requirements did not exceed 400 mW. ($V_{in} - V_o$), the load could be connected across R2. Transistor Q2 functions as a current amplifier to increase this rating to 25 Watts maximum dissipation. In theory, at least, the last two stages could regulate four Amperes at 24 Volts:

$$I_{Lmax} = \frac{P_{dmax}}{E_{in} - E_o} =$$

However, the heat removal system is seldom efficient enough to operate the transistor at maximum ratings safely. The losses in a single junction will cause low efficiency, also.

For high current applications several parallel connected transistors should be used as the current passing element. In Fig. 16 this is accomplished by using four 2N1136 transistors which will easily handle 150 Watts, and a maximum of 240 Watts with ideal heat sink conditions.

A potentiometer is incorporated in Fig. 16 to adjust the output voltage. As the output voltage is reduced, the drop across the transistors increases and the maximum output current must be reduced accordingly to avoid exceeding P_{dmax} of Q3.

This condition can be avoided by incorporating a separate supply for the Zener regulator. A Variac or adjustable autotransformer would be used at the AC input of the high current supply and would be mechanically ganged to the voltage adjustment potentiometer. With this current configuration the collector/emitter voltage of the current passing transistor would not exceed 3 to 6 Volts, thereby providing a great increase in output current at low voltages.

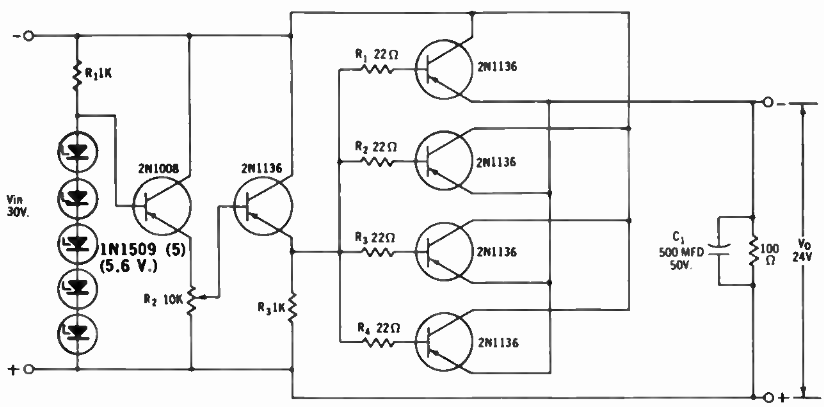


Fig. 16. A Zener diode controlled high current regulator.

High Voltage Supplies

Although Zener diodes are not usually associated with high voltage supplies (300 V), they may be used in conjunction with transistors having a high E_{ce} breakdown.

Such a transistor is the 2N1136B, which has an E_{ce} of 80 Volts and features 60 Watts dissipation. This transistor is used as the series element in the negative lead of a high voltage regulated supply[3] (Fig. 17). The difference between input and output voltage is 80 Volts with minimum load conditions. At the 200 mA maximum power supply rating, the transistor dissipates something less than 16 Watts.

In operation, an increase in the load current causes a decrease in the bias on Q2 through the compound connected pair (Q3, Q4). The action of Q2, in turn, increases the bias on Q1, decreasing its resistance and thereby returning the output voltage to its nominal value. A decrease in load current reverses the action just decribed maintaining the output voltage at the correct value.

The International Rectifier 1N2970, 10 Watt Zener series is ideal for use in vacuum tube regulated high voltage supplies. The 10 Watt Zener is superior to a gas discharge tube, for it is not subject to erratic voltage changes, but which may be compensated in critical applications with a thermistor. In addition, these solid state regulators have no ignition potential and therefore are not subject to relaxation oscillations in high capacity circuits.

A simple regulator is shown in Fig. 18, using a 6080 current passing tube. The 6080 resistance is controlled by a 12AX7 DC amplifier, which is stabilized by an IN3005A regulator.

The control grid (7) is held constant by the Zener diode. If the load increases, the voltage across the output terminals will tend to drop. This causes a reduction in the plate voltage on the right-hand triode. This change is coupled to the left-hand section as an increase in bias, which decreases the bias on the current passing tube. With these circuit condition changes, the output voltage returns to its nominal value.

Since this closed-loop is constantly working, the voltage does not deviate far from the design center. In operation this circuit will regulate within 0.2 Volts from no-load to full-load. A 10% line voltage variation will produce approximately 0.1 Volt output varition.

Current Regulators

A simple constant current regulator can be constructed by using a transistor as a variable series resistor. The circuit shown (Fig. 19) is optimized for a

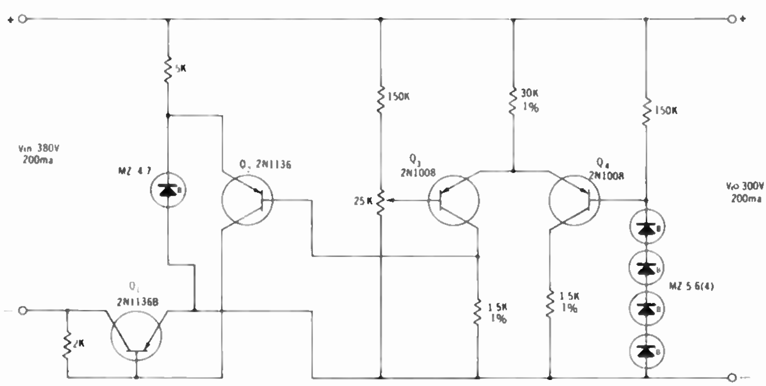


Fig. 17. 300 Volt, 200 mA solid state regulated supply.

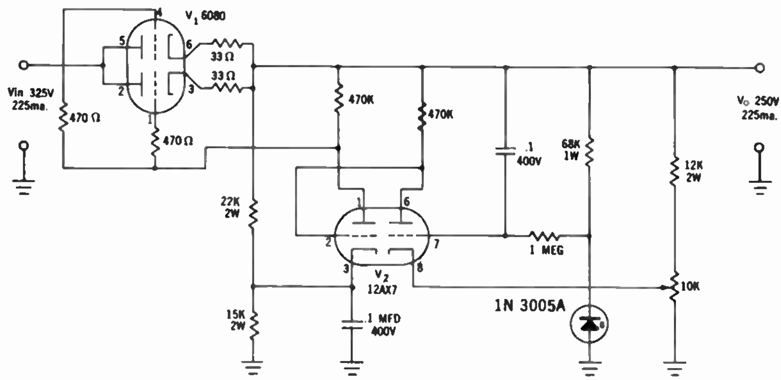


Fig. 18. Vacuum tube regulated supply using the 1N3005A Zener diode.

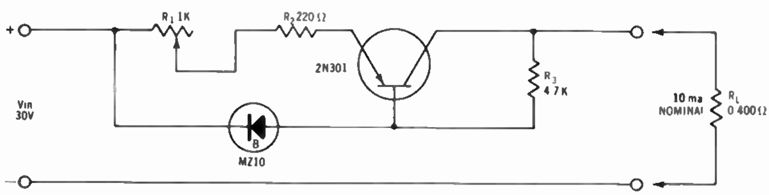


Fig. 19. A constant current regulated supply. The values shown are nominal for a 10 mA max. load current.

10 mA load current with the values specified. In this configuration two circuit paths exist; one through the regulator diode 1N3020 (10 V) in series with the bias resistor. Current also flows through R1, R2, and the transistor. Any change in the current through R3 causes a change in base bias. The transistor, in turn, changes resistance to correct the current flow. The net result is that for every change in R1 there is an equal and opposite change in the transistor junction resistance.

In operation the current will remain constant within 10% from short circuit conditions to a maximum load of 400 Ohms. A load resistance curve for this type of regulator is shown in Fig. 20.

A Utility Supply

An adjustable voltage regulated power supply is a useful piece of test equipment for the laboratory or repair shop. It may be used as a voltage source for calibrating meters or to power experimental transistor projects.

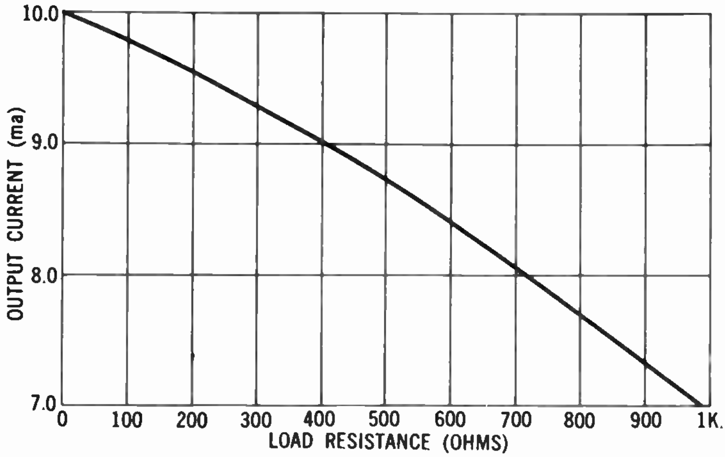


Fig. 20. A current regulator curve optimized at 10 mA.

Zener diode regulation provides an immunity to line voltage variations making the supply an excellent substitution for battery bias packs employed when measuring stage gain and aligning prototype if amplifiers.

Such a versatile supply is shown in Fig. 21, and the accompanying photographs. A multi-ratio transformer supplies an adjustable voltage to a bridge rectifier system, such as International Rectifier's 18DB2A molded rectifier bridge. The rectified output will be either 15, 30, or 45 Volts, depending on the setting of S1. This voltage is stabilized by a simple regulator system, similar to that shown in Fig. 14. On the 10 Volt range, a 1N3020 (10V) Zener diode is connected between the base of Q1 and the positive buss. Thus a regulated voltage equal to V_z appears across the potentiometer (R1) in the emitter circuit of Q1. A portion of this voltage is fed to the bases of the parallel connected current passing transistors (Q2 and Q3). Since these stages function as emitter followers, this voltage appears across the output terminals. It can be seen that the output voltage will be determined by the setting of R1.

On the 20 Volt range, a 1N3028 (22 V) replaces the 1N3020 (10V) and the maximum output voltage increases to 22 Volts.

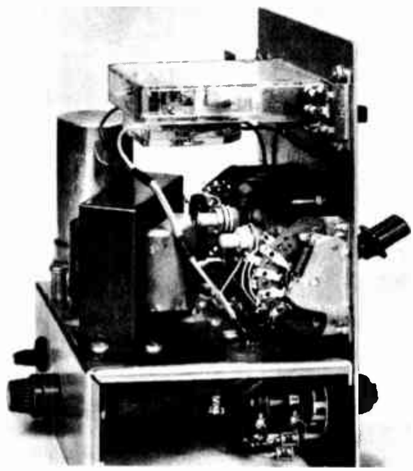


Photo 1: Interior Side View Of Power Supply.

On the 30 Volt range the two diodes are series connected, which allows a maximum output voltage of approximately 32 Volts. On each range the appropriate tap on the secondary of T1 is selected to minimize the voltage drop across Q2 and Q3. This, of course,

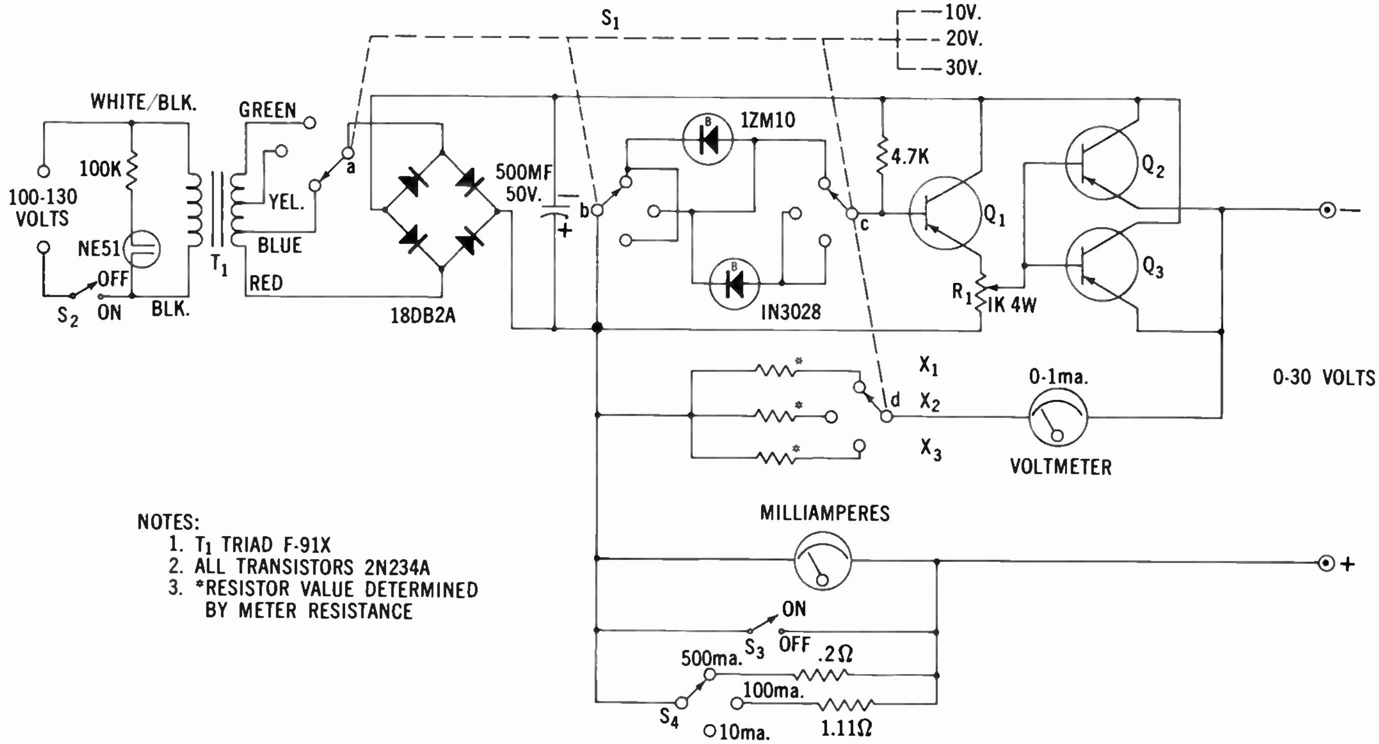
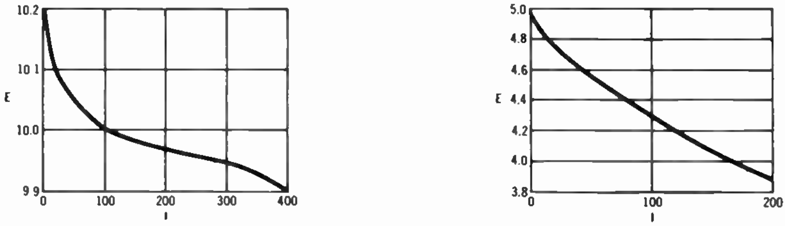
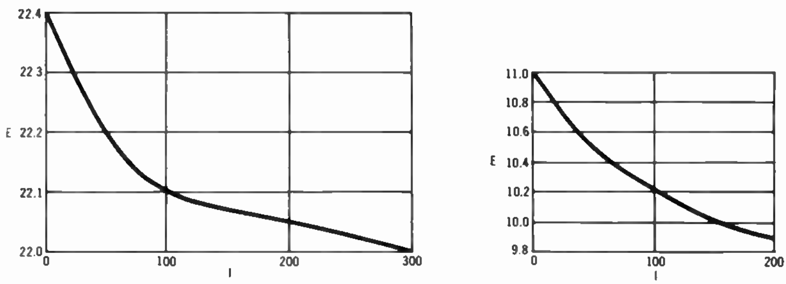


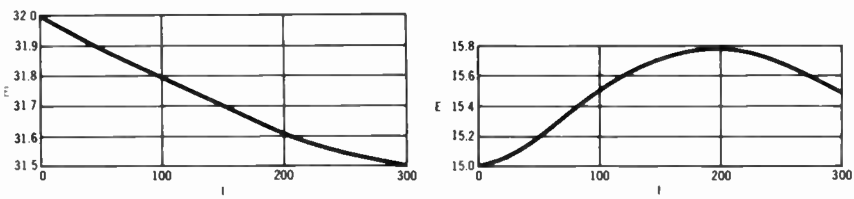
Fig. 21. A simple regulated supply suitable for meter calibration or transistor circuit testing.



10 Volt range, maximum and half voltage (adjustment potentiometer settings).



20 Volt range, maximum and half voltage (adjustment potentiometer settings).



30 Volt range, maximum and half voltage (adjustment potentiometer settings).

Fig. 22. Regulation curves for the power supply shown in Fig. 21.

reduces the dissipation and collector to emitter voltage. A fourth section of the switch selects the appropriate multiplier resistor for the voltmeter.

A current meter is connected in series with the positive terminal. A three position switch (S4) selects the proper current range. An additional switch (S3) shorts the meter when not in use as a safety precaution against overload.

Although not shown on the schematic, the unit includes a current regulator for special applications. The circuit is essentially the same as the one shown in Fig. 19.

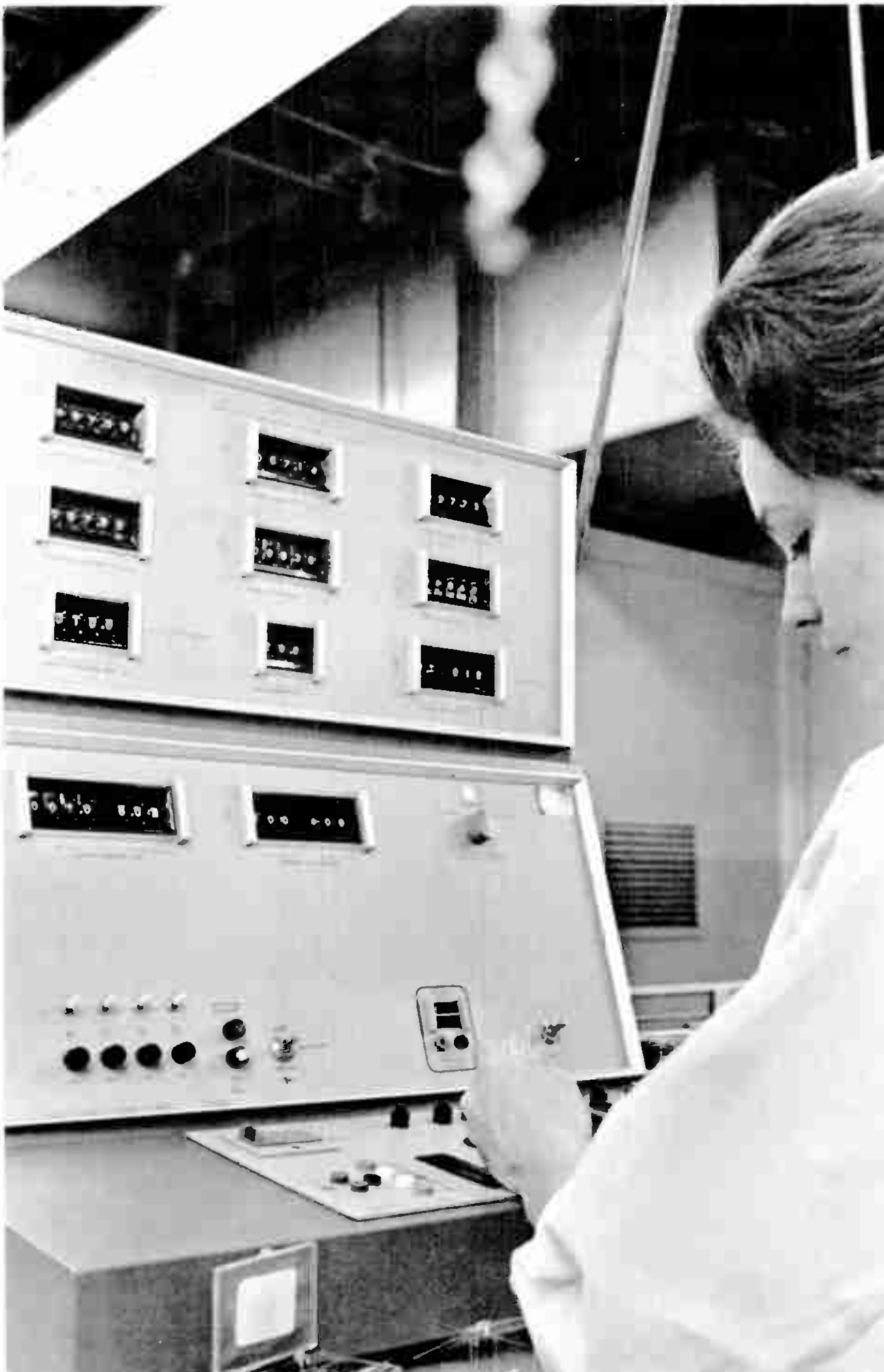
The multiplier resistors for the voltmeter, and the shunt resistors for the milliameter, must be determined for the particular meter used in the instrument.

The load regulation is relatively good at full scale, on each range, as indicated by the curves illustrated in Fig. 22. However, when the voltage is reduced (by adjustment of R1) the overall current gain is reduced and the regulating ability suffers accordingly.

The input regulation is excellent at any output voltage setting, and is in the order of 0.1% from less than 100 Volts to more than 130 Volts.

REFERENCES

1. J. H. Hershey, "Dynamic Impedance of Regulated Power Supplies," Bell Lab. Rec., vol. 27, June 1949, p. 216.
2. "Standards on Television: Methods of Measurement of Electronically Regulated Power Supplies, 1950," Proc. IRE, vol. 39, Jan. 1951, p. 29. See Par. 3.5.2 and fig. 3.
3. Courtesy of Bendix Application Notes.



Rating and sorting of Zener diodes.

Audio and RF Applications

Zener Diodes Are Useful As Bias And Coupling Elements In Vacuum Tube and Transistor Circuitry

The Zener diode may be used in audio and rf amplifier circuits wherever a source of stable voltage is required. In certain circuits it may prove useful as a switch or limiting device.

Bias

The Zener diode is an ideal bias element in amplifier applications. Its function is similar to a battery, however, it requires no maintenance, and cannot damage equipment due to deterioration.

The cathode circuit of a class AB_1 amplifier is shown in Fig. 1. Bias for the stage is developed by cathode current flow through resistor R. Any voltage or current variations will cause a shift in the tube operating point. Although the circuit is degenerative, the effect can create problems where stability is a criteria.

A Zener diode may be used in this application to provide a stable operating point. Cathode current flows through the Zener diode but, unlike the resistor, a stable bias voltage is developed across the diode and is equal to V_z. Variations in static conditions will change R_z, thereby maintaining the correct bias on the tubes.

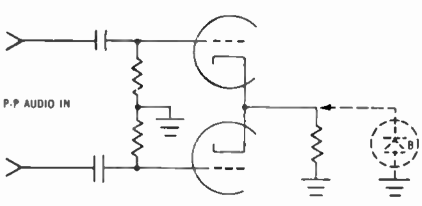


Fig. 1. Push-Pull class AB_1 stage illustrating how the Zener diode is used to provide fixed bias.

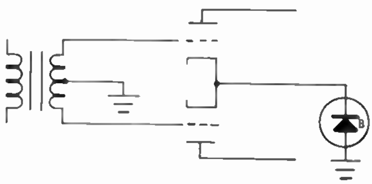


Fig. 2. A Zener diode bias Class B stage.

The application of Zener diodes as bias elements is particularly useful in class B stages. In this type of amplifier the cathode current will increase several hundred percent during peak signal periods. This makes self bias impossible, unless a Zener diode is used as in Fig. 2. As in the previous illustration the Zener element establishes the correct operating point.

A single-ended class B linear rf amplifier is shown in Fig. 3. Ordinarily a grid bias supply would be used with these tubes. The grid current swing is so severe, however, that a regulated supply is mandatory. A Zener regulated supply could be used (as shown in Chapter 6) although the solution in Fig. 3 is far simpler and in many cases, less expensive.

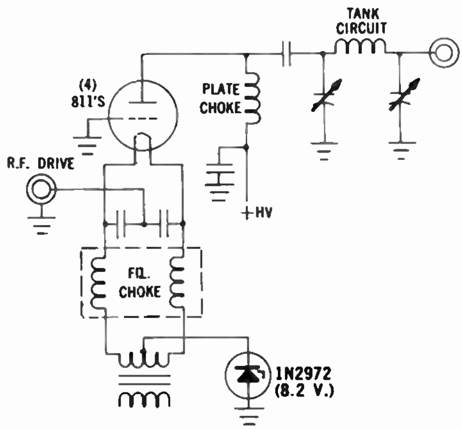


Fig. 3. A class B linear amplifier. The Zener diode determines the idling current.

In Class B, the diode must dissipate an average power at some point between the static and peak current. The amount of heating will be determined by the wave-form and duty cycle. For Class B audio applications the following equation will provide a liberal safety factor:

$$Pd = V_z \times 0.6 \text{ I peak}$$

A Zener diode in the cathode return may also be used in Class C rf amplifiers to provide a protective bias, should the drive signal fail.

In high power amplifier circuits the required dissipation may exceed that available in standard Zener diode packages. In these situations a transistor (or group of transistors) with suitable dissipation and breakdown ratings may be used as the current passing element. Such a circuit is shown in Fig. 4. The diode dissipation should not exceed one tenth the calculated dissipation of the transistor(s).

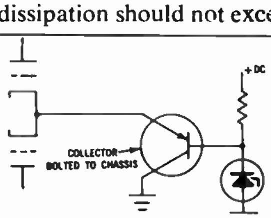


Fig. 4. A transistor may be used to regulate bias when employing low power Zener diodes.

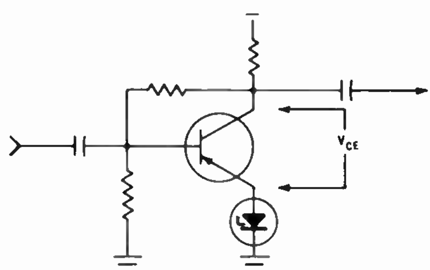


Fig. 5. The Zener element may also be used for emitter bias.

The Zener element is not confined to bias regulation in vacuum tube circuits. It may also be used with transistors wherever the ultimate in stabilization is required. Fig. 5 illustrates a single PNP amplifier stage, with a Zener diode replacing the usual emitter resistor. In this circuit the collector to emitter voltage (V_{CE}) is high enough to provide linear operation. Sufficient current must flow through Zener element to insure operation in the breakdown region.

In the preceding examples, the dynamic impedance of the Zener diodes will reduce degeneration to such a degree that bypass electrolytics are seldom necessary.

DC Amplifiers

A DC amplifier presents several knotty design problems, for each stage must operate at progressively higher potentials. If more than two or three stages are used, the associated circuitry becomes rather complex.

The Zener diode will provide a design solution, particularly in transistor circuits where the difference in collector and base potentials is relatively small.

A single section DC amplifier is shown in Fig. 6. As in Fig. 5, the bias is established with a Zener diode. The value of V_{CE} is dictated by linear operation with anticipated voltage changes. The collector current is selected near maximum beta. In this example, 10 Volts at 5 milliamperes was selected.

When these stages are cascaded, the voltage drop across R3 (due to I_{Z2}) must be approximately equal to V_E to ensure adequate bias for the following stage. Therefore V_{Z2} must have a breakdown equal to V_{CE} which in this case is 10 Volts. The current through both diodes should be sufficient to ensure operation in the breakdown region, beyond the knee of the curve. A 3 mA current through Z2 is adequate.

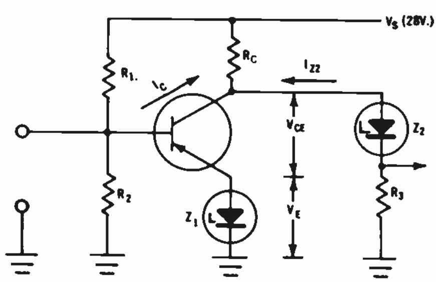


Fig. 6. As an inter-stage coupling element, the Zener diode allows direct cascading of stages.

With these parameters known, the value of V_E (and of course V_{Z1}) may be determined from the equation:

$$V_E = \frac{I_{Z2}}{I_c}(V_s - V_{CE})\left[\sqrt{\frac{I_c}{I_{Z2}}+1}-1\right]$$

where; V_E = emitter voltage ($V_E = V_{Z1}$)

I_{Z2} = current in Z2 (3 mA)

I_C = collector current (5 mA)

V_s = supply voltage (28 V)

V_{CE} = collector/emitter voltage (10 V)

solving: $V_E = 0.6 \times 18 \times 0.65$

$V_E = 7.02$ Volts

When V_E is known, the correct value for R_C may be obtained by applying the following equation:

$$R_C = \frac{V_s - V_{CE} - V_E}{I_C + I_{Z2}}$$

solving; $R_C = \frac{11}{.008}$

$R_C = 1{,}400$ Ohms

The resistor R3 may be determined, for both current and voltage are known:

$$R3 = \frac{V_E}{I_{Z2}} = 2{,}267 \text{ Ohms}$$

This figure includes the psysical resistance of R3, shunted by the base resistance of the following stage. The bias network consisting of R1 and R2 (Fig. 6) should be adjusted to obtain the required value of I_C.

From the preceding discussion it can be seen that the International Rectifier 1N1510 (6.8 V) Zener diode would be substituted for Z1, while a 1N1512 (10 V) could be used for Z2.

RF Circuits

The Zener diode may be used in frequency modulation equipment for both automatic gain control and for limiting.

A single diode connected across the coil and voltage decoupling resistor (Fig. 7) will conduct when the signal level exceeds V_z plus the voltage across the resistor. In addition to supplying a degree of limiting, diode conduction

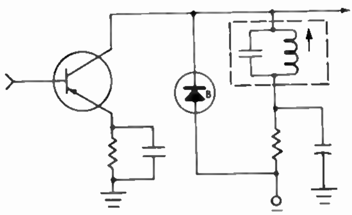


Fig. 7. Glass subminiature Zener diodes provide excellent limiting characteristics in transistorized FM receivers.

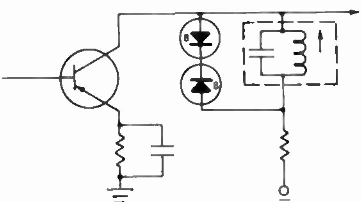


Fig. 8. A double anode version of the FM limiter. It should be noted that the diode junction capacity (10-20 μfd) will cause detuning, and is a circuit consideration.

places a low impedance in parallel with the coil. This effectively reduces the circuit Q and maintains the signal level in addition to avoiding overload in succeeding stages.

The customary system of limiting in FM equipment involves operating one or more stages at low voltage to produce saturation. A more efficient system incorporates the Zener diode, and allows the stage to operate at full gain. The circuit utilizes two Zener diodes connected in a "back-to-back" configuration as shown in Fig. 8. As soon as the incoming signal exceeds V_z the diodes conduct, effectively clipping the signal and reducing circuit Q.

Amateur Radio Circuits

The human voice contains many low energy peaks, and although they determine modulation percentage to a large degree, the average level is often very low. The effectiveness, or "talk power" of a radiotelephone transmitter can be greatly increased by using a peak or speech clipper to raise the average modulation level. The energy content is increased, but the 100% modulation point is not exceeded.

The characteristics of the Zener diode make them well suited as speech clippers in amateur radiotelephone transmitters. Such a circuit is shown in Fig. 9. The speech amplifier develops approximately 30 Volts of audio from a dynamic microphone. A small

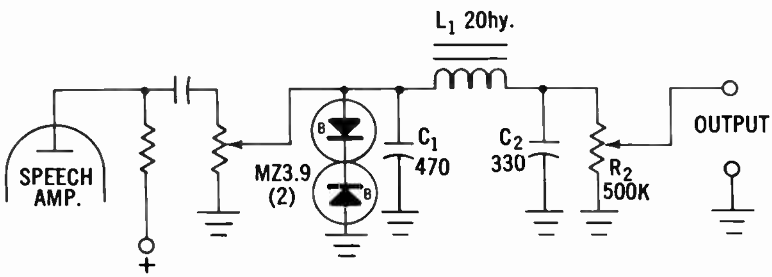


Fig. 9. A Zener diode speech clipper, useful in conjunction with amplitude modulated transmitters.

portion of this signal is applied to a pair of "back-to-back" Zener diodes, which clip any signal peak which exceeds 3.9 Volts. The harmonics produced by the clipping action are attenuated by the lowpass filter, consisting of L1, C1, and C2. The desired amount of audio signal is fed to the modulator by adjustment of R2.

The clipper need not be removed from the circuit to eliminate its effect on the voice energy. Simply turn R2 to its maximum clockwise position and control the audio level with R1. The amplitude should not exceed the clipping level and the audio will be unaffected by the Zener diode. The desired amount of clipping may be introduced by increasing the setting of R1, while reducing potentiometer R2.

A Zener diode may be utilized to switch a simple, tubeless, voice operated relay as in Fig. 10. The sensitivity control (R1) is set just below the point relay pull-in occurs. When the transmitter is modulated, audio peaks exceed V_z, energizing the relay. Capacitor C1 determines the dropout time of the relay. Larger capacitors will cause the relay to hold for a longer time. The relay contacts may be used to control the transmitter power circuits.

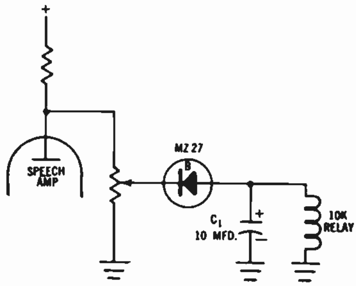
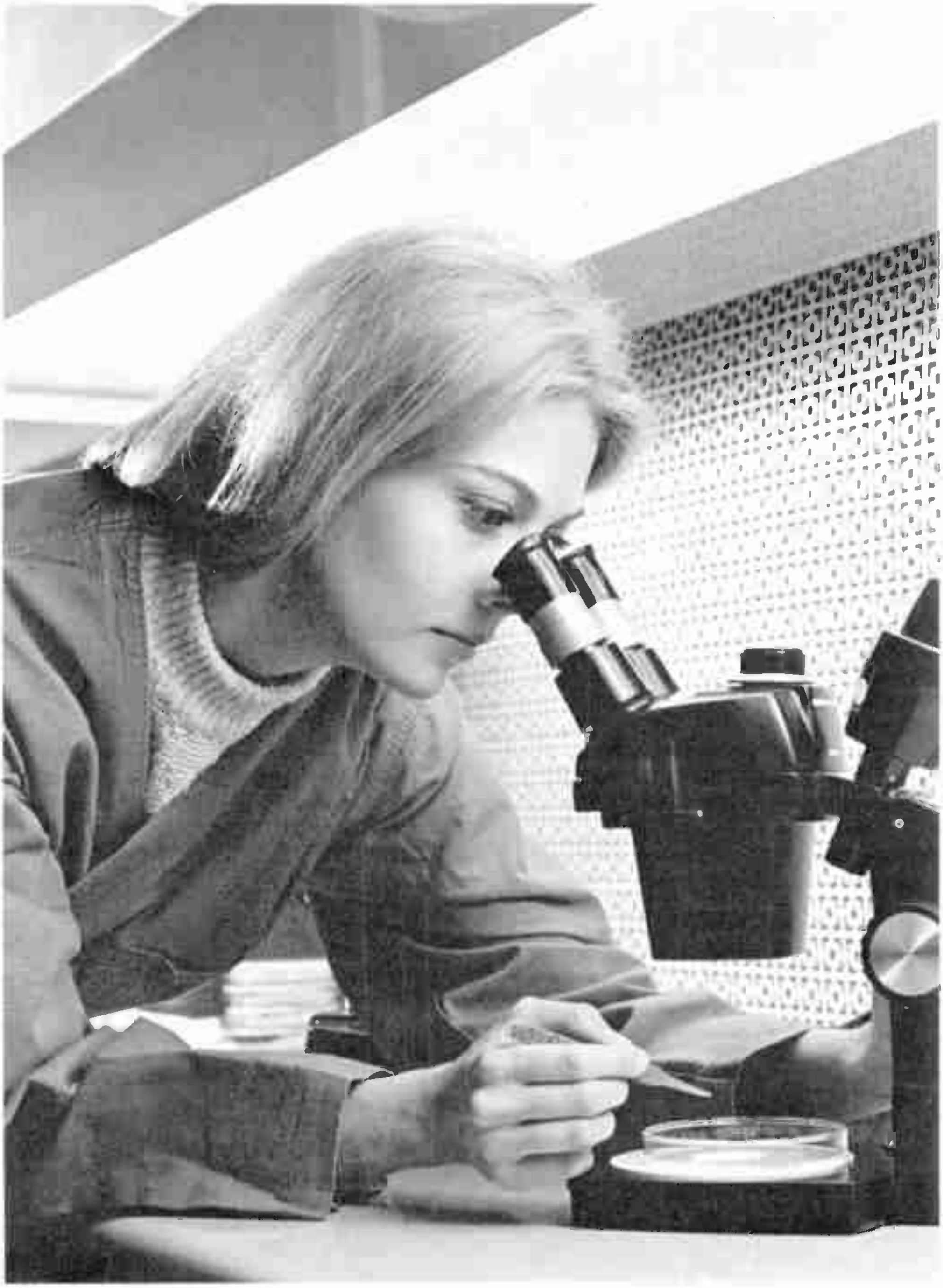


Fig. 10. A simple voice-operated relay circuit for single sideband equipment.



Careful quality control allows International Rectifier to offer a lifetime guarantee with each Zener diode.

Computer and Instrumentation Applications

The Switching Characteristic of A Zener Diode May Be Applied to Computers, Sorting Devices and Meter Scale Expansion or Compression

The Silicon Zener diode has been proven to be a tremendously useful design tool since its introduction. Over the years countless applications have been found to utilize the diode's unique characteristics.

In many ways the Zener diode resembles a switch and therefore one may find several logical applications in the computer and instrumentation fields.

Computers

The usual diode circuit configuration employs a slightly reverse biased germanium or silicon diode. The driving pulse, which represents information, initiates conduction. The diode may also be biased for forward conduction and the information bit used to cease conduction for the short interval. In either case the diode is cycled in and out of the forward conduction region.

This type of diode operation is not well suited to high speed digital computations, for it must be limited to data rates below 2.5 mHz, due to storage of minority carriers. It has been shown that switching times in excess of 0.1 microsecond are not practical with this system. When the information bit ceases, stored minority carriers diffuse back to the terminals causing a reverse transient and slow recovery time.

Zener Diode Computer Switch

These switching problems may be minimized by employing a Zener diode and switching about the avalanche point. When the diode is in an ava-

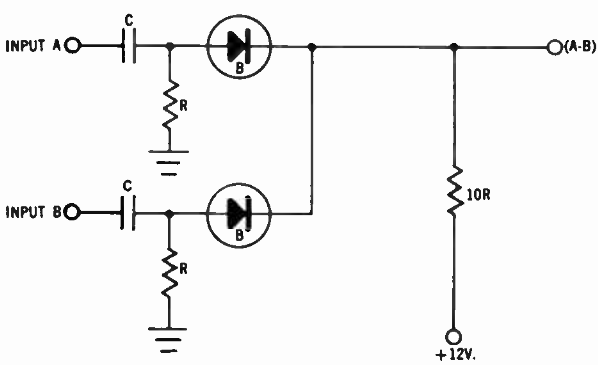


Fig. 1. Typical "AND" computer circuit employing Zener diodes to switch about the avalanche point.

lanche condition, current flow occurs due to majority carriers, and the condition created by storage of minority carriers does not exist. A strong electric field is present at all times in the region of the junction. Thus, a theoretical switching time equal to the relaxation time of the material could be obtained. In the case of silicon, this would increase the theoretical switching time to 10^{-9} seconds, or one nanosecond.

A typical "AND" circuit, using two 6.8 Volt Zener diodes, is shown in Fig. 1. The positive 12 Volt bias supply is applied to both cathodes through the common load resistor, maintaining the diodes in an avalanche condition. If a positive pulse is applied to input A, as an example, diode A will be gated out of the avalanche region. However, due to the low impedance of shunt connected diode B, the output (point A•B) remains clamped at approximately 6.8 Volts. The same conditions exist when a positive pulse is applied to input B. However, when an information bit arrives at terminals A & B simultaneously, both diodes are gated out of the avalanche region and the output potential rises to the supply voltage during this interval. This produces a positive coincidence pulse. A negative pulse, or transient, does not affect the circuit.

A logical "OR" computer circuit may be devised using the same technique, as shown in Fig. 2. In this circuit the diodes and the bias supply polarity are reversed, but positive information bits are retained. With this configuration either pulse will register in the output. When a positive bit is applied to input A, the diode is driven further into the avalanche region while the other diode switches from the avalanche to non-conducting state, producing a pulse in the output. Identical conditions occur when a bit is applied to input B.

There are several other circuit configurations that will no doubt occur to the reader. As an example, any number of diodes may be parallel connected for a variety of input gates. There are many applications for Zener diodes in high-speed switching applications when the concept of gating around the avalanche point is considered.

Sorting

The nature of Zener diodes also suggests an application in the realm of voltage controlled sorting. Several Zener diodes may be connected as shown in Fig. 3. As the input voltage increases, the relays will be progressively energized. The contacts may be arranged to open chutes and illuminate indicator lamps for rapid sorting.

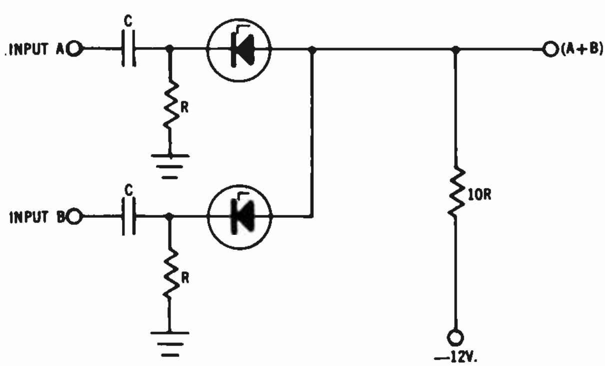


Fig. 2. Typical "OR" computer circuit employing Zener diodes to switch about the avalanche point.

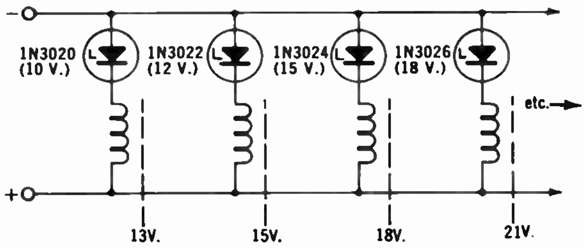


Fig. 3. A voltage sequence selector employing Zener diodes. Such a device would be useful for sorting electrically.

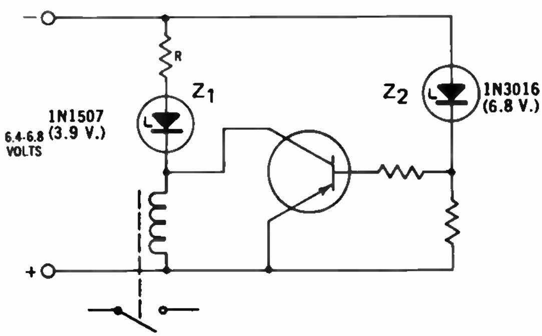


Fig. 4. A voltage selective device. The "selectivity" of the circuit may be improved as described in the text.

A voltage selective relay is shown in Fig. 4. Such a system would be useful for rapid "go - no go" sorting of batteries, Zener diodes or other voltage sources. It might also be used for voltage adjustment of power supplies in applications where delicate meters could be damaged due to rough handling.

If sufficient voltage is applied to the input, diode Z1 will conduct and energize the relay. In this example a 3.9 Volt diode (Z1) is used in conjunction with a relay set to pull in at 2.5 Volts. Thus a 6.4 Volt input signal will energize the relay. If the voltage exceeds this level, diode Z2 will conduct and apply forward bias to the transistor. This in turn reduces the collector-emitter junction resistance which de-energizes the relay. Thus the relay pulls in at 6.4 Volts and drops out at 6.8 Volts approximately.

The voltage "bandwidth" may be reduced by increasing the size of resistor "R". As this resistance is made larger, the voltage required to trip the relay increases, bringing the pull-in point closer to the conduction voltage of Z2. By careful selection of components it should be possible to resolve approximately 0.1 Volts.

The system may be used at higher voltages by employing other diode-relay combinations.

A system with resolving power in

excess of 0.1 Volts may be devised by using Zener diodes to provide a reference for transistor switches. Such a system of quantitizing would be useful for analog to decimal converters in statistical sorting applications.

Expanded Scale Instrumentation

The circuit shown in Fig. 5 illustrates an application of Zener diodes in expanded scale instrumentation. With this configuration a compressed lower scale and expanded upper scale is achieved, both of which may be calibrated. The relative areas and sensitivity of the scale will be determined by the component values selected. This circuit offers greater accuracy over the upper portion of the instrument scale, while still allowing rough measurements to be made in the lower scale range.

Fig. 6 illustrates an expanded scale with the lower scale portion completely suppressed. In this circuit, calibration may be achieved through adjustment of the meter needle below the low end of the scale.

The opposite effect of that shown in Fig. 6 may be accomplished as shown in Fig. 7. This circuit may be used to provide compressed indication of voltages above the critical ranges. The configuration may also be used to provide meter protection as described in Chapter 9, "Protection of Circuit Components with Zener Diodes."

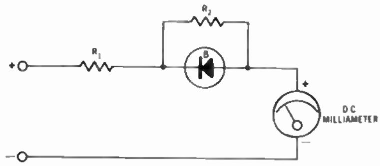


Fig. 5. A meter expander which partially suppresses the low end of the scale. Resistor R1 is used for upper scale calibration, while R2 is adjusted for the desired amount of low end suppression.

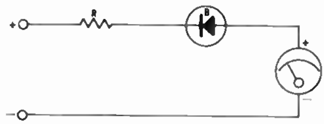


Fig. 6. A meter expander which completely suppresses the low end of the scale.

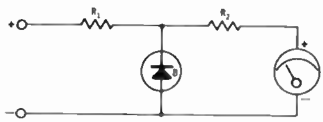


Fig. 7. A meter expander which completely suppresses the high end of the scale.

Protection of Circuit Components with Zener Diodes

The Ruggedness of The Zener Diode Permits It To Absorb Overloads And Protect Delicate Circuitry

The Zener diode has the ability to pass current when a specified breakdown voltage is exceeded, and therefore may be used to protect transistors, meters, or any device which may be damaged by excessive voltage.

Particularly critical in this regard are transistors. Like a diode, they too have an avalanche breakdown voltage which is somewhat higher than the specification V_{CE}. Unlike the Zener diode they cannot be cycled in and out of the avalanche region. Exceeding V_{CE} often results in the destruction of the transistor due to a total breakdown of the junction known as "punch-through."

Power Converters

The transistorized power converter provides an excellent illustration. These devices are quite popular in airborne electronic units for DC to AC and DC to DC conversion.

The aircraft electrical system is a notorious source of transients that may exceed V_{CE}. If these high voltage pulses are allowed to reach the power converter, they will appear on the transistors, for the low inductance of the transformer primary is not sufficient to provide a significant amount of attenuation. Even if the transient is less than V_{CE} it may occur simultaneously with a negative half-cycle, thereby creating a total collector potential which is destructive.

Fig. 1 shows a typical transistor power converter. When transistor Q1 is conducting, the other transistor will be cut off. However, due to auto-transformer action the collector-emitter potential of Q2 will be two-times V_{in}. Thus for a 12 Volt system, the collector to emitter potential would be 24 Volts, neglecting internally generated transients on the leading edge of the switching cycle.

If transistors with a 30 Volt V_{CE} were used in this example, a transient in the electrical system which exceeded 6 Volts could conceivably damage one or both transistors.

One solution is to employ transistors with 80 or even 100 Volt V_{CE} ratings. It is also possible to series connect transistors across each half of the switching winding. However, either solution is more expensive than the use of two Zener protection diodes, as shown in

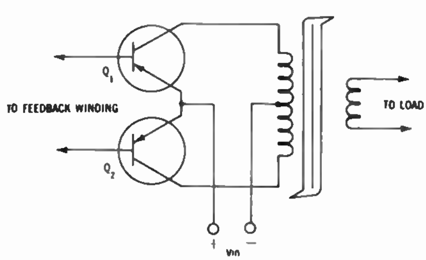


Fig. 1. The switching circuitry for a typical transistor power converter.

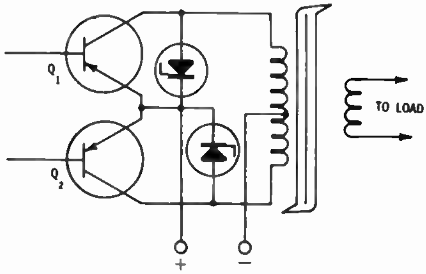


Fig. 2. A power converter modified for transient protection, with Zener diodes.

Fig. 2. These diodes are selected to have a V_Z which is slightly higher than two times maximum V_{in} and well below the maximum V_{CE} ratings for the transistors. Any transient, whether of internal or external origin, which exceeds the diode avalanche voltage will initiate conduction. This effectively clips or limits the transient well below the point where it could damage the equipment.

Audio Amplifiers

Class B servo and audio amplifiers are particularly susceptible to transient destruction. Should a pulse appear on the supply line (Fig. 3), it will pass unhindered to the transistor collectors due to the low primary inductance of transformer T2. The transient will cause an increase in the collector to base leakage since it flows through the base-emitter junction. Thus, the pulse is amplified by the transistor Beta and creates a sudden increase in collector current, almost simultaneously with the arrival of the transient. All design limits for the output stage are exceeded and transistor "punch-through" is usually the result.

The transient energy appearing at the base can destroy the base-emitter junction. Shunt resistors R1 and R2, in Fig. 3, will minimize this possibility in addition to reducing the amplitude of the pulse transformed to the collector circuit of Q1, through T1.

Complete transient protection may be obtained by incorporating Zener diodes between each collector and the common emitter circuit. As before, these diodes are selected with a V_Z which is somewhat greater than two times the supply voltage. Connected in this manner, the diodes limit destructive transients before they can damage the components.

The driver transistor is protected due to the resistance of R3, the primary

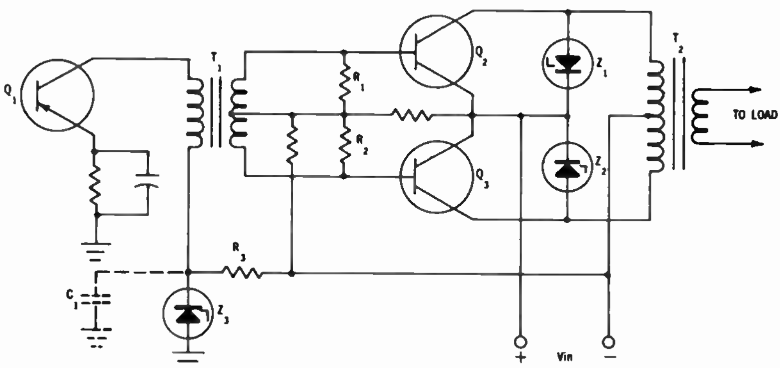
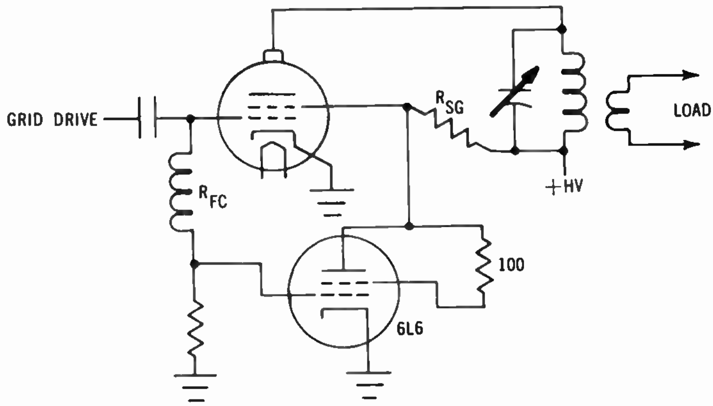


Fig. 3. A Class B audio or servo amplifier protected by Zener diodes. Even high voltage transients will not damage this amplifier.

Fig. 4. Class C rf power amplifier protected by a "clamp" tube.



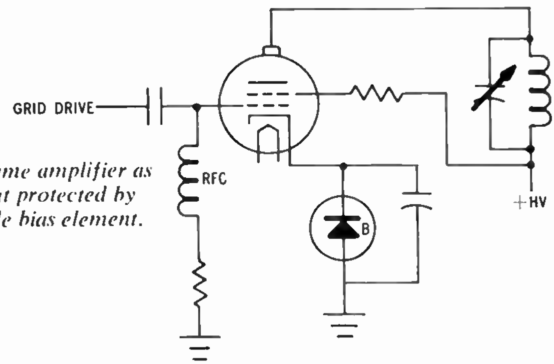


Fig. 5. The same amplifier as in Fig. 4, but protected by a Zener diode bias element.

of T1, and the low reactance of C1. Substituting a Zener diode (Z3) in the position normally occupied by the electrolytic capacitor will provide absolute protection in addition to greater decoupling (and improved low frequency response) due to the constant low impedance of Z3. The higher stabilized collector voltage for Q1 will allow less degeneration and consequently higher stage gain.

It has been shown[1] that this system affords complete protection even when the transistors are pulsed with 300 Volt transients, having a duration of one microsecond.

RF Amplifiers

The self biased Class C rf amplifier tube may be damaged if the rf drive energy should fail. For protection it is customary to connect a small pentode between screen grid and ground as shown in Fig. 4, with bias for this tube supplied from the potential which appears across the grid leak resistor. Should the drive fail, there is no bias for the "clamp" tube and it conducts heavily. This, of course, reduces the screen potential and plate current of the stage, insuring that its dissipation is not exceeded.

A less complicated, and in many cases less expensive, system is shown in Fig. 5. A Zener diode in the cathode circuit establishes the minimum bias consistent with maximum dissipation in the stage. In operation the grid bias is the sum of V_Z and the potential across R due to grid current flow. If the drive should cease, sufficient bias to protect the amplifier tube is provided by the

Zener diode. Since the protection device has no filament to burn out, reliability is greatly increased.

The correct diode for the application may be determined by consulting the tube characteristics to determine the maximum current which will not exceed the dissipation at a given plate potential:

$$I_p = \frac{P_d}{E_p}$$

where; I_p = maximum plate current
P_d = plate dissipation
E_p = plate voltage

When I_p is known, the correct bias for the current may be determined from the characteristic curves. The diode would be selected as near this value as possible. The diode wattage may be easily determined when voltage and current in the diode circuit are known. The next largest size should be selected.

Arc Protection

Whenever electric power to inductive devices must be controlled by contacts which initiate and interrupt the flow of current, the problem of protecting the contacts from erosion becomes important. This protection is more important in DC circuits, although it is often necessary in AC circuits, and commands greater attention at higher supply voltages.

Particular consideration must be given to the problem in transistor circuits. It has been shown that excessive potentials can destroy the transistor junctions. This discussion considers the deterioration of switch contacts, but in general the comments are equally applicable to transistor switches.

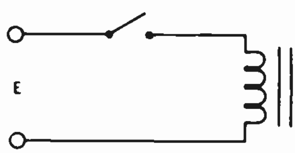


Fig. 6. Inductive circuit requiring arc suppression.

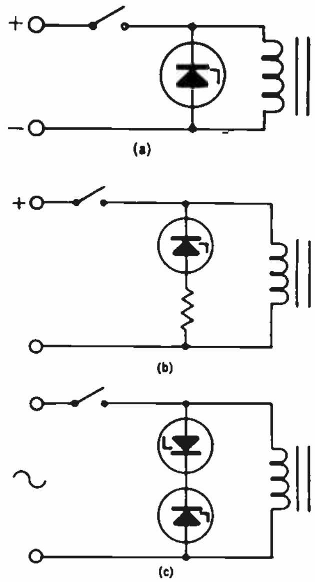


Fig. 7. Three common arc suppression circuits.

When the switch in the inductive circuit (Fig. 6) is opened, the magnetic field in the coil collapses and a voltage is generated equal to L di/dt, where L is the coil inductance and di/dt is the time rate of change of decay current. This voltage is frequently many times the supply voltage, enough to maintain an arc across the switch contacts. The arc may be a glow discharge, or even a small metallic bridge which becomes hot enough to vaporize a small portion of the contact metal, Repeated arcing causes erosion, pitting, and general deterioration of the contacts and results in high contact resistance and increased maintenance.

Several methods of contact protection have been employed with varying results. These methods have used various components such as arc blowout coils, permanent magnets, fixed resistors, capacitors and semiconductor devices. Zener diodes have proven to be most effective and are able to overcome some of the inherent disadvantages of other methods as noted below:

1. Blow-out coils or magnets, while

reducing arcing, tend to increase the voltage generated by the coils, sometimes to values that will puncture the coil or circuit insulation. Zener diodes both reduce the arcing and control peak voltages.

2. Fixed resistors can reduce the arcing and control peak coil voltages, but consume power while the coil is energized and sometimes increase drop-out time excessively.

3. Capacitors also provide arc reduction, but sometimes draw excessive charging currents, which cause contact heating and may also increase drop-out time.

Proper design and application of semiconductor arc suppressors can minimize or eliminate the above disadvantages, and may be applied to many types of magnetic devices.

To obtain maximum contact protection concurrent with optimum circuit operation, it is necessary to select the proper suppressor circuit (Fig. 7) and to specify the proper type and rating of the semiconductor device. On all AC circuits, the suppressor arrangement of Fig. 7c is necessary, since the suppressor must block voltages of both polarities.

In some applications, a delay in drop-out time of a relay or contactor may be an advantage. The duration of this delay can be controlled by selecting the correct suppressor, or made adjustable by using the circuit of Fig. 7b, with a variable resistor.

The Zener diode is an ideal arc or transient suppression device. For AC circuits, two diodes can be connected as shown in Fig. 7c. This back-to-back configuration can be used with either lead or stud mounted diodes.

A detailed discussion of semiconductor arc suppression techniques in specific magnetic circuits is given in Application Notes available from International Rectifier.

Meter Protection

D'Arsonval meter movements, due to their delicate construction, are easily damaged with even moderate overloads. Although the moving coil may not open, the sudden application of excessive current can cause the pointer to strike the end-stop violently. If the applied force is sufficient, the needle will be permanently bent.

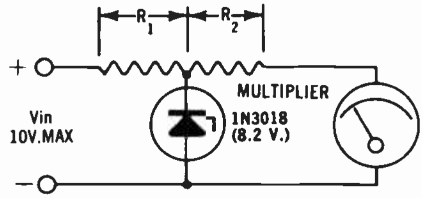


Fig. 8. A Zener diode meter protection circuit.

The Zener diode, with its ability to limit at a specific voltage, may be used to provide complete protection against physical damage due to overload.

A suitable circuit is shown in Fig. 8. The diode must be connected to a point on the voltage divider (or multiplier resistor) which is near V_Z.

For the purpose of illustration assume a protected 10 Volt DC meter is desired. To minimize circuit loading a basic 100 microampere movement with a 1,000 Ohm coil is to be used. A Zener diode rated in excess of 7 Volts and below 10 Volts should be used, for they exhibit a sharper "knee" and therefore do not affect full scale linearity. Therefore, an 8.2 Volt diode has been chosen for this application. The total resistance of the multiplier (R1 + R2) may be determined by using the formula:

$$R1 + R2 = \frac{E_{in}}{I_m} - R_m$$

where: E_{in} = maximum voltage to be measured
I_m = full-scale meter sensitivity
R_m = meter coil resistance

Thus, in the example, a 99,000 Ohm multiplier resistor would be used. The ratio of R1 to R2, which determines the voltage available for the Zener diode, is given by the formula:

$$R2 = \frac{R1 + R2 \times V_Z}{E_{in}}$$

Resistor R2 therefore is 81.18K, and of course, resistor R1 is the difference, or 17.82K.

For extreme accuracy, the tiny diode leakage current must be taken into consideration. The leakage current will cause a small voltage drop across R1, and will cause a slight decrease in the meter current. This effect will be least when a diode close to V_{in} is used, making R1 a small portion of the total resistance.

If full scale linearity proves to be a problem, the diode should conduct 5% above the maximum input voltage. Thus in the formula for R2, a figure of 10.5 Volts would be used for V_{in}. This will allow the pointer to deflect beyond full scale but it will not have sufficient inertia to damage it.

Recorders

Although not a meter in the strict sense, recorders and x-y plotters may be protected by employing the above technique. These devices use DC amplifiers and excessive voltage applied to the writing pens is almost an everyday occurrence, particularly in the hands of an unskilled operator. The new models usually incorporate some form of protection. Earlier recorders may be protected by determining the plate voltage of a low-level stage, when overload occurs. A Zener diode, slightly higher then this potential, may be connected between plate and cathode to limit pen movement.

REFERENCE

1. B. B. Daien, "Protect Transistors Against Destructive Transients," Electronic Design, Nov. 1959.

For further information on any of the International Rectifier products listed in this book, write directly to the factory or contact your local IR Distributor or local IR Field Office. All standard items listed are carried in stock by International Rectifier Industrial Distributors located throughout the United States.

INTERNATIONAL RECTIFIER IOR

Semiconductor Division, 233 Kansas Street, El Segundo, California 90245, Phone (213) 678-6281
Field offices and distributors in major cities around the world

INTERNATIONAL RECTIFIER

Semiconductor Division, 233 Kansas Street, El Segundo, California 90245, Phone (213) 678-6281
Field offices and distributors in major cities around the world